



<http://dx.doi.org/10.35596/1729-7648-2025-31-4-5-14>

УДК 378.147:004.8

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ СОЗДАНИЯ УЧЕБНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ «БАЗЫ ДАННЫХ»

О. Д. АСЕНЧИК

*Гомельский государственный технический университет имени П. О. Сухого
(Гомель, Республика Беларусь)*

Аннотация. Рассмотрена проблема использования больших языковых моделей (LLM) для создания учебных материалов. Предложена и верифицирована практическая методика с целью генерации качественного учебного контента для конкретной дисциплины «Базы данных». Представлена многоступенчатая методика, в которой одна LLM генерирует контент, а вторая, независимая «рассуждающая» модель, верифицирует его на предмет качества и корректности. Для проверки сгенерированных материалов на отсутствие фактических ошибок применялись метод сравнения с авторитетным источником и модифицированный алгоритм «цепочки верификаций». Результаты подтверждают, что данный подход при использовании современных производительных LLM (таких как DeepSeek, Gemini) позволяет создавать высококачественные учебные тексты с низкой вероятностью появления галлюцинаций. Методика способна значительно ускорить разработку надежных учебно-методических материалов и может быть оптимизирована путем сокращения числа итераций при высоком качестве первоначального ответа.

Ключевые слова: высшее образование, учебные материалы, большие языковые модели (LLM), искусственный интеллект, базы данных.

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Для цитирования. Асенчик, О. Д. Использование больших языковых моделей для создания учебных материалов по дисциплине «Базы данных» / О. Д. Асенчик // Цифровая трансформация. 2025. Т. 31, № 4. С. 5–14. <http://dx.doi.org/10.35596/1729-7648-2025-31-4-5-14>.

USING LARGE LANGUAGE MODELS TO CREATE EDUCATIONAL MATERIALS FOR THE “DATABASES” DISCIPLINE

ALEH D. ASENCHIK

Sukhoi State Technical University of Gomel (Gomel, Republic of Belarus)

Abstract. This paper examines the use of large language models (LLM) for creating educational materials. A practical methodology for generating high-quality educational content for the specific discipline of “Databases” is proposed and verified. A multi-stage methodology is presented, in which one LLM generates content, and a second, independent “reasoning” model verifies its quality and correctness. A comparison method with an authoritative source and a modified “verification chain” algorithm was used to check the generated materials for factual errors. The results confirm that this approach, when used with modern, high-performance LLMs (such as DeepSeek and Gemini), enables the creation of high-quality educational texts with a low probability of hallucinations. The methodology can significantly accelerate the development of reliable educational materials and can be optimized by reducing the number of iterations while maintaining a high-quality initial response.

Keywords: higher education, educational materials, large language models (LLM), artificial intelligence, databases.

Conflict of interests. The author declares that there is no conflict of interests.

For citation. Asenchik A. D. (2025) Using Large Language Models to Create Educational Materials for the “Databases” Discipline. *Digital Transformation*. 31 (4), 5–14. <http://dx.doi.org/10.35596/1729-7648-2025-31-4-5-14> (in Russian).

Введение

Создание учебных материалов – сложный и трудоемкий процесс, требующий значительных затрат времени и глубоких экспертных знаний. Появление больших языковых моделей (LLM), являющихся ведущим компонентом технологий искусственного интеллекта (ИИ), открывает новые горизонты для преобразования процесса создания образовательного контента. Эти модели, обученные на масштабных корпусах самых разных текстов, включая обучающие материалы и учебники, образцы программного кода, демонстрируют способность генерировать осмысленные, структурированные и адаптируемые к текущему содержанию тексты, что открывает новые перспективы для их использования в педагогической практике [1, 2]. Способность LLM понимать естественный язык позволяет рассчитывать на автоматизацию создания образовательного контента, существенно снижая трудозатраты на его подготовку. Также их способность генерировать текст на многих языках делает более доступной организацию обучения студентов из разных стран.

Интеграция LLM в образование может изменить практику преподавания и обучения [3]. В настоящее время LLM применяются в самых разных образовательных контекстах, включая персонализированное обучение, интеллектуальные системы обучения, генерацию контента и оценку результатов обучающих мероприятий [1, 3, 4]. Использование LLM для создания учебного контента является развивающимся направлением [5]. Исследования и публикации описывают различные методологии и результаты использования LLM для генерации образовательных материалов, включая планы уроков, викторины, объяснения кода, тексты лекций [5, 6]. Процессы генерации контента включают использование LLM в качестве самостоятельных инструментов, а также их интеграцию с внешними базами знаний и (или) экспертами в предметной области непосредственно в процесс [6]. Отмечаются такие результаты, как повышение эффективности, увеличение вовлеченности студентов и улучшение результатов обучения [7]. Однако в некоторых исследованиях указывается на необходимость тщательной оценки и проверки сгенерированного контента, в том числе для обеспечения глубины, надежности и правильности контента, сгенерированного LLM [7, 8].

