

DOI: 10.22363/2312-8631-2026-23-1-57-74

EDN: YJEOEV

УДК 378.147

Научная статья / Research article

## Методика применения больших языковых моделей при формировании фондов оценочных средств в высшем образовании

Д.М. Назаров  , С.В. Бегичева 

Уральский государственный экономический университет, Екатеринбург, Российская Федерация

 slup20005@mail.ru

**Аннотация.** *Постановка проблемы.* В современном вузовском обучении устойчиво растет потребность в масштабируемых и нормативно корректных фондах оценочных средств, обеспечивающих объективную проверку сформированности компетенций и соответствие аккредитационным требованиям. Практика показывает дефицит унифицированных процедур и значительные трудозатраты при ручной разработке заданий, а также различия методических подходов между кафедрами. На этом фоне востребованы технологии генеративного искусственного интеллекта, позволяющие ускорять подготовку разнотипных заданий без снижения их качества. *Методология.* Исследование опирается на принципы компетентностной модели оценивания и архитектуру трансформеров (механизм внимания) как теоретическую основу генерации заданий. Предложен воспроизводимый алгоритм: спецификация фондов оценочных средств в привязке к матрице компетенций и рабочих программ дисциплин; инженерия запросов с требованиями к точности и структуре; генерация; экспертная верификация и пилотное тестирование; доработка и интеграция. Апробация проведена на материале дисциплины «Алгоритмы нейронных сетей» направления «Бизнес-информатика». Сформированы тесты с выбором ответа, задания на соответствие и последовательность, а также кейсы с исправлением кода. Оценивались трудозатраты, понятность, уровень сложности и содержательная валидность. *Результаты.* Подход позволил существенно сократить цикл разработки: генерация банка заданий заняла часы вместо типичных нескольких дней; по итогам в фонд оценочных средств включены 50 вопросов. Из 120 сгенерированных позиций редакционных правок потребовали 17 вопросов и 23 ответа; апробация с участием студентов показала высокую понятность формулировок при среднем и высоком уровне сложности, а экспертная оценка подтвердила соответствие содержания целям дисциплины и требованиям объективности. *Заключение.* Представленная методика обеспечивает воспроизводимое формирование фонда оценочных средств, снижает трудоемкость подготовки материалов и повышает управляемость качества оценивания в условиях аккредитационных процедур. Универсальность подхода позволяет распространять его на иные дисциплины при адаптации

© Назаров Д.М., Бегичева С.В., 2026



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License <https://creativecommons.org/licenses/by-nc/4.0/legalcode>

спецификаций и шаблонов запросов. Практическая ценность усилится при интеграции с Learning Management System и процедурами внутреннего контроля качества. Перспективы связаны с расширением психометрической проверки (надежность, справедливость), разработкой предметно-ориентированных библиотек запросов и дальнейшей автоматизацией аналитики результатов.

**Ключевые слова:** компетентностная модель оценивания, аккредитационная экспертиза, инженерия запросов, спецификация оценочных материалов, механизм внимания, трансформеры, валидность, надежность

**Вклад авторов.** Д.М. Назаров – концепция, дизайн и разработка методологии исследования, предоставление ресурсов, написание текста, надзор и руководство за планированием и выполнением работы. С.В. Бегичева – реализация компьютерного кода и вспомогательных алгоритмов, сбор, верификация, анализ, синтез, администрирование и визуализация данных. Все авторы прочли и одобрили окончательную версию рукописи.

**Заявление о конфликте интересов.** Авторы заявляют об отсутствии конфликта интересов.

**История статьи:** поступила в редакцию 21 марта 2025 г.; доработана после рецензирования 13 сентября 2025 г.; принята к публикации 29 сентября 2025 г.

**Для цитирования:** Назаров Д.М., Бегичева С.В. Методика применения больших языковых моделей при формировании фондов оценочных средств в высшем образовании // Вестник Российского университета дружбы народов. Серия: Информатизация образования. 2026. Т. 23. № 1. С. 57–74. <http://doi.org/10.22363/2312-8631-2026-23-1-57-74>

## A methodology for using large language models to compile assessment tools in higher education

Dmitry M. Nazarov  , Svetlana V. Begicheva 

*Ural State University of Economics, Ekaterinburg, Russian Federation*  
slup20005@mail.ru

**Abstract.** *Problem statement.* In modern university education, there is a steadily growing need for scalable and normatively correct assessment tools that ensure objective verification of the development of competencies and compliance with accreditation requirements. Practice shows a shortage of unified procedures and significant labor costs in the manual development of tasks, as well as differences in methodological approaches between departments. Against this background, generative artificial intelligence technologies are in demand, allowing to accelerate the preparation of different types of tasks without compromising quality. *Methodology.* The study is based on the principles of the competence-based assessment model and the architecture of transformers (attention mechanism) as a theoretical basis for task generation. A reproducible algorithm is proposed: specification of assessment tools in relation to the matrix of competencies and work programs of disciplines; query engineering with requirements for accuracy and structure; generation; expert verification and pilot testing; refinement and integration. The pilot was conducted on the material of the discipline *Neural Network Algorithms* of the *Business Informatics* direction. Tests with multiple choice answers, assignments for compliance and sequence, as well as cases with code correction were formed. Labor costs, comprehensibility, level of complexity and content validity were assessed. *Results.* The approach made it possible

to significantly shorten the development cycle: generation of the task bank took about hours instead of the typical several days; as a result, 50 questions were included in the assessment tool fund. Of the 120 generated positions, 17 questions and 23 answers required editorial corrections; the pilot with the participation of students showed high comprehensibility of the wording at a medium and high level of complexity, and the expert assessment confirmed the compliance of the content with the goals of the discipline and the requirements of objectivity. *Conclusion.* The presented methodology ensures reproducible formation of a fund of assessment tools, reduces the labor intensity of preparing materials and increases the manageability of the quality of assessment in the context of accreditation procedures. The universality of the approach allows its extension to other disciplines by adapting the specifications and query templates. The practical value is enhanced by integration with Learning Management System and internal quality control procedures. Prospects are associated with the expansion of psychometric testing (reliability, fairness), the development of subject-oriented query libraries and further automation of results analysis.

