

РАЗРАБОТКА ШАБЛОНА КОРРЕКТНЫХ ЗАПРОСОВ ДЛЯ ГЕНЕРАЦИИ УЧЕБНЫХ ЗАДАНИЙ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Аннотация. Проект посвящен разработке шаблона запроса, который позволяет отправлять корректные запросы к большим языковым моделям для генерации заданий различных форматов, предназначенных для использования в образовательном процессе. Основная цель работы — создание удобного и эффективного инструмента для преподавателей и разработчиков учебных материалов, который поможет автоматизировать процесс подготовки заданий и упражнений, адаптированных под конкретные учебные цели и потребности студентов. Ожидается, что реализация проекта значительно облегчит процесс подготовки учебных материалов, повысит их качество и разнообразие, а также способствует более глубокому и взаимодействию студентов с учебным контентом.

Ключевые слова: генерация заданий, разработка шаблона запросов к большим языковым моделям.

Введение. В настоящее время преподаватели сталкиваются с проблемой создания тестовых заданий для проверки знаний студентов, ведь это требует значительных усилий и времени. В частности, возникла необходимость в автоматизированных методах проверки знаний. Решение данной проблемы позволит преподавателям тратить меньше времени на создание тестовых заданий.

В статье Загорской А. В. рассматривался вопрос о применении больших языковых моделей в научных исследованиях [1]. В статье говорится, что примером такого инструмента является GPT (Generative Pre-trained Transformer), чат-бот, который стоит в ряду с самыми продвинутыми и эффективными алгоритмами для создания текста.

Однако существенным аспектом использования искусственного интеллекта, а именно ChatGPT, является формулировка корректного запроса, об этом говорится в статье Захаровой И. Г. [2]. Это важно, поскольку корректность полученных результатов напрямую зависит от точности заданного вопроса. В научной работе [3] утверждается, что корректный запрос к ChatGPT — это четко сформулированный вопрос или команда, направленная к модели языка GPT, который используется для получения информации, совета или выполнения определенной текстовой задачи. А корректный результат — это ответ, который полностью соответствует заданному вопросу или проблеме, учитывая контекст и требования к ответу, без дополнительных уточнений и дополнительных запросов. В работе [4] автор утверждает, что, чем более конкретен и точен запрос, тем более релевантные и полезные ответы можно получить от чат-бота. Программное обеспечение, использующее модель GPT, обладает способностью автоматического генерирования учебных заданий различного уровня сложности — от базовых тестов до комплексных задач, требующих аналитического подхода и специализированных знаний.

Таким образом, использование технологий искусственного интеллекта и машинного обучения в образовании открывает новые перспективы для улучшения качества образования и эффективности процесса оценки знаний студентов.

Проблема исследования. Следовательно, целью данного исследования является разработка шаблона запроса с целью получения корректного ответа от модели GPT для автоматизации генерации заданий разного формата.

Для достижения этой цели необходимо исследовать функциональность ChatGPT, а также разработать структуру шаблона для формирования корректных запросов к ChatGPT, способствующего получению качественных и образовательно значимых заданий, по курсу «Программирование».

Постановка задачи.

Дано Q — множество запросов, где q_j — запрос (текст), $q_j \in Q$,

m — количество запросов, $m \in M$, $j \in [1, m]$.

P — множество промтов, где p_i — это параметр (часть текста), влияющий на исход ответа, здесь n — это количество параметров, $n \in N$, $i \in [1, n]$.

Каждый q_j может состоять из разнообразных комбинаций из множества P , то есть запрос может состоять из любого набора промтов: $C_n^1 + \dots + C_n^n$ комбинаций из P .

A — это множество ответов на запросы, где a_i — ответ на запрос (текст), $a_i \in A$, k — количество ответов, $k \in K$, $j \in [1, k]$. Для каждого q_j существует $C_n^1 + \dots + C_n^n$ ответов a_i , где a_i имеет оценку релевантности g_i , $g_i \in [1, 5]$, где 1 — ответ полностью не соответствует запросу, а 5 — ответ полностью соответствует запросу.

Найти для любого заданного запроса $q_j = \{p_1, \dots, p_n\}$ фиксированный набор промтов p_1, \dots, p_n и h — количество значимых промтов, таким образом, чтобы оценка релевантности g была максимальна для a_i .

Материалы и методы

Метод оценки влияния промтов на ответ

Пусть запрос q_j , а также промты, из которых состоит $q_j = \{p_1, \dots, p_n\}$ определены, и даны оценки релевантности соответствия запроса ответу g_j . Предположим, что связь между q_j и g_j является линейной, а также существует некоторая случайная ошибка (1):

$$g_j = \langle q_j, w \rangle + \varepsilon, \quad (1)$$

где $\langle q_j, w \rangle$ — стандартное скалярное произведение $(b_1 w_1 + \dots + b_n w_n)$,

ε — случайная ошибка,

b — бинарное значение определяющие существование промта,

w — коэффициент регрессии каждого составляющего параметра запроса.

