***УДК 378.147:004.8:811.124***

***Е. Ф. Асенчик, О.Д. Асенчик***

*г. Гомель, ГГУ имени Ф. Скорины,*

*г. Гомель, ГГТУ имени П.О. Сухого*

**ИСПОЛЬЗОВАНИЕ БОЛЬШОЙ ЯЗЫКОВОЙ МОДЕЛИ DEEPSEEK ДЛЯ СОЗДАНИЯ УЧЕБНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ "ЛАТИНСКИЙ ЯЗЫК"**

Интеграция больших языковых моделей (LLM), являющихся ведущим компонентом технологий искусственного интеллекта (ИИ), в образовании может изменить практику преподавания и обучения [1]. LLM используются для автоматизации различных образовательных задач, включая создание учебных материалов, предоставление обратной связи и выставление оценок. Они управляемым образом могут генерировать материалы, ориентированные на обучающегося, помогать в разработке учебных программ.

Применение LLM для создания разнообразных учебных материалов остается малоизученной областью. Существующие подходы к созданию учебных пособий чаще всего основываются на традиционных методах, таких как ручное составление текстов, использование существующих учебников и адаптация их под конкретные программы. В настоящей работе представляются результаты применения выбранной LLM для автоматизированного создания учебного материала по учебной дисциплине «Латинский язык».

В качестве используемой LLM был выбрана DeepSeek v3 – версия большой языковой модели, разработанная компанией DeepSeek AI. На данный момент она обладает рядом особенностей, выводящих ее на лидирующие позиции в мире и подходящих для выбранной задачи: модель обучалась на огромном наборе текстовых данных и имеет 671 миллиард параметров; демонстрирует высокую производительность в различных задачах, включая понимание и генерацию текста, работу с различными языками, ответы на вопросы; имеет открытый исходный код; поддерживает длинные контексты – обрабатывает до 128 тысяч токенов [2].

Выбор дисциплины «Латинский язык» в качестве объекта обусловлен тем, что это классическая дисциплина. Латинский язык обладает стабильностью грамматических правил, что должно позволить минимизировать ошибки в сгенерированных материалах. Для построения конкретного учебного курса по этой дисциплине в распоряжении преподавателя имеется большое количество устоявшихся и авторитетных источников, в том числе расположенных в Интернет и проиндексированных поисковыми системами и LLM при их обучении.

Роль преподавателя в этом процессе заключается в формулировании подходящих запросов (промптов), рецензировании ответов и адаптации материалов к учебной программе.

Новизна предлагаемого подхода заключается в использовании LLM DeepSeek для генерации учебных материалов на основе заранее подготовленных преподавателем вопросов по дисциплине и запросов (промптов), описывающих, как отвечать на поставленные вопросы и формировать окончательный ответ. Этот метод позволяет систематизировать информацию в соответствии с учебной программой, обеспечить высокую точность и актуальность материалов, сократить время на подготовку пособий.

Целью настоящего исследования была разработка методики создания учебного пособия по латинскому языку для студентов с использованием LLM на примере DeepSeek и оценка качества полученных материалов для студентов специальности русская филология (профилизация: литературно-редакционная деятельность).

Для создания учебного пособия был использован доступ к API DeepSeek из программного кода на языке Python [3]. Язык формулирования запросов – русский, язык ответов – русский и латинский. Для получения текста пособия было испытано несколько методик.

Прямое получение ответов на все вопросы сразу – этот подход предполагает генерацию ответов на все вопросы одновременно. Он оказался неэффективным из-за недостаточного размера контекстного окна 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 ответ либо не удавалось получить, либо ответы на вопросы были слишком краткие.

В дальнейшем применялась методика ответа на каждый вопрос по отдельности. Было проанализировано два варианта: прямое получение ответа от LLM на поставленный вопрос и получение ответа после обсуждения. Во втором случае полученный прямой вопрос дополнительно анализировался и оценивался LLM, и на основании первичного ответа и рекомендаций LLM формировала окончательный ответ.

Первый способ оказался менее эффективным и ответы менее полными. Например, в ответах на вопросы присутствовали фактические и стилистические ошибки («грамматика латинского языка была стандартизирована благодаря *усилиям таких филологов* как Варрон и Квинтилиан»; «к первому склонению относились и некоторые *мужские существительные*») и др.

Подход с обсуждением позволил улучшить качество ответов, например, были исправлены допущенные ранее недочёты: «это литературный язык, зафиксированный *в произведениях таких авторов*, как Цицерон, Цезарь, Вергилий, Овидий и Гораций»; «к первому склонению могут относиться и *существительные мужского рода*: это названия профессий, рода занятий, национальностей».

Кроме перечисленного, был реализован метод обсуждения с дополнительным использованием в качестве контекста сжатой информации из файла имеющегося учебного пособия. Он продемонстрировал соответствие ответов эталонному источнику.

Таким образом, анализ качества полученных ответов показал, что наиболее эффективной методикой является последовательное получение ответов после обсуждения. Использованные вопросы, полученные учебные пособия, журналы обсуждений и эталонный источник приведены в репозитории [3], там же, в файлах программного кода, приведены использованные настройки модели и промпты.

Следует отметить, что стилистические ошибки могут присутствовать в разных методиках, например: «существительные первого *склонения склоняются* *по следующим окончаниям*». Возможно, это связано с тем, что “рабочим” языком DeppSeek, является английский: все вопросы и запросы (промпты) перед обработкой переводятся на него, и на окончательном этапе ответы переводятся на русский со вставками латинского языка.

Как показало исследование, качество полученных ответов достаточно высокое и зависит от: формулировок запросов и вопросов; использованных настроек модели; учета специфики латинского языка (например, влияющие на качество перевода и определение грамматических категорий позиция глагола в абсолютном конце предложения, перемежающийся порядок слов, беспредложное употребление существительных и др.); использования дополнительных источников для проверки генерации информации.

Дальнейшие исследования могут быть направлены на оптимизацию методики, применения ее к другим дисциплинам и интеграцию с другими образовательными технологиями и системами. Это позволит повысить эффективность использования ИИ и обеспечить студентов современными и качественными учебными материалами.

**Литература**

1. Bonner E., Lege R., Frazier 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, № 1. – P. 23–41.

2. DeepSeek v3: расширенная языковая модель искусственного интеллекта [Электронный ресурс]. – URL: <https://deepseekv3.org/ru> (дата обращения: 30.01.2025).

3. Репозиторий latin\_book на GitHub [Электронный ресурс]. – URL: <https://github.com/Olgasn/latin_book> (дата обращения: 30.01.2025).