**Keywords:** competency-based assessment model, accreditation expertise, query engineering, assessment materials specification, attention mechanism, transformers, validity, reliability

**Authors' contribution.** *Dmitry M. Nazarov* – conceptualization, resources, writing – review and editing, supervision. *Svetlana V. Begicheva* – software, data curation, validation, formal analysis, project administration, visualization. All authors have read and approved the final version of the manuscript.

**Conflicts of interest.** The authors declare that there is no conflict of interest.

**Article history:** received 21 March 2025; revised 13 September 2025; accepted 29 September 2025.

**For citation:** Nazarov DM, Begicheva SV. A methodology for using large language models to compile funds of assessment tools in higher education. *RUDN Journal of Informatization in Education*. 2026;23(1):57–74. (In Russ.) <http://doi.org/10.22363/2312-8631-2026-23-1-57-74>

**Постановка проблемы.** Актуальность использования больших языковых моделей (LLM) для составления фондов оценочных средств (ФОС) в образовательном процессе обусловлена необходимостью постоянного повышения качества и эффективности проверки знаний студентов как в рамках промежуточной аттестации, так и со стороны Рособрнадзора – уполномоченного органа Министерства образования. В условиях современных требований при прохождении процедуры аккредитации образовательные организации обязаны обеспечивать проверяющий орган фондами оценочных средств для проверки качества знаний студентов по профессиональным компетенциям. Автоматизация процесса создания оценочных материалов с помощью LLM позволяет значительно сократить время и трудозатраты преподавателей, обеспечивая при этом высокую точность и разнообразие заданий, как это требуется со стороны Рособрнадзора. Такие модели способны генерировать задания различного уровня сложности, адаптированные под индивидуальные потребности студентов, что способствует не только выполнению требований надзорных органов, но и организации процесса персонализации в рамках обучения студентов по различным направлениям бакалавриата и магистратуры. Разработка и внедрение методики составления заданий для студентов с помощью

LLM-моделей в образовательный процесс особенно актуальны в условиях постоянно развивающихся технологий дистанционного и гибридного обучения, которые используют инновационные подходы в контексте подачи и проверки знаний обучающихся. Отметим и тот факт, что применение больших языковых моделей в режиме реального времени позволит обеспечить соответствие образовательных программ современным требованиям рынка труда. Таким образом, использование LLM для создания ФОС – важный шаг к совершенствованию системы оценки знаний и навыков студентов в образовательных организациях.

LLM – это модель, основанная на архитектуре нейронных сетей, которые обучаются на больших объемах текстовых данных для того, чтобы понимать семантику текстовых запросов и генерировать тексты на естественном языке. Архитектуру нейронных сетей, используемых в LLM-моделях, предложил Ашиш Васвани в 2017 г. и назвал такой тип искусственного интеллекта (ИИ) трансформерами [1]. Несмотря на широкие возможности LLM-модели в сфере образования, которые включают перевод, резюмирование (суммирование) текста, умение давать ответы на вопросы, генерировать код на разных языках программирования и многое другое, количество публикаций в рецензируемых изданиях как в России, так и за рубежом ограничено.

Российские авторы Д.И. Ермилов, С.А. Хлебников, Т.Н. Юдина в основном описывают возможности технологий GPT-3 и GPT-4, объясняют, почему эти модели произвели революцию в области ИИ, а также анализируют их применение в различных сферах деятельности, но не формулируют методику работы с этими инструментами [2–4]. Множество работ зарубежных авторов, в том числе S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. Ribeiro, Y. Zhang, K. Cheng, Z. Sun, Y. He, S. Gu, H. Wu, посвящено различным экспериментам, связанным с применением больших языковых моделей в медицинской деятельности. Эти ученые выявляют возможности и ограничения LLM-моделей в сфере здравоохранения, в частности обсуждают различные аспекты использования ИИ в профессиональной деятельности хирургов и делают выводы о приближении к созданию общего искусственного интеллекта [5; 6]. В образовательном контексте можно выделить исследования M. Firdaus, J. Wibawa, F. Rahman, которые обсуждают внедрение педагогической технологии персонализированного обучения с учетом образовательных потребностей поколения Z на примере использования производной от LLM-модели – чата GPT (технология GPT-4) [7]. В работах W. Gan, Z. Qi, J. Wu, C. Lin, M. Parker, C. Anderson, C. Stone, Y. Oh демонстрируется, что LLM-модели могут качественно выполнять задачи классификации, тематического анализа текста и анализа настроений, которые затем авторы используют для оценки образовательных достижений, анализа обратной связи в опросах обучающихся [8; 9]. Нами обнаружено лишь одно исследование J. Oppenlaender, R. Linder, J. Silvennoinen, в котором авторы обсуждают методику создания запросов (промтов) для ге-

нерации релевантных ответов сервисами, основанными на LLM-моделях. В этой работе авторы проводят анализ подходов к составлению запросов и того, как различные формулировки запросов к сервисам по созданию изображений влияют на качество и креативность генерируемых графических файлов [10–12]. В российских источниках таких научных работ нами обнаружено не было.