Методика генерации и верификации материалов

Анализ проблем и ограничений LLM при генерации текстов выявляет ряд серьезных ограничений. Большие языковые модели полагаются на статические данные обучения, которые могут быть устаревшими или не содержать специфической информации [9]. В результате LLM могут выдавать некорректные сведения, известные как галлюцинации, особенно если они не знают правильного ответа или если информация в их обучающем наборе потеряла актуальность. Более того, LLM не имеют доступа к информации о событиях или данных, которые появились после завершения их обучения. Эти проблемы могут негативно сказаться на качестве, актуальности и корректности генерируемого материала.

Устоявшимся способом повышения точности ответов, снижения риска галлюцинаций является использование технологии генерации с дополненным поиском (Retrieval Augmented Generation, RAG) [10]. Эта технология подразумевает наличие специально подбираемого набора эталонных источников, покрывающих предметную область, в которой будут задаваться вопросы, а также построение специфического программного решения и экспериментального подбора оптимальных параметров, таких как размер фрагментов текста, на которые разбиваются документы, методы поиска наиболее релевантной информации и др. Использование RAG повышает сложность системы, требует тщательной настройки и увеличивает затраты на обслуживание и поддержание в актуальном состоянии. Учитывая постоянно возрастающие возможности современных реализаций LLM, актуальными являются поиск и апробация более простых в реализации методик, основанных на использовании собственно возможностей LLM, для получения релевантных предметной области учебных материалов.

В статье предлагается, описывается и верифицируется простая методика создания с использованием LLM учебных материалов в виде текста теоретической части курса для изучения учебной дисциплины на примере конкретной дисциплины «Базы данных». Выбор данной дисциплины в качестве объекта обусловлен тем, что она классическая; занимает важное место в подготовке студентов, изучающих информационные технологии и программирование, обеспечивая фундаментальные знания и навыки, необходимые для разработки современных информационных систем. Для построения конкретного учебного курса по этой дисциплине имеется обширный корпус авторитетных учебников и пособий, справочных материалов или официальной документации, расположенных в интернете и проиндексированных поисковыми системами и LLM при их обучении. Учебные материалы по дисциплине «Базы данных», как правило, имеют четко определенную и устоявшуюся структуру, что существенно облегчает их генерацию с использованием LLM. Все это обеспечивает надежную основу для создания точного и проверенного контента.

Основными этапами реализации предлагаемой методики являются: разработка наборов вопросов, конкретизирующих и детализирующих изучение тем учебной дисциплины; выбор набора реализаций LLM и предварительная оценка (самооценка) возможностей этих LLM ответить на данные вопросы без галлюцинаций; многоступенчатая генерация ответов на каждый из полученных вопросов с предварительной верификацией получаемых ответов; оценочная финальная верификация текстов полученных ответов.

При создании теоретической части курса для изучения учебной дисциплины «Базы данных» было подготовлено 64 вопроса, охватывающих как классические темы (реляционная модель данных, нормализация таблиц, язык структурированных запросов SQL, методология проектирования баз данных), так и относительно современные подходы (NoSQL базы данных, BIG DATA и обработка больших объемов данных, аналитическая обработка данных, искусственный интеллект и машинное обучение в базах данных). Ожидаемые ответы на вопросы должны содержать соответствующий текст, иногда код примеров на языке SQL, иногда математические выражения из теории реляционных баз данных.

В качестве используемых для проведения экспериментов LLM были выбраны DeepSeek V3 (далее – deepseek), DeepSeek R1 (далее – deepseek-reasoner), Llama 3.3 70B (далее – llama), Gemini 2.0 Flash (далее – gemini), Qwen QwQ (далее – qwen) и Gemma 3 1B (далее – gemma). Каждая из этих моделей¹ представляет значительное достижение в области искусственного интеллекта, разработанное ведущими компаниями и исследовательскими группами. 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^{2,3} с качеством, приемлемым для решения задач. Все модели, за исключением gemini, также доступны для локального развертывания и использования на компьютерах под управлением Windows, macOS и Linux без взимания оплаты, например, с применением фреймворка ollama.

Для получения текстов учебного пособия был использован доступ из программного кода на языке Python к моделям deepseek-reasoner, deepseek, gemini, llama, qwen посредством сетевых вызовов через API соответствующего внешнего провайдера или доступ к модели gemma, расположенной на локальном компьютере, с использованием фреймворка ollama. Язык формулирования запросов к LLM – русский, язык ответов – русский. С целью сделать ответы LLM более детерминированными и близкими к данным обучающего набора параметр «температура» для всех моделей устанавливался равным нулю.