Для оценки того, насколько хорошо данные исследуются моделью множественной регрессии, используется коэффициент детерминации.

Возможности модели ChatGPT

Модель ChatGPT была разработана на основе архитектуры трансформеров, концепции, первоначально предложенной учеными из Google. Трансформеры являются методом глубокого обучения, который использует механизм внимания, или attention mechanism, для более точного определения значимых частей входных данных при формировании ответа [5]. Этот подход, также известный как контекстуальное моделирование, позволяет модели лучше понимать семантические и синтаксические связи во вводимом тексте и создавать более естественные и точные ответы [6].

Технология GPT, основанная на принципах машинного обучения и обучения с подкреплением, способна анализировать большие объемы текстовых данных и генерировать на их основе собственные тексты. Это открывает новые горизонты для создания образовательного контента и автоматизации процесса подготовки учебных материалов [7].

Применение GPT может быть особенно релевантно для студентов и преподавателей информационных технологий, которые регулярно сталкиваются с комплексными задачами программирования, и оперативной обратной связи [8].

Проектирование и разработка шаблона запросов

Был проведен детальный анализ параметров, которые могут быть использованы при формировании запросов к ChatGPT. Это включало изучение влияния различных параметров на качество и релевантность получаемых ответов. В ходе анализа были выделены ключевые параметры, которые необходимо учитывать при разработке шаблона запросов. Особое внимание было уделено разработке шаблона на языке программирования Python, что позволило автоматизировать процесс генерации запросов и обеспечить высокую степень гибкости и масштабируемости.

Для реализации задачи был создан специализированный скрипт на Python, который использует принципы объектно-ориентированного программирования для эффективного управления параметрами запросов. Скрипт включает в себя классы и методы, которые позволяют динамически формировать запросы на основе заданных пользователем критериев и предпочтений. Это включает возможность задавать ключевые слова, сложность вопроса, и другие параметры, что делает процесс генерации запросов максимально гибким и адаптируемым к конкретным потребностям пользователя.

Составленные запросы были разработаны таким образом, чтобы охватить широкий спектр возможных вариаций и условий, с которыми пользователи могут столкнуться при общении с ИИ — системами.

В процессе анализа данных мы выделили шесть ключевых параметров, которые, по нашему мнению, играют решающую роль в определении качества и релевантности ответов, получаемых от системы (табл. 1).

Таблица 1

Разметка данных для проведения линейного регрессионного анализа

| Номер запроса | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ | Y |
|---------------|----------------|----------------|----------------|----------------|----------------|----------------|-----|
| 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| 2 | 0 | 1 | 1 | 0 | 0 | 0 | 3 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 150 | 1 | 1 | 0 | 1 | 1 | 0 | 5 |

Ответы Y оценены по пятибалльной шкале:

- ответ полностью не соответствует запросу — 1;
- ответ содержит некоторые подходящие элементы — 2;
- ответ удовлетворяет половине критериев — 3;
- ответ почти полностью соответствует запросу — 4;
- ответ полностью соответствует запросу — 5.

Параметры запросов были представлены в виде бинарных переменных.

Коэффициент детерминации, является статистической мерой, которая используется для оценки насколько хорошо данные исследуются моделью множественной регрессии.

Он представляет собой долю дисперсии зависимой переменной Y , которая объясняется независимой переменной в модели регрессии (2).

$$R^2 = \frac{\sum_{i=0}^n (q_i - g_i)^2}{\sum_{i=0}^n (q_i - \bar{q}_i)^2}, \quad (2)$$

где q_i — наблюдаемые значения, g_i — модельные (предсказанные) значения, \bar{q}_i — среднее значение наблюдаемых значений.

Результаты анализа следующие:

- X_1 — коэффициент регрессии — 0.29;
- X_2 — коэффициент регрессии — 1.19;
- X_3 — коэффициент регрессии — 1.05;
- X_4 — коэффициент регрессии — 1.1;
- X_5 — коэффициент регрессии — 1.05;
- X_6 — коэффициент регрессии — 0.5.

Все параметры оказывают влияние на зависимую переменную. Это подчеркивает важность каждого из рассмотренных параметров для получения ожидаемого ответа от ChatGPT, и подтверждает гипотезу о том, что правильно составленный промт способствует получению качественного результата.