Таким образом, можно констатировать факт о том, что вопросы использования LLM-моделей в образовательной сфере сегодня остаются открытыми, поскольку эффективность работы таких технологий недостаточно изучена и в российской, и в зарубежной науке. Не существует опубликованных методик и принципов создания эффективных запросов к сервисам такого рода.

Цель статьи – формулировка принципов написания эффективных запросов к большим языковым моделям (LLM) и разработка методики составления фондов оценочных средств для оценки уровня сформированности компетенций студентов по требованиям Рособнадзора.

**Методология.** Для достижения заявленной цели были проанализированы современные теоретические и практические аспекты использования LLM-моделей в образовании, изучен механизм внимания, являющийся ключевым компонентом трансформерных нейронных сетей. В основе работы трансформера лежит механизм внимания, позволяющий модели выделять наиболее важные части входного текста (запроса) при генерации ответа. Он оперирует тремя матрицами: запросами (Q), ключами (K) и значениями (V). Для каждого токена входной последовательности вычисляются соответствующие векторы, а итоговая релевантность определяется умножением Q на K, что помогает модели сосредоточиться на смысле. Значения (V) несут фактическую информацию, используемую для генерации ответа.

Изучение и анализ этого механизма внимания позволили авторам сформулировать принципы написания эффективных запросов к LLM-моделям.

1. Четкость и конкретность формулировок.
2. Целеполагание в образовательном контексте.
3. Структурирование запроса: краткость, логичность, ясность.
4. Учет специфики образовательного запроса и целевой аудитории.

Компетентностный подход в образовании подразумевает оценку сформированности у студентов практических навыков и профессиональных компетенций. Фонд оценочных средств каждой дисциплины включает тесты с открытыми и закрытыми вопросами, контрольные и практические задания, кейсы и проекты, позволяющие объективно измерить уровень владения компетенциями.

Составление ФОС вручную требует серьезных временных затрат и нередко сопровождается различными трудностями: отсутствием унифицированных подходов к формированию заданий среди преподавателей разных кафедр; недостатком методических материалов по новым дисциплинам; необходимостью составления большого количества тестовых вопросов с вариантами ответов;

проверкой корректности эталонных решений. Применение разработанной нами пошаговой методики, основанной на использовании LLM-моделей, устраняет многие из этих проблем.

Этапы предложенной методики:

- 1) подготовительный – определяются типы оценочных средств и дисциплины, формирующие заданную компетенцию;
- 2) создание оценочных заданий – формулируются структурированные запросы (промпты) к LLM-модели;
- 3) генерация заданий – модель выдает варианты вопросов и ответов, которые проверяются и корректируются (при необходимости);
- 4) валидация и тестирование – проводится экспертиза преподавателями, пилотное тестирование на малой группе студентов и сбор обратной связи;
- 5) формирование итоговой версии ФОС – внесение правок и интеграция заданий в учебные курсы.

**Результаты и обсуждение.** Примеры больших языковых моделей – ChatGPT, разработанный компанией OpenAI, Yandex GPT, основанный на модели YaLM (Yandex Language Model) от Яндекс, и GigaChat от корпорации Сбер. Последние две являются одними из крупнейших языковых моделей, адаптированных для генерации текста на русском языке. Все перечисленные модели в работе используют архитектуру трансформера, которая включает в себя механизм внимания, предназначенный для качественной обработки текстового запроса. Для того чтобы сформулировать принципы написания эффективных запросов, необходимо разобраться, как устроен механизм внимания. Сделаем это на конкретном примере.

Рассмотрим, как GPT генерирует ответ на вопрос: «Как работает механизм распределения нагрузки в компьютерных сетях?»

Механизм внимания позволяет модели выделять наиболее важные части входного текста (запроса пользователя) при генерации ответа. Это помогает LLM-модели сосредоточиться на ключевых элементах запроса и генерировать более точные и релевантные ответы.

Когда модель получает запрос, она разбивает его на отдельные токены (слова или части слов). Для каждого токена вычисляются три матрицы: запросы (Q) – представляют текущий токен и его контекст, как бы задавая вопрос, на что следует обратить внимание; ключи (K) – содержат информацию о всех токенах в последовательности, выступая в роли «ответа» на запрос; значения (V) – хранят фактическую информацию, которая будет использоваться для генерации ответа [13–15].

1. Токенизация. Запрос разбивается на токены: [«как», «работает», «механизм», «распределения», «нагрузки», «в», «компьютерных», «сетях», «?»].

2. Вычисление Q, K и V. Эти матрицы помогают модели определять, на какие части текста нужно обратить внимание при обработке каждого токена.

Простыми словами, механизм внимания помогает модели понять, какие части запроса важны для понимания смысла и формирования ответа.

Приведем примеры значений.

- Токен «Как»:  $Q: [0.5, 0.1, -0.3]$  – представляет контекст и смысл токена «как»;  $K: [0.4, 0.2, -0.1]$  – определяет важность других слов относительно токена «как»;  $V: [0.7, 0.0, 0.4]$  – хранит фактическую информацию о слове «как».
- Токен «работает»:  $Q: [0.6, -0.2, 0.8]$ ;  $K: [0.5, -0.1, 0.6]$ ;  $V: [0.9, -0.4, 0.3]$ .
- Токен «механизм»:  $Q: [0.7, 0.3, -0.5]$ ;  $K: [0.6, 0.1, -0.2]$ ;  $V: [1.0, 0.2, -0.1]$ .

Числа, указанные в примере, находятся на основе весов, которые получаются во время процесса обучения (тренировки) на больших объемах текстов, чтобы модель в дальнейшем могла эффективно определять контекст и значение слов в разных предложениях. В процессе обучения модель оптимизирует эти веса, чтобы минимизировать ошибку предсказания.