Для проверки возможности LLM ответить на все вопросы ей отправлялся запрос следующего содержания: «Вы специалист и преподаватель по базам данных и системам управления базами данных, информационным системам и разработке программных приложений баз данных, создающий учебное пособие для студентов университета, изучающих информационные технологии, по дисциплине «Базы данных». Вы исключительно хорошо владеете английским и русским языками, знаете содержание документации, связанной с дисциплиной, знакомы с содержимым

¹ Models [Electronic Resource] // Artificial Analysis. Mode of access: <https://artificialanalysis.ai/leaderboards/models>. Date of access: 23.06.2025.

² OpenRouter [Electronic Resource] // OpenRouter. Mode of access: <https://openrouter.ai/>. Date of access: 22.06.2025.

³ Rate Limits [Electronic Resource] // Google for Developers. Gemini API Documentation. Mode of access: <https://ai.google.dev/gemini-api/docs/rate-limits>. Date of access: 23.06.2025.

большого количества признанных учебных пособий. Ответьте, достаточно ли тех источников, на которых ты обучен, и достаточно ли они компетентные, чтобы дать развернутые, корректные и правильные ответы на каждый из прилагаемых вопросов в списке по дисциплине?»).

Подход, предполагающий генерацию ответов на все вопросы одновременно в рамках одного запроса к LLM, оказался неэффективным из-за недостаточного размера контекстного окна почти всех моделей для передачи больших объемов информации, большой нагрузки на ресурсы серверов с установленной LLM и неустойчивости сетевого соединения между сервером и клиентом, которое нужно было поддерживать. В результате ответы либо не удавалось получить, либо ответы на вопросы были слишком краткие. В процессе исследований автор применял методику ответа на каждый вопрос по отдельности⁴. Ответы LLM по каждому вопросу не увязывались с результатами ответов на предыдущие и последующие вопросы.

При генерации ответов применялся модифицированный алгоритм цепочки верификаций (Chain of Verification, CoVe) [11]. Алгоритм CoVe был разработан для того, чтобы дать языковым моделям возможность обдумывать свои ответы и исправлять ошибки, а также для минимизации фактических ошибок в ответах LLM. В основе методологии CoVe лежат принципы самокоррекции, декомпозиции и независимой верификации.

Оригинальный алгоритм CoVe состоит из четырех основных этапов: генерация базового ответа, который может содержать неточности; планирование верификаций путем использования запроса и базового ответа для генерации списка проверочных вопросов; выполнение верификаций путем ответа на каждый проверочный вопрос независимо для проверки на наличие несоответствий или ошибок по сравнению с исходным ответом; генерация итогового верифицированного ответа – на основе результатов верификации LLM генерирует модифицированный ответ, который включает исправления и устраняет выявленные несоответствия.

В рассматриваемом случае выполнение верификаций осуществлялось совместно с этапом планирования с использованием независимой LLM. Данное объединение возможно, когда в качестве такой LLM выбирается производительная современная рассуждающая модель, способная сама задавать необходимые вопросы. В исследовании в качестве LLM была выбрана *deepseek-reasoner* – рассуждающая языковая модель, разработанная для решения сложных задач через цепочки логических шагов. Эта модель сама разбивает сложные задачи на шаги, анализирует пробелы в данных, выявляет противоречия и обрабатывает неоднозначные запросы. Целями этапа оценочной финальной верификации полученных текстов были как проверка отсутствия галлюцинаций в ответах, так и дополнительная проверка полноты и качества ответов, полученных в результате описанной выше многоступенчатой схемы генерации.

При оценочной финальной верификации текстов полученных ответов на отсутствие галлюцинаций применялись метод сравнения с эталонным источником и модифицированный CoVe-метод. При оценочной финальной верификации ответов на отсутствие галлюцинаций полученный ответ сравнивался с информацией из эталонного источника. В качестве такого источника использовался современный учебник⁵. Содержание передаваемой LLM соответствующей инструкции для проверки: «Ты опытный специалист по базам данных, который выступает в качестве судьи. Есть вопрос и ответ на него, которые приводятся ниже. Оцени содержание ответа по шкале от 1 до 5. Критерии оценки: правильность утверждений и фактов, корректность использования терминов. Оценку проводи исходя из информации в контексте прилагаемого pdf-файла компетентного источника». В качестве LLM использовалась *gemini*, которая обладает большим контекстным окном для учета содержания объемного эталонного источника. Содержание упомянутого учебника, по экспертной оценке преподавателей Гомельского государственного технического университета имени П. О. Сухого, было полностью релевантно 37 вопросам из 64 (58 %), частично релевантно 15 вопросам (23 %) и полностью не релевантно 12 вопросам (19 %).

Для оценки отсутствия галлюцинаций в ответах на упомянутые выше вопросы из числа частично релевантных и нерелевантных использовался модифицированный CoVe-метод. Модифика-

⁴ Асенчик, Е. Ф. Использование большой языковой модели DeepSeek для создания учебных материалов по дисциплине «Латинский язык» / Е. Ф. Асенчик, О. Д. Асенчик // Современное образование: преемственность и непрерывность образовательной системы «школа – университет – предприятие»: сб. матер. XV Междунар. науч.-метод. конф., Гомель, 20–21 февр. 2025 г. Гомель: Гомельский гос. ун-т им. Ф. Скорины, 2025. С. 25–27.

