

Мечта ленивого студента: ChatGPT самостоятельно сдает курс инженерии

Дата: 2025-02-23 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2503.05760>

Рейтинг: 85

Адаптивность: 90

Ключевые выводы:

Основная цель исследования - оценить способность языковой модели ChatGPT успешно пройти семестровый курс по системам управления в аэрокосмической инженерии (AE 353) с минимальными усилиями. Исследование показало, что ChatGPT смог получить оценку B (82,24%), что близко, но не превышает средний балл класса (84,99%). LLM показала наилучшие результаты в структурированных заданиях и имела наибольшие ограничения в открытых проектах.

Объяснение метода:

Исследование предоставляет непосредственно применимые методы промптинга и четкое понимание возможностей LLM в решении технических задач. Результаты показывают, где LLM эффективны (структурированные задания) и где ограничены (открытые проекты), что помогает пользователям формировать оптимальные стратегии использования. Методология "минимальных усилий" и добавление контекста универсально применимы в любых образовательных и профессиональных сценариях.

Ключевые аспекты исследования: 1. **Всестороннее тестирование LLM в образовательном контексте:** Исследование оценивает способность ChatGPT (GPT-4) пройти полный семестровый курс по системам управления в аэрокосмической инженерии, используя 115 различных учебных заданий.

Методология "минимальных усилий": Авторы симулировали сценарий, когда студент использует LLM с минимальными затратами времени и усилий, просто копируя и вставляя задания в ChatGPT без дополнительных указаний.

Разнообразные методы оценки: Тестирование проводилось с использованием трех подходов к промптингу: на основе изображений, с упрощенной математической нотацией и с добавлением контекста из лекционных материалов.

Количественные результаты по категориям заданий: LLM достиг общего балла 82,24% (оценка B), приближаясь к среднему показателю класса (84,99%), с

наилучшими результатами в структурированных заданиях и наибольшими ограничениями в открытых проектах.

Рекомендации по адаптации курсов: Исследование предлагает переосмыслить стратегии оценки в инженерном образовании, делая акцент на интегрированной проектной работе, объяснении рассуждений и задачах, требующих практического суждения.

Дополнение:

Применимость методов в стандартном чате

Исследование **не требует** дообучения или специального API для применения основных методов. Все подходы могут быть реализованы в стандартном чате с LLM, так как авторы намеренно использовали общедоступную версию ChatGPT (GPT-4).

Ключевые концепции для адаптации

Контекстное обогащение промптов: Добавление релевантного материала из лекций/учебников перед заданием значительно повышает качество ответов (из 82.24% общего балла с контекстом против более низких результатов без контекста).

Стратегии для различных типов задач:

Для структурированных заданий (MCQ, числовые задачи): прямой запрос работает хорошо
Для сложных математических задач: упрощение нотации повышает точность
Для программирования: разбиение задачи на последовательные части

Многошаговый промптинг: Использование обратной связи для уточнения ответов (особенно эффективно для MCQ, где успешность выросла с 89.5% до 96.5% с контекстом)

Ожидаемые результаты

- Повышение эффективности на ~5-10% для структурированных заданий
- Существенное улучшение для математических задач при добавлении контекста
- Понимание ограничений для открытых проектов, где даже с оптимальными промптами LLM достигает только ~65% эффективности

Анализ практической применимости: 1. **Всестороннее тестирование LLM в образовательном контексте** - Прямая применимость: Пользователи могут использовать описанные подходы к промптингу для решения учебных задач по инженерным дисциплинам. - Концептуальная ценность: Исследование показывает, где LLM справляется хорошо (структурированные задачи), а где хуже (открытые проекты), что помогает пользователям понять ограничения моделей. - Потенциал для адаптации: Понимание различий в эффективности разных типов промптов

может быть применено к любым техническим предметам.

Методология "минимальных усилий" Прямая применимость: Пользователи могут применять простой подход к использованию LLM, не требующий специальных знаний или сложного инструктирования. Концептуальная ценность: Демонстрирует, что даже с минимальными усилиями LLM может достигать хороших результатов, но также показывает пределы такого подхода. Потенциал для адаптации: Методология может быть использована для оценки эффективности LLM в любых других образовательных контекстах.