В конечном итоге эти вычисления приводят к созданию числовых векторов  $Q$ ,  $K$  и  $V$ , которые затем используются в механизме внимания для определения релевантности и значимости каждого токена (слова или его части) в контексте всего предложения.

Модель вычисляет внимание (attention) как произведение запросов ( $Q$ ) и ключей ( $K$ ). Это помогает определить, насколько каждый токен важен в контексте других токенов запроса. Например, если токен «механизм» имеет высокую важность при умножении  $Q$  и  $K$  по сравнению с другими токенами, то модель будет обращать больше внимания на этот токен при генерации ответа.

3. Генерация ответа. После вычисления показателя «внимание» модель использует значения ( $V$ ) для формирования ответа. В нашем примере модель будет учитывать, что токены «механизм», «распределения» и «нагрузки» являются ключевыми для понимания вопроса, и на основе этого генерировать ответ, описывающий, как работает механизм распределения нагрузки в компьютерных сетях.

4. Ответ LLM-модели. Механизм распределения нагрузки в компьютерных сетях предназначен для равномерного распределения трафика между несколькими серверами или узлами, чтобы оптимизировать использование ресурсов и предотвратить перегрузки. Основные методы включают балансировку на основе системы DNS, использования прокси-серверов и применения алгоритмов хеширования [12].

Таким образом, механизм внимания и вычисление  $Q$ ,  $K$  и  $V$  позволяют технологии GPT эффективно обрабатывать сложные запросы и генерировать точные и информативные ответы.

Исходя из описанного выше механизма генерации ответа на вопрос становится понятным, как необходимо строить запрос пользователя к LLM-модели, реализованной GPT-технологии.

Как показывают исследования, содержание запросов пользователей играет решающую роль в определении качества и точности ответов, получаемых от модели. Поэтому правильная формулировка запроса должна основываться на определенных правилах и принципах.

1. Четкость и конкретность формулировок. Этот принцип необходим для того, чтобы модель могла правильно интерпретировать информацию пользователя. Следует включить все важные детали, такие как тема, уровень сложности, цели задания. Например, запрос: «Создайте тест по математике для учеников 10 класса, включающий задачи на логарифмы, тригонометрию и вероятность. Каждое задание должно иметь один правильный ответ и четыре неправильных».

2. Целеполагание в образовательном контексте. Необходимо указать контекст и цели запроса, чтобы модель могла понять, что от нее ожидается в итоге и сгенерировать релевантный и точный ответ. Пример запроса: «Составьте задания для итогового экзамена по истории для студентов первого курса. Включите вопросы о ключевых событиях Второй мировой войны и их значении».

3. Структурирование. Требуется структурировать запрос таким образом, чтобы модель могла легко определить ключевые элементы. Используйте краткие и ясные предложения. Пример: «Разработайте 5 открытых вопросов для теста по биологии для 8 класса. Вопросы должны охватывать темы: фотосинтез, клеточное дыхание, генетика».

4. Учет специфики образовательного запроса и целевой аудитории. Запросы от преподавателей для модели могут включать специализированную терминологию и формулироваться более сложно, в то же время запросы от студентов могут быть более простыми. Пример запроса от преподавателя: «Разработайте 5 открытых вопросов для экзамена по физике для студентов, включающие задачи по специальным разделам квантовой механики и термодинамики». Пример запроса от студента: «Создайте краткий тест из трех открытых вопросов по основам физики для 10 класса, включающий вопросы о законе сохранения энергии».

Применение этих принципов поможет эффективно использовать большие языковые модели для создания качественных и релевантных фондов оценочных средств, улучшая образовательный процесс и обеспечивая адекватную оценку знаний студентов.

Компетентностный подход в образовании направлен на развитие у студентов конкретных знаний, умений и навыков, которые необходимы для успешного выполнения профессиональных задач в той сфере деятельности, по которой проходит обучение. В этом подходе акцент делается не только на усвоении теоретических знаний, но и на их практическом применении при решении профессиональных задач. Фонды оценочных средств представляют собой набор инструментов, предназначенных для оценки сформированности компетенций у студентов, которые разрабатываются в рамках каждой рабочей программы дисциплины. ФОС включают большое количество разнообразных упражнений, представленных в виде тестов с открытыми и закрытыми вопросами, ситуационных задач, кейсов, контрольных работ, практических заданий, проектов и др., которые, по мнению разработчика, позволяют объективно измерить уровень владения определенными компетенциями.

В учебных планах, составляемых кафедрами в рамках подготовки бакалавров, магистров и специалистов, выбор и путь формирования компетенций определяется целями подготовки студентов, федеральными государственными образовательными и профессиональными стандартами, требованиями контролирующих органов в системе образования. Для каждого учебного плана создается матрица компетенций, которая ставит в соответствие каждой компетенции множество дисциплин, ее формирующих.

При этом возникает целый клубок проблем, который связан как с содержательными аспектами, так и с организационными. Перечислим основные.

Первая проблема связана с тем, что различные дисциплины ведут разные преподаватели, возможно даже с нескольких кафедр, а поэтому подходы для составления ФОС могут отличаться. Вторая проблема касается содержательных аспектов и заключается в том, что по некоторым дисциплинам учебного плана еще не наработано достаточно методического материала в целом, например по LLM-моделям. Третья проблема – это сложность и большие временные затраты на составление закрытых вопросов теста, где нужно подобрать, помимо вопроса, и адекватного уровня сложности ответы. Четвертая проблема заключается в формировании эталонных ответов на каждый открытый вопрос. Все это в первую очередь связано с большими временными затратами.