⁵ Волк, В. К. Базы данных. Проектирование, программирование, управление и администрирование / В. К. Волк; 3-е изд., стер. СПб.: Лань, 2022.

ция заключалась в том, что на последнем этапе вместо генерации итогового верифицированного ответа давалась оценка точности полученного ответа. Содержание передаваемой LLM инструкции для проверки было следующим: «Вы эксперт по базам данных и системам управления базами данных, информационным системам и разработке программных приложений баз данных. Даны ВОПРОС, ОТВЕТ и ПРОВЕРЕННЫЙ_ИСТОЧНИК. Оцените по шкале от 1 до 10 точность ОТВЕТА на заданный ВОПРОС, используя информацию из проверенных источников (ПРОВЕРЕННЫЙ_ИСТОЧНИК). При выставлении оценки учитывайте, насколько информация в ответе на вопрос соответствует информации из проверенного источника».

Для выполнения верификаций каждый проверочный вопрос обрабатывался с использованием соответствующего API поискового сервиса Google, производящего поиск информации из источников в интернете (ПРОВЕРЕННЫЙ_ИСТОЧНИК). Число таких вопросов колебалось в диапазоне от 8 до 20 шт. в зависимости от длины проверяемого ответа. По этой причине данный метод верификации является относительно затратным и длительным, поскольку требует генерации большого количества запросов к внешним ресурсам, получения и обработки ответов.

При оценочной финальной верификации текстов полученных ответов на полноту и правильность использовалась независимая рассуждающая LLM – deepseek-reasoner. Содержание передаваемой LLM инструкции для проверки было следующим: «Вы эксперт по базам данных и системам управления базами данных, информационным системам и разработке программных приложений для них. Проверьте, насколько ответ соответствует вопросу. Перечислите основные подтемы вопроса. Насколько ответ раскрывает каждую подтему? Насколько информация и терминология в ответе соответствуют фактам и терминологии в компетентных источниках? Если в ответе присутствует программный код – проверьте его правильность. Оцените предоставленный ответ на вопрос на: правильность утверждений и фактов, корректность использования специальных терминов, полноту ответа, наличие примеров, качество и корректность кода, в случае если программный код приведен. Выставьте оценку ответу от 1 до 10. Напишите рекомендации, содержащие предложения по уточнениям, правкам или дополнениям, если это требуется».

Результаты исследований и их обсуждение

Тексты сгенерированных учебных материалов каждой моделью deepseek, llama, gemini, qwen, gemma можно найти в папке results в github-репозитории⁶. Результаты проведенной оценочной финальной верификации ответов на отсутствие галлюцинаций с использованием эталонного источника приведены в табл. 1.

Таблица 1. Результаты верификации ответов на отсутствие галлюцинаций с использованием эталонного источника

Table 1. Verification results of hallucination-free responses using a reference source

Название модели	Средняя оценка отсутствия галлюцинаций
gemini	4,62
llama	4,38
qwen	4,32
deepseek	4,30
gemma	2,43

Оценка отсутствия галлюцинаций проводилась с использованием 5-балльной шкалы. Высокая оценка соответствует ответу, полностью согласующемуся с содержимым эталонного источника. Из табл. 1 видно, что ответы, полученные с помощью легковесной локальной модели gemma, в целом содержат не релевантные вопросу ответы. Например, на вопросы «Унарные и бинарные операции реляционной алгебры. Характеристика и примеры операций выборки (селекции), проекции, деления, соединения», ответы на которые у всех моделей, кроме gemma, полностью согласовывались с эталонным источником, модель gemma предоставила неверный ответ. Полученное заключение от оценивающей модели выглядит так: «Ответ совершенно не соответствует поставленному вопросу и информации из предоставленного источника, игнорирует ключевые

⁶ IT-GSTU. db_textbook [Electronic Resource] // Gomel State Technical University. Mode of access: https://github.com/IT-GSTU/db_textbook. Date of access: 23.06.2025.

аспекты вопроса (унарные/бинарные операции, характеристика деления и проекции), смешивает понятия реляционной алгебры и SQL, использует некорректную терминологию и примеры. Большая часть ответа посвящена сравнению PA и SQL, что не требовалось». Непосредственная верификация ответа путем экспертной оценки также подтверждает это заключение. Кроме того, данная модель допускала в текстах ответов смешение русского и английского языков.

Остальные модели показали достаточно высокий средний уровень соответствия ответов информации в эталонном источнике, что подтверждается высокой средней оценкой. Зависимости процентной доли ответов при оценке на отсутствие галлюцинаций от величины этой оценки для различных моделей представлены на рис. 1. Процентная доля вычислялась от общего количества вопросов.

