

Потенциал применения больших языковых моделей для автоматизации контроля знаний в школах и вузах

Пайдершев Алексей Константинович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В статье исследуются возможности интеграции большой языковой модели Gemini от Google в образовательный процесс на примере проектирования тестовых материалов по отечественной истории. Выбор данной платформы, предназначенной для глубокой аналитической обработки информации, обусловлен актуальным запросом на цифровые инструменты, способные автоматически создавать контрольно-оценочные задания закрытого типа на основе учебников. Применение этого метода позволяет существенно оптимизировать временные затраты преподавателя на формирование фондов оценочных средств. В то же время использование экспертной верификации сгенерированных вопросов по критериям классической отечественной тестологии способствует минимизации фактических ошибок ИИ, логическому структурированию заданий и повышению надежности оперативной проверки знаний студентов.

Ключевые слова: Gemini, тестовые задания, нейросети, автоматизация контроля, история России.

The potential of using large language models to automate knowledge assessment in schools and universities

Paydershev Alexey Konstantinovich

Sholom-Aleichem Priamursky State University

Student

Abstract

The article explores the possibilities of integrating Google's large language model, Gemini, into the educational process through the design of test materials in Russian history. The choice of this platform, designed for deep analytical processing of information, is driven by the current demand for digital tools that can automatically generate closed-ended assessment tasks based on textbooks. This method significantly reduces the time required for teachers to create assessment tools. At the same time, the use of expert verification of generated questions based on the criteria of classical Russian testology helps to minimize actual AI errors, logically structure tasks, and increase the reliability of operational testing of students' knowledge.

Keywords: Gemini, test tasks, neural networks, control automation, and Russian history.

1 Введение

1.1 Актуальность

Актуальность исследования обусловлена высокой нагрузкой на преподавателей вузов и школ, тратящих много времени на ручное создание тестов.

Использовать нейросети для подготовки тестов вслепую рискованно из-за фактических ошибок и галлюцинаций ИИ. Модель Gemini в связке с последующей верификацией решает эту проблему, позволяя быстро генерировать качественные вопросы строго на основе учебников. Изучение способов автоматизации разработки тестов с помощью Gemini имеет высокую практическую ценность, помогая разгрузить педагогов и повысить качество оценочных материалов.

1.2 Обзор исследования

Практический потенциал ИИ детально раскрывают Е. Э. Шаповалов и В.С. Шаповалова, доказывая пользу генеративных нейросетевых инструментов как доминирующего фактора оптимизации и повышения эффективности преподавательской деятельности [1]. Траекторию развития технологий дополняет Т. В. Букина, выделяя макротренды современного состояния и долгосрочных перспектив эволюции интеллектуальных систем в дидактике [2]. Комплексный критический анализ Н. А. Иванько предлагает сбалансированную оценку внедрения технологий, исследуя искусственный интеллект сквозь призму социальных инноваций и когнитивных рисков [3]. В сфере профессионального развития кадров А. И. Минаков видит в нейросетях эффективный инструмент и актуальное стратегическое направление подготовки современных педагогов [4]. Исследователи Т. А. Бороненко и В.С.Федотова предлагают систематизацию новых задач генеративного искусственного интеллекта и подчеркивают, что ключевым условием успеха является развитие у педагогов качественно новых цифровых компетенций [5].

1.3 Цель исследования

Цель исследования – теоретически обосновать и экспериментально оценить эффективность применения большой языковой модели Gemini в качестве инструмента автоматизации разработки тестовых материалов для школ и вузов, а также верифицировать качество сгенерированного контента на основе принципов классической отечественной тестологии.

2 Материалы и методы

В качестве эмпирического материала исследования использовался фрагмент учебного текста из профильного университетского курса. Инструментом автоматического анализа и генерации текстового контента выступила большая языковая модель Gemini.

В основу методологии легли принципы классической отечественной тестологии, разработанные В. С. Аванесовым [6], а также современные подходы к промт-инжинирингу без предварительного обучения модели.

3 Результаты

Первый шаг исследования – загрузка и постановка задачи нейросетевой модели Gemini. В качестве источника использовался сертифицированный учебник по истории России за 11 класс В. Р. Мединского и А. В. Торкунова [7].

Промт для генерации теста: «Загруженный файл - учебник по истории России за 11 класс. На основании текста учебника составь контрольный тест по теме «СССР в 1945-1991 гг». Тест должен состоять из 10 вопросов с 4 вариантами ответа (1 - правильный, 3 - неправильных)».

Результат генерации:

Контрольный тест по теме: «СССР в 1945–1991 гг.»

Инструкция: Выберите один правильный ответ из четырех предложенных.