Со всеми этими проблемами помогает справиться разработанная нами методика использования LLM, реализованная на базе GPT-технологии.

*Применение методики использования LLM для составления ФОСов по дисциплине «Интеллектуальный анализ данных» учебного плана по направлению 38.03.05 «Бизнес-информатика».* Выберем компетенцию ПК-4. Проведем аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика. Для сокращения объема изложения покажем пример составления фондов оценочных средств только для одной дисциплины.

Рассмотрим тематический план по дисциплине «Алгоритмы нейронных сетей». Он включает пять основных тем.

1. Введение в нейронные сети. Базовые алгоритмы и основные понятия нейронных сетей.

2. Основные типы нейронных сетей.

3. Принципы обучения нейронных сетей. Переобучение (overfitting).

4. Deep Learning. Принципы и особенности процесса глубокого обучения.

5. Применение нейронных сетей в экономике и управлении.

Далее определимся с формами оценки уровня сформированности. Это все виды тестов и кейсы. Немаловажный вопрос заключается в том, какое количество заданий необходимо составлять по каждой теме. Опыт формирования и проведения диагностических работ показывает, что 4–5 заданий каждого вида вполне достаточно для достижения целей, связанных с оценкой сформированности компетенций.

Перейдем ко второму этапу и приведем несколько запросов к LLM-моделям.

1. Составьте 5 тестовых вопросов с выбором одного или нескольких вариантов ответа по теме «Введение в нейронные сети». Количество вариантов ответов равно 4. Укажите верный ответ. Верные ответы должны быть выделены жирным шрифтом. Вопросы должны охватывать основные концепции и историю развития нейронных сетей.

2. Составьте 5 заданий на установление соответствия между ключевыми терминами и их определениями по теме «Введение в нейронные сети». В каждом вопросе должно быть 4 варианта ответа. Укажите правильные соответствия.

3. Составьте 5 заданий на установление правильной последовательности шагов в процессе обучения нейронных сетей. В каждом вопросе должно быть 4 варианта ответа. Укажите правильную последовательность через запятую.

4. Составьте 5 заданий на установление соответствия между типами нейронных сетей и их применением в различных задачах. В каждом вопросе должно быть 4 варианта ответа. Укажите правильные соответствия.

5. Сформулируйте 5 вопросов открытого типа по теме Deep Learning, которые позволят студентам описать основные преимущества и особенности, связанные с глубоким обучением. Дайте правильный ответ. Не более трех предложений.

Составьте 5 кейсов по теме «Обучение нейронных сетей», которые требуют исправить ошибку в коде обработки данных для обучения нейронной сети. Приведите пример неверного фрагмента кода, который нужно исправить в первой или второй строчках. Дайте верный ответ и поясните, почему необходимо внести именно такие исправления.

Приведенные примеры запросов удовлетворяют всем принципам составления запросов и позволяют сгенерировать релевантные ответы.

Перейдем к третьему этапу – генерации заданий для составления ФОСов. Загрузим эти запросы в модель по одному. Фрагмент ответа на первый запрос приведен на рис. 1.

#### Вопрос 1

Что такое искусственная нейронная сеть (ИНС)?

1. Модель, основанная на биологических нейронах, предназначенная для имитации процесса обучения.
2. Компьютерная система, состоящая из взаимосвязанных узлов (нейронов), которые могут обучаться на основе входных данных.
3. Программа для создания графических изображений.
4. Система для хранения и управления большими объемами данных.

#### Вопрос 2

Кто является одним из основателей концепции нейронных сетей?

1. Уоррен Мак-Каллок и Уолтер Питтс.
2. Алан Тьюринг.
3. Клод Шеннон.
4. Джон фон Нейман.

**Рис. 1.** Фрагмент ответа на первый запрос

*Источник:* создано Д.М. Назаровым, С.В. Бегичевой.

**Question 1**

What is an artificial neural network (ANN)?

1. A model based on biological neurons, designed to imitate the learning process.
2. **A computer system consisting of interconnected nodes (neurons) that can learn based on input data.**
3. A program for creating graphic images.
4. A system for storing and managing large volumes of data.

**Question 2**

Who is one of the founders of the concept of neural networks?

1. **Warren McCulloch and Walter Pitts.**
2. Alan Turing.
3. Claude Shannon.
4. John von Neumann.

**Figure 1.** Fragment of the response to the first request

Source: created by Dmitry M. Nazarov, Svetlana V. Begicheva.

Анализ показал, что формулировка запроса, предложенная нами, была вполне адекватно распознана по смыслу LLM-моделью, и она предложила релевантный ответ, полностью соответствующий запросу. Фрагменты ответов системы на второй и третий запросы приведены на рис. 2 и 3.

**Задание 1**

Установите соответствие между термином и его определением.

Термин	Определение
1. Нейрон	A. Процесс настройки весов на основе входных данных и ошибок
2. Обратное распространение ошибки	B. Функция, используемая для активации нейрона
3. Активационная функция	C. Основной элемент нейронной сети, обрабатывающий входные данные
4. Синапс	D. Связь между нейронами, через которую передаются сигналы

**Правильные соответствия:**

1C, 2A, 3B, 4D

**Задание 2**

Установите соответствие между термином и его определением.

Термин	Определение
1. Перцептрон	A. Совокупность всех возможных значений, которые может принимать нейрон
2. Обучение	B. Основной алгоритм нейронных сетей для бинарной классификации
3. Входной слой	C. Первый слой нейронной сети, принимающий входные данные
4. Область определения	D. Процесс настройки параметров модели для минимизации ошибки

**Правильные соответствия:**

1B, 2D, 3C, 4A

**Рис. 2.** Фрагмент ответа на второй запрос

Источник: создано Д.М. Назаровым, С.В. Бегичевой.