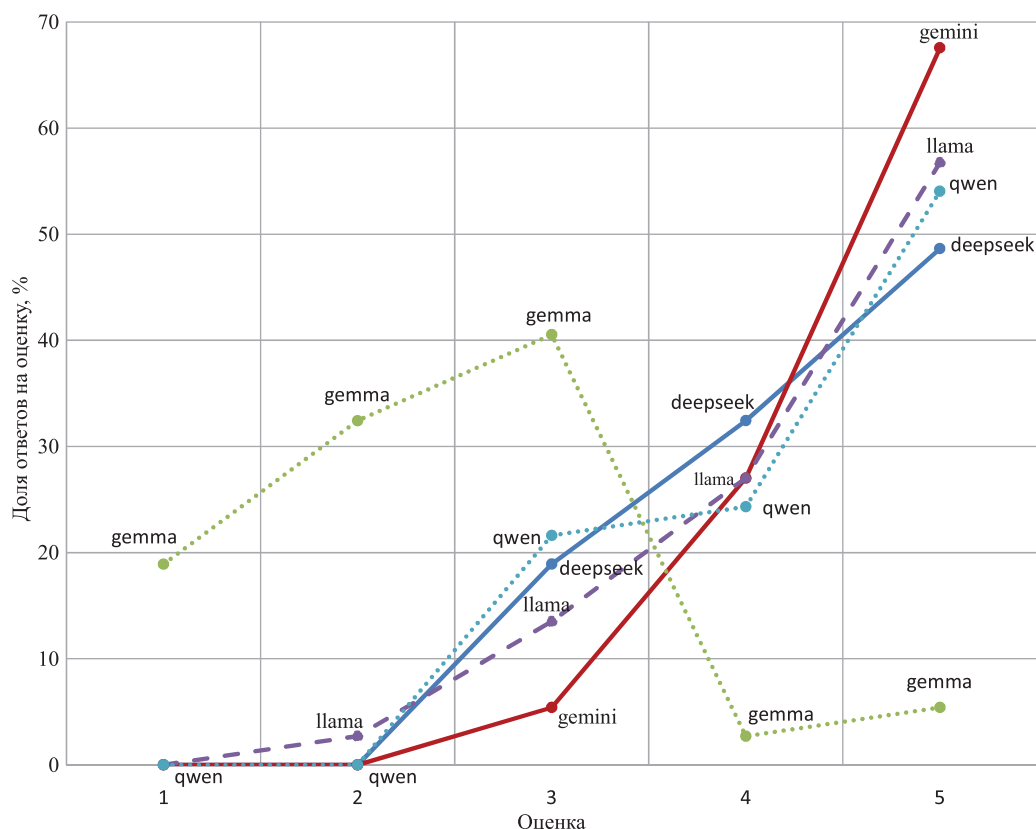


Рис. 1. Зависимость процентной доли ответов при оценке на отсутствие галлюцинаций от величины оценки для различных моделей

Fig. 1. Dependence of the percentage of responses in the assessment of the absence of hallucinations on the assessment value for different models

Данные на рис. 1 показывают, что все модели, за исключением gemma и llama, не галлюцинируют, и при ответах на отдельные вопросы их оценки составили 3 и более. Модель llama при ответе на один вопрос, касающийся нормальной формы Бойса – Кодда (BCNF), сгенерировала в целом достаточно полный и правильный ответ, но начальная часть его содержала ошибочное утверждение: «Нормальная форма Бойса – Кодда – это метод нормализации базы данных, который гарантирует, что каждая нетривиальная зависимость функциональности в таблице базы данных является полной зависимостью. Другими словами, каждый неключевой атрибут в таблице должен зависеть от всего первичного ключа, а не от части его». Данное утверждение не согласовывалось с содержимым эталонного источника, другими источниками и экспертным мнением. Этот ответ получил оценку 2. При этом следует отметить, что при проведении последующей повторной генерации ответа на данный вопрос такая ошибка не появлялась.

Непосредственный экспертный анализ отзывов рецензирующей модели при выставлении оценок 3 и 4 показал, что снижение оценки происходило не из-за наличия галлюцинаций в ответах, а из-за превышения информации в ответе относительно источника, присутствия деталей, не подтвержденных источником, но не противоречащих ему.

Ответы на вопросы, отнесенные к группе частично релевантных и нерелевантных эталонному источнику, на предмет отсутствия галлюцинаций были проверены модифицированным CoVe-методом. В силу его временной и вычислительной затратности оценки были проведены только для deepseek, gemini и qwen. Полученные оценки составили 10 и 9 по 10-балльной шкале, за исключением одной оценки 8 у модели qwen по вопросу об искусственном интеллекте и о машинном обучении в базах данных. Средняя оценка получилась равной 9,8, т. е. 5 по 5-балльной шкале, что является подтверждением низкой вероятности галлюцинаций в ответах на вопросы этих категорий. Таким образом, можно сделать вывод, что при генерации учебных материалов с использованием описанной выше методики по выбранной дисциплине «Базы данных» моделями gemini, deepseek и qwen вероятность наличия галлюцинаций низкая.