Вопросы:

1. Какая пятилетка вошла в историю СССР как период послевоенного восстановления народного хозяйства?

- а) Третья (1938–1942 гг.)
- б) Четвёртая (1946–1950 гг.)
- в) Пятая (1951–1955 гг.)
- г) Седьмая (1959–1965 гг.)

2. В каком году в СССР была одновременно проведена послевоенная денежная реформа и полностью отменена карточная система распределения продуктов питания?

- а) 1945 г.
- б) 1953 г.
- в) 1947 г.
- г) 1949 г.

3. С каким докладом выступил Н. С. Хрущёв на закрытом заседании XX съезда КПСС в феврале 1956 года?

- а) «О культе личности и его последствиях»
- б) «О задачах послевоенного восстановления»
- в) «О переходе к элементам рыночной экономики»
- г) «О контрольных цифрах развития народного хозяйства»

4. Что из перечисленного являлось характерной чертой экономических и управленческих реформ Н. С. Хрущёва?

- а) Введение частной собственности на землю и поддержка фермерства
- б) Привлечение иностранного капитала через систему концессий
- в) Полная ликвидация государственного планирования экономики
- г) Упразднение центральных отраслевых министерств и создание совнархозов

5. Какое ключевое изменение вводилось в промышленность СССР в ходе экономической реформы А. Н. Косыгина 1965 года?

- а) Передача заводов и фабрик в акционерную собственность трудовых коллективов

б) Ликвидация плановой системы и переход к свободному рыночному ценообразованию

в) Восстановление отраслевых министерств и внедрение элементов хозрасчёта

г) Разрешение свободной купли-продажи тяжелого оборудования частными лицами

6. Какое неофициальное название получил период в истории СССР (середина 1970-х – начало 1980-х гг.), связанный с именем Л. И. Брежнева?

а) Эпоха «великого перелома»

б) Период «застоя»

в) Период «ускорения»

г) Эпоха «военного коммунизма»

7. Какое положение о роли КПСС было официально закреплено в статье 6 Конституции СССР 1977 года?

а) Партия признавалась руководящей и направляющей силой общества, ядром политической системы

б) Партия объявлялась дискуссионным клубом без права участия в управлении

в) В стране легализовывалась многопартийная система с равенством всех сил

г) Члены партии получали исключительное право на приватизацию предприятий

8. Какое международное событие 1962 года привело к самому опасному обострению холодной войны, поставившему человечество на грань ядерного конфликта?

а) Начало затяжной войны во Вьетнаме

б) Берлинский кризис и blockade 1948 года

в) Ввод советских войск в Афганистан

г) Карибский кризис

9. Как называлась внешнеполитическая концепция М. С. Горбачёва, предполагавшая приоритет общечеловеческих ценностей над классовыми интересами?

а) «Доктрина Брежнева»

б) «Новое политическое мышление»

в) «Политика мирного сосуществования»

г) «Доктрина Трумэна»

10. Какое историческое событие, произошедшее в декабре 1991 года, юридически зафиксировало прекращение существования СССР?

а) Августовский путч и выступление ГКЧП

б) Проведение Всесоюзного референдума о сохранении Союза

в) Подписание Беловежских соглашений

г) Принятие Декларации о государственном суверенитете РСФСР

Ключ к тесту (Правильные ответы):

1. б (Четвёртая пятилетка)

2. в (1947 г.)

3. а («О культе личности и его последствиях»)
4. г (Упразднение министерств и создание совнархозов)
5. в (Восстановление отраслевых министерств и хозрасчёт)
6. б (Период «застоя»)
7. а (Партия — руководящая и направляющая сила)
8. г (Карибский кризис)
9. б («Новое политическое мышление»)
10. в (Подписание Беловежских соглашений)

Следующий этап – проверка теста на валидность. Для этого использовалась методика В. С. Аванесова, адаптированная для проверки ИИ-тестов. Согласно методике, каждый вопрос теста должен отвечать нескольким параметрам:

- Фактическая точность – правильный ответ должен соответствовать тексту учебника.
- Качество неправильных ответов – неправильные ответы должны быть логичными и не выглядеть глупо, чтобы студент или ученик не мог угадать правильный ответ.
- Однозначность – под формулировку вопроса должен подходить только один вариант ответа (для тестов с одним правильным вариантом ответа).
- Стиль и логика – вопрос должен быть естественным для темы, по которой составляется тест.

Итоговый вердикт по каждому вопросу делится на три категории:

- Годен – идеальный вопрос.
- Нужна правка – неправильные ответы слишком простые.
- Брак – ИИ перепутал дату или сделал 2 правильных ответа.