**Task 1**

Match the term with its definition.

Term	Definition
1. Neuron	A. The process of adjusting weights based on input data and errors
2. Backpropagation	B. A function used to activate a neuron
3. Activation function	C. The basic element of a neural network that processes input data
4. Synapse	D. A connection between neurons through which signals are transmitted

**Correct matches:**

1C, 2A, 3B, 4D

**Task 2**

Match the term with its definition.

Term	Definition
1. Perceptron	A. The set of all possible values that a neuron can take
2. Training	B. The main algorithm of neural networks for binary classification
3. Input layer	C. The first layer of a neural network that receives input data
4. Domain	D. The process of adjusting model parameters to minimize error

**Correct matches:**

1B, 2D, 3C, 4A

**Figure 2.** Fragment of the response to the second request  
*Source:* created by Dmitry M. Nazarov, Svetlana V. Begicheva.

**Задание 1**

Установите правильную последовательность шагов в процессе обучения нейронной сети.

1. Обработка входных данных.
2. Обратное распространение ошибки.
3. Инициация весов.
4. Обновление весов.

Правильная последовательность: 3, 1, 2, 4.

**Рис. 3.** Фрагмент ответа на третий запрос  
*Источник:* создано Д.М. Назаровым, С.В. Бегичевой.

**Task 1**

Establish the correct sequence of steps in the process of training a neural network.

1. Processing input data.
2. Backpropagation.
3. Weight initialization.
4. Weight update.

Correct sequence: 3, 1, 2, 4.

**Figure 3.** Fragment of the response to the third request  
*Source:* created by Dmitry M. Nazarov, Svetlana V. Begicheva.

Результаты ответов LLM-модели вполне релевантные, а формулировки вопросов и вариантов ответов нетривиальны. Перейдем к шестому запросу. Фрагмент ответа модели приведен на рис. 4.

Таким образом, сгенерированные вопросы и задания включают в себя вопросы с выбором ответа, задания на установление соответствия, установление последовательности и исправление кода по тематике дисциплины «Алгоритмы нейронных сетей», обеспечивают достаточно глубокую проверку знаний студентов. Задания с исправлением кода ориентированы на практическое применение знаний, что помогает студентам лучше подготовиться к решению реальных задач в области нейронных сетей. Примеры неверных фрагментов кода и их исправления представлены моделью четко и ясно.

Приведем результаты работы по первому, второму и третьему этапу разработанной методики с учетом всех дисциплин, которые формируют компетенцию ПК-4. Время на выполнение всех этапов составления ФОСов представлено в табл. 1, а результаты пилотного тестирования с учетом обратной связи – в табл. 2.

**Кейс 1: Ошибка в нормализации данных**

**Описание:**  
В коде предобработки данных для обучения нейронной сети была допущена ошибка в нормализации данных.

**Неверный фрагмент кода:**

```
python Копировать код  
  
import numpy as np  
  
def normalize_data(data):  
    return (data - np.min(data)) / np.max(data)  
  
data = np.array([1, 2, 3, 4, 5])  
normalized_data = normalize_data(data)  
print(normalized_data)
```

**Верный ответ:**

```
python Копировать код  
  
def normalize_data(data):  
    return (data - np.min(data)) / (np.max(data) - np.min(data))
```

**Пояснение:**  
Необходимо вычесть минимальное значение из данных и разделить результат на диапазон данных (разницу между максимальным и минимальным значениями). Это обеспечит правильную нормализацию данных в диапазоне от 0 до 1.

**Рис. 4.** Фрагмент ответа на шестой запрос  
Источник: создано Д.М. Назаровым, С.В. Бегичевой.

**Case 1: Data Normalization Error**

**Description:**  
In the data preprocessing code for training a neural network, an error in data normalization was made.

**Incorrect code snippet:**

```
python Копировать код

import numpy as np

def normalize_data(data):
    return (data - np.min(data)) / np.max(data)

data = np.array([1, 2, 3, 4, 5])
normalized_data = normalize_data(data)
print(normalized_data)
```

**Correct answer:**

```
python Копировать код

def normalize_data(data):
    return (data - np.min(data)) / (np.max(data) - np.min(data))
```

**Explanation:**  
It is necessary to subtract the minimum value from the data and divide the result by the data range (the difference between the maximum and minimum values). This ensures correct normalization of the data within the range from 0 to 1.

**Figure 4.** Fragment of the response to the sixth request  
Source: created by Dmitry M. Nazarov, Svetlana V. Begicheva.

Итак, процесс генерации заданий и валидации ответов с использованием LLM-модели показал свою временную эффективность. Внутренняя экспертиза и пилотное тестирование помогли выявить сильные и слабые стороны сгенерированных заданий и внести ряд корректировок, касающихся в основном правки текста в соответствии с правилами русского языка. Внесенные изменения и анализ обратной связи позволили улучшить качество и релевантность оценочных средств. В целом методика использования LLM-моделей для создания фондов оценочных средств доказала свою применимость и высокую эффективность в образовательном процессе. В итоге в фонд оценочных средств по компетенции ПК-4 были включены 50 вопросов. В качестве реализации LLM-модели был использован ChatGPT-4o. Российские аналоги пока отстают от мирового лидера в части написания программного кода, формулировки кейсов. Однако тестовые вопросы и ответы в сервисах GigaChat<sup>1</sup> и Yandex GPT<sup>2</sup> были сформулированы вполне адекватно. На наш взгляд, сервис Yandex GPT формулировал тестовые задания лучше, чем GigaChat.