В табл. 2 с использованием deepseek-reasoner приведены результаты оценочной (по 10-балльной шкале) промежуточной и финальной верификации текстов полученных ответов на полноту и правильность. Результаты промежуточной верификации получались путем оценки deepseek-reasoner первичного ответа с той же инструкцией, что и при финальной верификации.

Таблица 2. Результаты оценочной промежуточной и финальной верификации ответов на полноту и правильность

Table 2. Evaluative results from the intermediate and final verification of answer completeness and accuracy

Название модели	Средняя оценка верификации	
	промежуточной	финальной
gemini	8,38	8,59
deepseek	8,34	8,55
qwen	8,14	8,35
llama	7,31	7,81
gemma	5,30	5,70

Данные табл. 2 демонстрируют достаточно высокие оценки по всем моделям, за исключением gemma и llama. При этом видно, что для всех моделей средняя оценка финального ответа после этапа предварительной оценки первичного ответа и учета его замечаний увеличивается (столбцы 2 и 3).

Зависимости процентной доли ответов при оценке на полноту и правильность от величины оценки для различных моделей приведены на рис. 2. Процентная доля ответов вычислялась от общего количества вопросов. В силу представленных в табл. 1, 2 и на рис. 1 результатов данные по модели gemma, как не подходящей для выбранной задачи, в статье не приводятся.

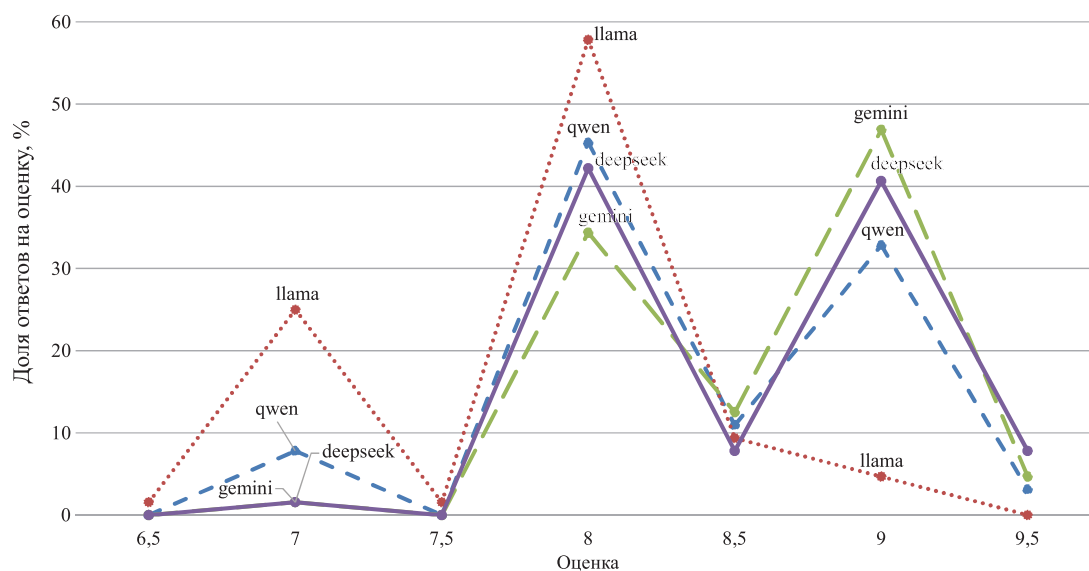


Рис. 2. Зависимость процентной доли ответов при оценке на полноту и правильность от величины оценки для различных моделей

Fig. 2. Dependence of the percentage of responses in the assessment of completeness and correctness on the assessment value for different models

Из табл. 2 и рис. 2 видно, что модели gemini и deepseek при оценке их ответов третьей верифицирующей моделью демонстрируют наибольшую полноту и корректность, а модель llama – наименьшую. Примерно 27 % ответов llama были оценены на 7 баллов. Непосредственный анализ содержания отзывов на ответы llama показал, что в целом они позитивные, но при этом были выявлены некоторые неверные команды, упрощенные примеры, смешение синтаксисов различных систем управления базами данных, несогласованность утверждений.

В табл. 3 представлены данные о среднем увеличении оценки на полноту и правильность ответов в разрезе моделей и величин оценок. Значение увеличения рассчитывалось как разность между финальной и промежуточной оценками.