В основе запроса лежали шесть ключевых компонентов: задание, контекст, уточнение, роль, форма и настрой:

1. Задание. Основное правило — всегда инициировать запрос с конкретного действия (создавать, управлять, составлять, анализировать и т.д.) и ясно определять итоговую цель.

2. Контекст. Контекст имеет ключевое значение для успешного взаимодействия с GPT или любым другим ИИ.

3. Уточнение. Примеры значительно способствуют обучению алгоритма, особенно когда задание сложное или требует уникального формата ответа. Они показывают ожидаемый исход, что улучшает понимание моделью поставленной задачи.

4. Роль. Позиция определяет метод взаимодействия модели пользователем. Модель может выступать в роли советника, специалиста, наставника и т.д.

5. Формат. Формат устанавливает структуру ответа. Это может быть перечень, таблица, текстовый абзац или даже поэма. В контексте ИИ форма может принимать вид кода, схемы, графика или математического выражения.

6. Настрой. Настрой отражает эмоциональный окрас и стиль ответа. Он может быть официальным, неофициальным, юмористическим, серьезным и т.д. В контексте ИИ настрой может быть научным, пояснительным, критическим или восторженным.

Пример сконструированного шаблона запросов:

Роль: Ты профессор в университете.

Контекст: Преподаешь дисциплину «Программирование на Python».

Задача: Нужно разработать комплекс тестовых заданий для проверки знаний с акцентом на уровень сложности *средний* по теме *работа со строками*.

Уточнения: Необходимо предусмотреть задания про: *Реализация функций для обработки строк: переворот строки, подсчет определенного символа в строке, разделение строки на слова*.

Формат: Для заданий с выбором ответа ответы должны быть представлены в виде списка вариантов, где правильный ответ будет отмечен точкой в начале. Это поможет структурировать ответы и облегчит проверку.

Структура Теста:

- Первая часть: *10 тестовых заданий* с выбором ответа. Каждое задание должно иметь один правильный ответ из предложенных вариантов,
- Вторая часть: *2 задания с открытым ответом.*

| | |
|--|---|
| <p>**Часть 1: Задания с выбором ответа**</p> <p>1. Какая функция в Python используется для переворота строки?</p> <ul style="list-style-type: none">- A. reverse()- B. flip()- *C. [::-1]- D. revert() <p>2. Какой метод позволяет подсчитать количество вхождений определенного символа в строке?</p> <ul style="list-style-type: none">- A. countChar()- *B. count()- C. findChar()- D. search() <p>3. Какой метод разделяет строку на список слов?</p> <ul style="list-style-type: none">- A. splitWords()- *B. split()- C. divide()- D. cut() <p>4. Что делает следующий код: list("Python")?</p> <ul style="list-style-type: none">- A. Создает список из слова "Python"- *B. Преобразует строку "Python" в список ['P', 'y', 't', 'h', 'o', 'n']- C. Генерирует ошибку- D. Ничего из перечисленного <p>5. Какой результат выполнения следующего кода: ", ".join(["Python", "Programming"])?</p> <ul style="list-style-type: none">- A. "Python Programming"- B. ["Python", "Programming"]- *C. "Python Programming"- D. "Python-Programming" | <p>7. Что возвращает следующий код: "Python".find("t")?</p> <ul style="list-style-type: none">- A. 't'- *B. 2- C. True- D. -1 <p>8. Какой метод удаляет пробельные символы в начале и конце строки?</p> <ul style="list-style-type: none">- *A. strip()- B. trim()- C. cut()- D. delete() <p>9. Какой из следующих способов является правильным для объявления функции в Python?</p> <ul style="list-style-type: none">- A. function myFunction():- *B. def myFunction():- C. function: myFunction()- D. create myFunction(): <p>10. Какое ключевое слово используется для генерации исключений в Python?</p> <ul style="list-style-type: none">- A. throw- *B. raise- C. error- D. exception |
|--|---|

Рис. 1. Ответ, полученный от ChatGPT

Как можно заметить на рис. 1, комбинирование этих шести компонентов в запросе гарантирует, что ответы от GPT будут полностью соответствовать заданному запросу.

Результаты. На основе изученных материалов был разработан общий шаблон запросов, предназначенный для генерации образовательных заданий различного формата. Шаблон включает в себя определение структуры запроса таким образом, чтобы максимизировать вероятность получения корректного и содержательного ответа от модели.

После тщательного проектирования и разработки шаблона запросов, была проведена серия экспериментов для оценки его эффективности. Используя специализированный скрипт на Python, был проведен процесс генерации запросов к ChatGPT, что позволило тестировать различные комбинации параметров и оценивать качество получаемых ответов.