<sup>1</sup> ГигаЧат – русскоязычная нейросеть от Сбера. URL: <https://developers.sber.ru/gigachat> (дата обращения: 10.02.2025).

<sup>2</sup> Чат с Алисой. URL: <https://a.ya.ru/chat> (дата обращения: 10.02.2025).

**Таблица 1. Оценка времени на выполнение этапов разработанной методики**

Этап	Время на выполнение	Описание
Генерация запросов и ответов	2 часа	Создание тестовых вопросов, заданий на установление соответствия и последовательности, кейсов
Валидация и оценка качества	2 дня	Проведение внутренней экспертизы преподавателями или экспертами. Были внесены корректировки по формулировке ответов и вопросов. Из 120 сгенерированных вопросов и ответов небольших корректировок требовали 17 вопросов и 23 ответа
Пилотное тестирование	1 день	Проведение тестирования на малой группе студентов и сбор обратной связи
Анализ обратной связи	4 часа	Обработка и анализ обратной связи с помощью LLM-модели

Источник: составлено Д.М. Назаровым, С.В. Бегичевой.

**Table 1. Estimated time for completing the stages of the developed methodology**

Stage	Time to complete	Description
Generation of requests and responses	2 hours	Creating test questions, assignments for establishing correspondence and sequence, cases
Validation and quality assessment	2 days	Conducting internal examination by teachers or experts. Adjustments were made to the wording of answers and questions. Of the 120 questions and answers generated, 17 questions and 23 answers required minor adjustments
Pilot testing	1 day	Conducting testing on a small group of students and collecting feedback
Feedback analysis	4 hours	Processing and analyzing feedback using the LLM model

Source: compiled by Dmitry M. Nazarov, Svetlana V. Begicheva.

**Таблица 2. Результаты пилотного тестирования**

Задание	Количество студентов	Средний балл	Уровень сложности	Понятность смысла задания
Тестовые вопросы	10	8,5/10	Средний	Высокая
Задания на установление соответствия	10	7,8/10	Средний	Высокая
Задания на установление последовательности	10	7,5/10	Средний	Высокая
Кейсы с исправлением кода	10	6,9/10	Высокий	Средняя

Источник: составлено Д.М. Назаровым, С.В. Бегичевой.

**Table 2. Pilot testing results**

Task	Number of students	Average score	Difficulty level	Clarity of the task meaning
Test questions	10	8.5/10	Medium	High
Matching tasks	10	7.8/10	Medium	High
Sequencing tasks	10	7.5/10	Medium	High
Code correction cases	10	6.9/10	High	Average

Source: compiled by Dmitry M. Nazarov, Svetlana V. Begicheva.

**Заключение.** В ходе исследования были сформулированы такие принципы составления запросов к LLM-моделям, как четкость и конкретность, целеполагание, структурирование, учет специфики запроса и целевой аудитории, которые основаны на сути работы механизма внимания.

Разработанная методика использования LLM-моделей для составления ФОСов включает несколько ключевых этапов: подготовительный; создание оценочных заданий; генерация заданий; валидация и тестирование; формирование текста. ФОС показала свою эффективность в организации и проведении работ для составления различных типов учебных заданий.

Использование LLM для создания фондов оценочных средств в образовательном процессе по разработанной методике показало, что автоматизация создания тестов и кейсов значительно сократила временные затраты на разработку оценочных материалов.

Технология составления методики способствует повышению объективности оценки уровня сформированности компетенций студентов. Внутренняя экспертиза преподавателей кафедры бизнес-информатики Уральского государственного экономического университета и пилотное тестирование подтвердили высокое качество и релевантность заданий, соответствующих заданным учебным целям и структуре рабочей программы дисциплины. Полученные отзывы от экспертов и студентов в ходе пилотного тестирования позволили внести необходимые изменения и улучшить качество заданий ФОС. Сгенерированные кейсы с исправлением кода на языке Python позволили оценить развитие практических навыков студентов. Результаты исследования подтверждают перспективность использования LLM-моделей для дальнейшего совершенствования образовательного процесса и оценки знаний.

Дальнейшее направление исследований в этой сфере научного поиска может включать разработку и тестирование LLM-моделей для создания оценочных средств на разных языках и для студентов с различными культурными и языковыми особенностями. Перспектива ближайших лет – интеграция LLM-моделей с другими образовательными платформами и технологиями, такими как системы управления обучением (Learning Management System) и виртуальными классами. Долгосрочные исследования влияния использования LLM-моделей на качество образования и успешность студентов в профессиональной деятельности помогут оценить их реальное воздействие. Наконец, создание и тестирование специализированных LLM-моделей для различных дисциплин и уровней образования, включая начальное, среднее и высшее, может существенно расширить их применение и повысить эффективность обучения.

### Список литературы / References

- [1] Назаров Д.М., Бегичева С.В., Ковтун Д.Б., Назаров А.Д. *Data Science и интеллектуальный анализ данных* : учеб. пособие. М. : Ай Пи Ар Медиа, 2023. 304 с. EDN: VQKMUJ  
Nazarov DM, Begicheva SV, Kovtun DB, Nazarov AD. *Data Science and Data Mining: textbook*. Moscow: IPR Media Publ.; 2023. 304 p. (In Russ.) EDN: VQKMUJ
- [2] Ермилов Д.И. Chat GPT: новая технология общения с клиентами в банковской сфере // *Финансовые рынки и банки*. 2023. № 7. С. 18–22. EDN: IWGGEZ  
Ermilov DI. Chat GPT: a new technology for communicating with customers in the banking sector. *Financial Markets and Banks*. 2023;(7):18–22. (In Russ.) EDN: IWGGEZ