Таблица 3. Среднее увеличение оценки в разрезе моделей и величин оценок
Table 3. Average scores increase breakdown by model and score value

Оценка промежуточной верификации	Среднее увеличение оценки модели				Среднее увеличение оценки
	deepseek	gemini	llama	qwен	
3			4,50		4,50
4			2,50		2,50
6			1,40	2,00	1,50
6,5			1,00		1,00
7	1,00	1,63	0,65	0,83	0,81
7,5	1,00	0,50	0,50	0,67	0,63
8	0,50	0,41	0,13	0,34	0,37
8,5	0,00	0,00	-0,50	0,00	-0,09
9	-0,31	-0,20	-0,88	-0,44	-0,34
9,5	-0,50	-0,50		-0,50	-0,50
3–9,5	0,20	0,21	0,5	0,21	0,28

Согласно табл. 3, в случаях, когда первичный ответ был оценен относительно низким баллом, меньшим или равным 8, финальная оценка повышалась тем сильнее, чем ниже была первичная оценка. Замечания, которые высказывались рецензирующей моделью, учитывались основной моделью при повторной генерации, и ответ становился более полным и корректным. Особенно заметно это проявляется на примере модели llama, которая имела низкую среднюю оценку первичных ответов. В случаях, когда первичный ответ был оценен баллом, большим или равным 8,5, финальная оценка понижалась тем сильнее, чем выше была первичная оценка. Непосредственный экспертный анализ отзывов в этих случаях показывает, что первичные ответы были уже достаточно полные и объемные, и второй вариант ответов часто расширял пояснения, вставлял дополнительные уточнения и примеры. Расширение ответов приводило к их «размытию» и появлению свежих недочетов. Ответы становились слишком многословными, появлялись описательные фрагменты, не относящиеся напрямую к полученному вопросу. А это свидетельствует об избыточности повторной генерации ответа для ответов с высокой первичной оценкой.

Заключение

1. Рассмотрены возможности применения больших языковых моделей (LLM) для создания текстов учебных пособий на примере конкретной дисциплины «Базы данных» для студентов университета, изучающих информационные технологии.

2. Предложена и апробирована циклическая многоступенчатая методика генерации текстов, основанная на использовании двух независимых LLM, одна из которых («рассуждающая») выступает в роли эксперта. Методика позволяет добиться полноты и правильности ответов, минимизировать появление нерелевантных запросам ответов (галлюцинаций). Проведена верификация этой методики с использованием метода эталонного источника и модифицированного CoVe-алгоритма. Показано, что современные производительные LLM Gemini, DeepSeek, Qwen способны отвечать с низкой вероятностью галлюцинаций при формировании содержимого курса «Базы данных». Это связано с корректным выбором предметной области в силу наличия сохраненных в памяти конкретных LLM непротиворечивых материалов, релевантных вопросам, формирующим содержимое этого курса.

3. Показано, что многоступенчатую методику следует упрощать, сведя количество этапов генерации вариантов ответов только к одному этапу при высокой промежуточной оценке ответа на вопрос. Это повысит скорость генерации текстов и их качество. Следует также отметить, что человеческая экспертиза остается важнейшим элементом контроля качества учебных материалов, позволяя выявлять ошибки и предвзятость. Роль преподавателя заключается в формулировании списка вопросов, подходящих текстов запросов (промптов), выборочном рецензировании ответов и адаптации материалов к учебной программе. Использование LLM при написании учебных материалов позволяет повысить возможности отдельного преподавателя, поскольку будут задействованы известные мировые источники, о которых преподаватель может и не знать.

4. Результаты исследования могут быть использованы для создания учебных материалов и по другим учебным дисциплинам при условии корреляции их содержимого со «знаниями» выбираемых LLM и подготовки набора вопросов, конкретизирующих их содержимое. Использование методики позволит существенно ускорить создание качественных электронных учебно-методических комплексов, а также предоставит возможность создания материалов на разных языках, используя мультязычные LLM.

5. Дальнейшие исследования могут быть направлены на оптимизацию методики, применение ее к иным дисциплинам и интеграцию с другими образовательными технологиями и системами. Это позволит повысить эффективность использования LLM и обеспечить студентов современными, адаптивными и качественными учебными материалами.