Для оценки эффективности шаблона запросов было сформировано 150 уникальных запросов, охватывающих разнообразные темы и сложность. Каждый запрос был направлен в ChatGPT, и полученные ответы оценивались на основе предварительно определенной пяти-балльной шкалы, отражающей степень соответствия ответа запросу. Коэффициенты регрессии

для ключевых параметров запросов подтвердили их значимое влияние на качество ответов. Наибольшее влияние оказывали такие параметры, как задача и уточнение. Все параметры запросов оказались значимыми, что доказывает необходимость комплексного подхода к формулировке запросов для достижения оптимальных результатов.

Интеграция такого шаблона в образовательные платформы и системы управления обучением (LMS) может значительно упростить создание контента, автоматизировать рутинные процессы и повысить качество обучения.

Кроме того, использование структурированных запросов облегчает аналитику данных об обучении, позволяя образовательным учреждениям и онлайн-платформам лучше понимать потребности своих пользователей. Это создает основу для постоянного улучшения качества образования.

Заключение. Разработанный шаблон запросов объединил методы искусственного интеллекта с потребностями преподавателей и доказал свою эффективность в создании тестовых заданий через исследование функциональных возможностей чата-GPT.

Шаблон запросов охватывает несколько ключевых аспектов: формулировку четкого и конкретного вопроса и, при необходимости, предоставление дополнительных уточнений или параметров для более точного ответа. Таким образом, шаблон предусматривает гибкость и адаптивность к различным темам и уровням сложности заданий, обеспечивая эффективное взаимодействие с моделью.

Для автоматизации и упрощения процесса создания образовательно значимых заданий был разработан скрипт на Python 3.10. Использование библиотек Flask и Requests позволило интегрировать шаблон запросов в веб-приложение путем написания API, обеспечивая удобный интерфейс для преподавателей и возможность мгновенного получения готовых заданий. Таким образом, разработанный шаблон запросов стал не только эффективным инструментом для создания тестовых заданий, но и основой для создания полноценного программного решения, способного значительно упростить и автоматизировать процесс подготовки учебных материалов.

В дальнейшем планируется:

- Расширение функционала разработанного шаблона, добавление более глубоких уточнений.
- Внедрение шаблона в web-приложение.
- Интеграция с различными образовательными платформами.

Это позволит сократить время преподавателям для подготовки тестов, и сделать процесс обучения более гибким и индивидуализированным.

СПИСОК ЛИТЕРАТУРЫ

1. Загорская А.В. Применение GPT-чата в научном исследовании: возможности и ограничения / А.В. Загорская // Строительное производство. — 2023. — № 3. — С. 41-48. — Текст: непосредственный.
2. Захарова И.Г. Большие языковые модели в образовании: создание тестовых заданий с помощью ChatGPT / И.Г. Захарова // Информатизация образования и методика электронного обучения: цифровые технологии в образовании: материалы VII Международной научной конференции. — Красноярск, 2023. — С. 1100-1104. — Текст: непосредственный.

3. Liu P., Yuan W., Fu J., Jiang Z., Hayashi H., Neubig G. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing // ACM Computing.
4. Плахотнюк О.С. Особенности разработки вариативных тестовых заданий с использованием чат-бота с искусственным интеллектом ChatGPT в обучении будущих педагогов иностранному языку / О.С. Плахотнюк // Информатизация образования и методика электронного обучения: цифровые технологии в образовании. — Москва, 2023. — С. 107-115. — Текст: непосредственный.
5. Моргунова А. Исследование эффективности использования ChatGPT в обучении программированию / А. Моргунова // Материалы II Всероссийской научно-практической конференции «Современные цифровые технологии». — Барнаул, 2023. — С. 307-311. — Текст: непосредственный.
6. Pyko A., Sokolovskaya Y., Stolyarova E., Provotorova E. ChatGPT and its influence on modern business / А. Pyko, Y. Sokolovskaya, E. Stolyarova, E. Provotorova // Бизнес. Образование. Экономика. Материалы Международной научно-практической конференции. — Минск, 2023. — С. 115-118. — Текст: непосредственный.
7. Хасбатов Д.Ш., Халиулин, И.М. Возможности использования искусственного интеллекта чат-ботов на базе GPT в образовательном процессе / Д.Ш. Хасбатов, И.М. Халиулин // Материалы II Всероссийской научно-практической конференции «Современные цифровые технологии». — Барнаул, 2023. — С. 77-82. — Текст: непосредственный.
8. Educational AI Chatbots for Content and Language Integrated Learning // MDPI: [сайт]. — URL: <https://www.mdpi.com/2076-3417/12/7/3239> (дата обращения: 10.10.2023).