Результаты оценки теста представлены в таблице (таб. 1):

Таблица 1. Результаты оценки теста

№ вопроса	Фактическая точность	Качество неправильных ответов	Однозначность	Стиль и логика	Итоговый вердикт
					Годен. Отличная проверка знания хронологии пятилеток.
					Годен. Неправильные ответы – реальные важные даты эпохи.
					Годен. Названия вымышленных докладов звучат очень правдоподобно.
					Годен. Совнархозы – классическая тема для проверки этого периода.

№ вопроса	Фактическая точность	Качество неправильных ответов	Однозначность	Стиль и логика	Итоговый вердикт
					Годен. Элементы хозрасчёта отражены фактологически верно.
					Годен. Все варианты – реальные исторические термины
		+-			Нужна правка. Варианты «в» и «г» слишком очевидно не подходят для 1977 года. Студент может угадать ответ методом исключения, даже не зная Конституцию.
					Годен. Хорошая проверка на знание дат ключевых событий Холодной войны.
					Годен. Терминология подобрана корректно.
					Годен. Все события из вариантов произошли в 1990–1991 гг., что заставляет студента именно думать и вспоминать точный месяц.

Анализ сгенерированного массива вопросов показал, что модель Gemini демонстрирует высокий уровень фактологической точности. Особый интерес представляет качество альтернативных вариантов ответа. В 90% случаев модель не просто выдумывает случайные даты или названия, а использует реальные исторические концепты из смежных периодов. Это усложняет задачу для тестируемого и исключает простое угадывание.

Единственная уязвимость была обнаружена в вопросе №7. Модель включила в варианты ответов понятия «многопартийность» и «приватизация». Данные явления исторически относятся к концу 1980-х и 1990-м годам, что делает их заведомо ложными для студента и снижает ценность конкретного вопроса.

Тем не менее, общая доля валидных вопросов составила 90%, что является высоким показателем для систем без предварительного обучения.

4 Вывод

Таким образом, проведённое исследование демонстрирует, что нейросетевая модель Gemini обладает достаточной точностью составления вопросов для контрольных тестов. Её применение позволит существенно сократить время, затрачиваемое преподавателями на создание итоговых и промежуточных работ, так как весь процесс генерации теста занял меньше 5 минут реального времени, учитывая скачивание учебника и написание промта, когда составителю на это требуется не менее часа.

Библиографический список

1. Шаповалов Е. Э., Шаповалова В. С. Генеративные нейросетевые инструменты как фактор повышения эффективности преподавательской деятельности: анализ современных подходов //International Journal of Open Information Technologies. 2025. Т. 13. №. 12. С. 42-51. URL: <https://cyberleninka.ru/article/n/generativnye-neyrosetevye-instrumenty-kak-faktor-povysheniya-effektivnosti-prepodavatelskoy-deyatelnosti-analiz-sovremennyh>
2. Букина Т. В. Искусственный интеллект в образовании: современное состояние и перспективы развития //Общество: социология, психология, педагогика. 2025. №. 1. С. 76-83. URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-v-obrazovanii-sovremennoe-sostoyanie-i-perspektivy-razvitiya>
3. Иванько Н. А. Искусственный интеллект в образовании: социальная инновация или когнитивная деградация? //Вестник РГГУ. Серия: Литературоведение. Языкознание. Культурология. 2025. №. 7. С. 125-138. URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-v-obrazovanii-sotsialnaya-innovatsiya-ili-kognitivnaya-degradatsiya>
4. Минаков А. И. Искусственный интеллект в образовании—актуальное направление подготовки современных педагогов //Мир науки. Педагогика и психология. 2024. Т. 12. №. 1. С. 56. URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-v-obrazovanii-aktualnoe-napravlenie-podgotovki-sovremennyh-pedagogov>
5. Бороненко Т. А., Федотова В. С. Генеративный искусственный интеллект в образовании: новые задачи и компетенции педагога //Мир науки, культуры, образования. 2025. №. 2 (111). С. 228-233. URL: <https://cyberleninka.ru/article/n/generativnyy-iskusstvennyy-intellekt-v-obrazovanii-novye-zadachi-i-kompetentsii-pedagoga>
6. Аванесов В. С. Композиция тестовых заданий. М.: Адепт, 1998. 219 с. URL: <https://viperson.ru/uploads/attachment/file/949931/avanesov.pdf>
7. Мединский В. Р., Торкунов А. В. История России. 1945 год — начало XXI века: 11-й класс: базовый уровень: учебник. 3-е изд. М.: Просвещение, 2024. 447 с. URL: <https://no.11klasov.net/index.php?do=download&id=26223>