Список литературы

1. Large Language Models in Education: Vision and Opportunities / W. Gan [et al.] // 2023 IEEE International Conference on Big Data. 2023. P. 4776–4785. <https://doi.org/10.1109/BigData59044.2023.10386291>.
2. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education / E. Kasneci [et al.] // Learning and Individual Differences. 2023. Vol. 103. <https://doi.org/10.1016/j.lindif.2023.102274>.
3. Bonner, E. Large Language Model-Based Artificial Intelligence in the Language Classroom: Practical Ideas for Teaching / E. Bonner, R. Lege, E. Frazier // Teaching English with Technology. 2023. Vol. 23, No 1. P. 23–41.
4. Baidoo-Anu, D. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning / D. Baidoo-Anu, L. Owusu Ansah // Journal of AI. 2023. Vol. 7, No 1. P. 52–62. <https://doi.org/10.61969/jai.1337500>.
5. A Comprehensive Review on Generative AI for Education / U. Mittal [et al.] // IEEE Access. 2024. Vol. 12. P. 142733–142759. <https://doi.org/10.1109/ACCESS.2024.3468368>.
6. Enlarged Education – Exploring the Use of Generative AI to Support Lecturing in Higher Education / D. Hennekeuser [et al.] // International Journal of Artificial Intelligence in Education. 2024. <https://doi.org/10.1007/s40593-024-00424-y>.
7. Generative AI in Education: Perspectives Through an Academic Lens / I. Întorsureanu [et al.] // Electronics. 2025. Vol. 14. <https://doi.org/10.3390/electronics14051053>.
8. Evaluating the Quality of AI-Generated Digital Educational Resources for University Teaching and Learning / Q. Huang [et al.] // Systems. 2025. Vol. 13. <https://doi.org/10.3390/systems13030174>.
9. Survey of Hallucination in Natural Language Generation / Z. Ji [et al.] // ACM Computing Surveys. 2023. Vol. 55, No 12. P. 1–38. <https://doi.org/10.1145/3571730>.
10. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis [et al.] // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474.
11. Chain-of-Verification Reduces Hallucination in Large Language Models / S. Dhuliawala [et al.] // arXiv:2309.11495. 2023. <https://doi.org/10.48550/arXiv.2309.11495>.

Поступила 15.07.2025

Принята в печать 30.08.2025

Доступна на сайте 12.01.2026

References

1. Gan W., Wang Z., Li J., Li Y., Zheng Y., Yu J. (2023) Large Language Models in Education: Vision and Opportunities. *2023 IEEE International Conference on Big Data*. 4776–4785. <https://doi.org/10.1109/BigData59044.2023.10386291>.
2. Kasneci E., Seßler K., Küchemann S., Bannert M., Dementieva D., Fischer F., et al. (2023) ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*. 103. <https://doi.org/10.1016/j.lindif.2023.102274>.

3. Bonner E., Lege R., Frazier E. (2023) Large Language Model-Based Artificial Intelligence in the Language Classroom: Practical Ideas for Teaching. *Teaching English with Technology*. 23 (1), 23–41.
4. Baidoo-Anu D., Owusu Ansah L. (2023) Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*. 7 (1), 52–62. <https://doi.org/10.61969/jai.1337500>.
5. Mittal U., Sai S., Chamola V., Sangwan D. (2024) A Comprehensive Review on Generative AI for Education. *IEEE Access*. 12, 142733–142759. <https://doi.org/10.1109/ACCESS.2024.3468368>.
6. Hennekeuser D., Vaziri D. D., Golchinfar D., El-Khatib K. (2024) Enlarged Education – Exploring the Use of Generative AI to Support Lecturing in Higher Education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-024-00424-y>.
7. Întorsureanu I., Oprea S.-V., Bâra A., Vespan D. (2025) Generative AI in Education: Perspectives Through an Academic Lens. *Electronics*. 14. <https://doi.org/10.3390/electronics14051053>.
8. Huang Q., Lv C., Lu L., Tu S. (2025) Evaluating the Quality of AI-Generated Digital Educational Resources for University Teaching and Learning. *Systems*. 13. <https://doi.org/10.3390/systems13030174>.
9. Ji Z., Lee N., Frieske R., Yu T., Su D., Xu Y., et al. (2023) Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*. 55 (12), 1–38. <https://doi.org/10.1145/3571730>.
10. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., et al. (2020) Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems*. 33, 9459–9474.
11. Dhuliawala S., Komeili M., Xu J., Raileanu R., Weston J., Roller S. (2023) Chain-of-Verification Reduces Hallucination in Large Language Models. *arXiv:2309.11495*. <https://doi.org/10.48550/arXiv.2309.11495>.

Received: 15 July 2025

Accepted: 30 August 2025

Available on the website: 12 January 2026

Сведения об авторе

Асенчик О. Д., канд. физ.-мат. наук, доц., доц. каф. информационных технологий, Гомельский государственный технический университет имени П. О. Сухого

Адрес для корреспонденции

246029, Республика Беларусь,
Гомель, просп. Октября, 48
Гомельский государственный технический
университет имени П. О. Сухого
Тел.: +375 29 696-75-01
E-mail: olgasn@tut.by
Асенчик Олег Даниилович

Information about the author

Asenchik A. D., Cand. Sci. (Phys. and Math.), Associate Professor, Associate Professor at the Department of Information Technology, Sukhoi State Technical University of Gomel

Address for correspondence

246029, Republic of Belarus,
Gomel, Oktyabrya Ave., 48
Sukhoi State Technical
University of Gomel
Tel.: +375 29 696-75-01
E-mail: olgasn@tut.by
Asenchik Aleh Daniilovich