Разнообразные методы оценки

Прямая применимость: Пользователи могут выбирать между тремя описанными методами промптинга в зависимости от типа задачи. Концептуальная ценность: Понимание того, что добавление контекста из лекционных материалов значительно улучшает результаты работы LLM. Потенциал для адаптации: Методы промптинга могут быть адаптированы для любых образовательных задач, включая нетехнические дисциплины.

Количественные результаты по категориям заданий

Прямая применимость: Пользователи могут понять, для каких типов заданий LLM наиболее эффективны (MCQ, автоматически оцениваемые задания) и наименее эффективны (проектные работы). Концептуальная ценность: Показывает закономерности в успешности LLM в зависимости от типа задачи, что может помочь в формировании ожиданий. Потенциал для адаптации: Эти закономерности могут быть экстраполированы на другие области знаний.

Рекомендации по адаптации курсов

Прямая применимость: Преподаватели и учащиеся могут использовать эти рекомендации для более эффективного внедрения LLM в образовательный процесс. Концептуальная ценность: Предлагает переосмыслить оценку в инженерном образовании, фокусируясь на задачах, которые LLM выполняет хуже человека. Потенциал для адаптации: Рекомендации могут быть применены к разработке курсов в различных технических дисциплинах.

Prompt:

Применение знаний из исследования ChatGPT в инженерном образовании ##
Ключевые инсайты из исследования

Исследование показало, что ChatGPT может достигать хороших результатов в инженерных курсах (оценка B, 82.24%), но его эффективность значительно варьируется в зависимости от типа заданий и способа формулировки запросов.

Пример улучшенного промпта на основе исследования

[=====] # Запрос по системам управления в аэрокосмической инженерии

Контекст [Вставить краткое содержание соответствующей лекции/материала]

В системах стабилизации полета используются пропорционально-интегрально-дифференциальные (ПИД) контроллеры для поддержания заданной траектории. Основное уравнение ПИД-контроллера: $u(t) = K_p e(t) + K_i \int e(t)dt + K_d de(t)/dt$, где $e(t)$ - ошибка отклонения от целевого состояния.

Задача Разработайте ПИД-контроллер для стабилизации беспилотного летательного аппарата при боковом ветре. Система описывается следующим уравнением состояния: $\dot{x} = Ax + Bu$ $y = Cx + Du$

где матрицы: $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$ $D = \begin{bmatrix} 0 \end{bmatrix}$

Вопросы (отвечайте последовательно) 1. Определите передаточную функцию системы 2. Проведите анализ устойчивости системы 3. Подберите параметры ПИД-контроллера для обеспечения времени установления менее 5 секунд 4. Напишите код на Python для моделирования поведения системы с разработанным контроллером [=====]

Почему этот промпт эффективен согласно исследованию

Контекстное обогащение: Исследование показало, что добавление контекста из лекций значительно улучшает качество ответов. Промпт включает релевантную теоретическую информацию.

Текстовая математическая нотация: Вместо изображений с формулами использована упрощенная текстовая запись математических выражений, что повышает точность обработки.

Структурированный подход: Задача разбита на последовательные шаги, что помогает преодолеть ограничения модели в системной интеграции сложных проектов.

Четкая формулировка: Промпт содержит конкретные требования и ожидаемые результаты, что снижает вероятность получения расплывчатых или неточных ответов.

Практическое применение результатов исследования

- Для технических и инженерных задач всегда предоставляйте контекст перед вопросом
- При работе с математикой используйте текстовую нотацию вместо изображений

- Разбивайте сложные задачи на последовательные подзадачи
- Будьте критичны к ответам, особенно в открытых проектных заданиях
- Используйте итеративный подход с обратной связью для улучшения результатов

Такой подход к формулировке запросов позволит максимально использовать возможности ChatGPT для решения инженерных задач, учитывая выявленные в исследовании сильные и слабые стороны модели.