- [3] Хлебников С.А. Использование технологий искусственного интеллекта в сфере экономики: Chat GPT // Инженерные кадры – будущее инновационной экономики России : материалы IX Всерос. студ. конф., Йошкар-Ола, 7–10 ноября 2023 г. Йошкар-Ола : Поволжский гос. технол. ун-т, 2023. С. 1107–1110. EDN: LOMDNO  
Khlebnikov SA. The use of artificial intelligence technologies in the field of economics: Chat GPT. In: *Engineering Staff – The Future of Russia’s Innovative Economy: Proceedings of the IX All-Russian Student Conference, 7–10 November 2023, Yoshkar-Ola*. Yoshkar-Ola: Volga State University of Technology; 2023. p. 1107–1110. (In Russ.) EDN: LOMDNO
- [4] Юдина Т.Н. Тенденции и перспективы развития цифровых технологий. Искусственный интеллект GPT // Тенденции развития Интернет и цифровой экономики : труды VI Международной научно-практической конференции, Симферополь – Алушта, 1–3 июня 2023 г. Симферополь : ИП Зуева, 2023. С. 67–69. EDN: LVKLLKQ  
Yudina TN. Trends and prospects for the development of digital technologies. GPT artificial intelligence. In: *Trends in the Development of the Internet and Digital Economy: Proceedings of the VI International Scientific and Practical Conference, 1–3 June 2023, Simferopol – Alushta*. Simferopol: IP Zueva Publ.; 2023. p. 67–69. (In Russ.) EDN: LVKLLKQ
- [5] Bubeck S, Chandrasekaran V, Eldan R, et al. Sparks of artificial general intelligence: early experiments with GPT-4. *ArXiv*. 2023:2303.12712. <https://doi.org/10.48550/arXiv.2303.12712>
- [6] Cheng K, Sun Z, He Y, Gu S, Wu H. The potential impact of ChatGPT/GPT-4 on surgery: will it topple the profession of surgeons? *International Journal of Surgery*. 2023;109(5):1545–1547. <https://doi.org/10.1097/JS9.0000000000000388>
- [7] Firdaus MF, Wibawa JN, Rahman FF. Utilization of GPT-4 to improve education quality through personalized learning for generation Z in Indonesia. *IT for Society. Journal of Information Technology*. 2023;8(1):6–14. <https://doi.org/10.33021/itfs.v8i1.4728>
- [8] Gan W, Qi Z, Wu J, Lin JC-W. Large language models in education: vision and opportunities. *ArXiv*. 2023:2311.13160. <https://https://doi.org/10.48550/arXiv.2311.13160>
- [9] Parker M, Anderson C, Stone C, Oh Y. A Large Language Model approach to educational survey feedback analysis. *ArXiv*. 2023:2309.17447. <https://doi.org/10.48550/arXiv.2309.17447>
- [10] Oppenlaender J, Linder R, Silvennoinen J. Prompting AI art: an investigation into the creative skill of prompt engineering. *ArXiv*. 2023:2303.13534. <https://doi.org/10.48550/arXiv.2303.13534>
- [11] Caines A, Benedetto L, Taslimipoor S, et al. On the application of Large Language Models for language teaching and assessment technology. *ArXiv*. 2023:2307.08393. <https://doi.org/10.48550/arXiv.2307.08393>
- [12] Jojic A, Wang Z, Jojic N. GPT is becoming a Turing machine: here are some ways to program it. *ArXiv*. 2023:2303.14310. <https://doi.org/10.48550/arXiv.2303.14310>
- [13] Marr B. What is GPT-3 and why is it revolutionizing artificial intelligence? *Archives of Applied Science Research*. 2020;12:13–15.
- [14] Wu Y, Wang S, Yang H, Zheng T, Zhang H, Zhao Y, Qin B. An early evaluation of GPT-4V(ision). *ArXiv*. 2023:2310.16534. <https://doi.org/10.48550/arXiv.2310.16534>
- [15] Matzakos N, Doukakis S, Moundridou M. Learning mathematics with Large Language Models. *International Journal of Emerging Technologies in Learning (IJET)*. 2023;18(20):51–71. <https://doi.org/10.3991/ijet.v18i20.42979>

#### Сведения об авторах:

Назаров Дмитрий Михайлович, доктор экономических наук, доцент, заведующий кафедрой бизнес-информатики, Уральский государственный экономический университет,

Российская Федерация, 620144, Екатеринбург, ул. 8 Марта, д. 62. ORCID: 0000-0002-5847-9718; SPIN-код: 2424-3457. E-mail: slup20005@mail.ru

*Бегичева Светлана Викторовна*, кандидат экономических наук, доцент кафедры бизнес-информатики, Уральский государственный экономический университет, Российская Федерация, 620144, Екатеринбург, ул. 8 Марта, д. 62. ORCID: 0000-0002-0551-1622; SPIN-код: 1924-9572. E-mail: begichevas@mail.ru

**Bio notes:**

*Dmitry M. Nazarov*, Doctor of Economic Sciences, Associate Professor, Head of the Department of Business Informatics, Ural State University of Economics, 62 Marta 8 St, Yekaterinburg, 620144, Russian Federation. ORCID: 0000-0002-5847-9718; SPIN-code: 2424-3457. E-mail: slup20005@mail.ru

*Svetlana V. Begicheva*, Candidate of Economic Sciences, Associate Professor at the Department of Business Informatics, Ural State University of Economics, 62 Marta 8 St, Yekaterinburg, 62620144, Russian Federation. ORCID: 0000-0002-0551-1622, SPIN-code: 1924-9572. E-mail: begichevas@mail.ru