

ПОПИШИ: Структурированное рассуждение Больших Языковых Моделей с экстраполяцией достоверности, вдохновленной графами знаний

Дата: 2025-02-08 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2410.08475>

Рейтинг: 70

Адаптивность: 75

Ключевые выводы:

Исследование представляет новый метод GIVE (Graph Inspired Veracity Extrapolation) для улучшения рассуждений больших языковых моделей (LLM) путем объединения параметрической и непараметрической памяти. Основная цель - повысить точность рассуждений LLM с минимальным внешним вводом. Результаты показывают, что GIVE значительно улучшает производительность LLM разных размеров, позволяя даже меньшим моделям превосходить более крупные в научных задачах.

Объяснение метода:

GIVE предлагает мощный метод структурированного рассуждения с использованием ограниченной внешней информации. Хотя полная реализация технически сложна, ключевые концепции (разбиение запроса, экстраполяция на основе ограниченных фактов, контрфактуальное рассуждение) могут быть адаптированы обычными пользователями для улучшения взаимодействия с LLM и получения более достоверных ответов в сложных областях знаний.

Ключевые аспекты исследования: 1. **Структурированное рассуждение с графовым подходом:** Исследование представляет метод GIVE (Graph-Inspired Veracity Extrapolation), который объединяет параметрическую память LLM с непараметрическими знаниями для улучшения рассуждений в задачах, требующих специализированных знаний.

Экстраполяция достоверности: Метод не просто извлекает информацию из внешних источников, а использует ограниченные экспертные данные как отправную точку для дивергентного мышления, позволяя LLM связывать запрос с неполной информацией.

Многоэтапное структурированное рассуждение: GIVE создает группы связанных сущностей, устанавливает внутригрупповые и межгрупповые связи, а также

определяет промежуточные сущности для многошагового рассуждения.

Контрфактуальное рассуждение: Метод включает проверку потенциальных связей, отбрасывая неверные, что помогает избегать галлюцинаций модели при недостаточности знаний.

Прогрессивная генерация ответов: GIVE использует поэтапный подход к формированию ответа, сначала с утвердительными знаниями, затем с контрфактуальными, и наконец с экспертными знаниями.

Дополнение: Исследование GIVE не требует обязательного дообучения или специального API для своей работы, это метод инференса, который может быть адаптирован для стандартного чата. Авторы использовали стандартные LLM (GPT-3.5, GPT-4, Llama 3) без дообучения, просто направляя их с помощью специальных промптов.

Концепции и подходы, которые можно применить в стандартном чате:

Структурированное разбиение вопроса - пользователь может попросить модель выделить ключевые понятия и отношения в вопросе перед ответом.

Формирование групп связанных концепций - можно предложить модели сначала перечислить связанные концепции для каждого ключевого понятия.

Двухэтапное рассуждение - сначала установить связи внутри групп концепций, затем между группами.

Контрфактуальная проверка - попросить модель не только подтвердить возможные связи, но и опровергнуть неверные.

Прогрессивное формирование ответа - сначала получить предварительный ответ, затем уточнить его с учетом дополнительных соображений.

Пример адаптации: при ответе на медицинский вопрос пользователь может сначала попросить модель выделить ключевые термины, затем для каждого термина перечислить связанные понятия, установить связи между ними, проверить потенциальные утверждения и сформировать итоговый ответ. Это позволит получить более структурированное и достоверное рассуждение даже без доступа к графам знаний.

Результаты: значительное улучшение качества ответов в сложных областях знаний, снижение галлюцинаций, более прозрачное рассуждение, которое пользователь может проследить и проверить.

Анализ практической применимости: **Структурированное рассуждение с графовым подходом** - Прямая применимость: Средняя. Обычным пользователям сложно реализовать полную архитектуру, но принцип структурирования запроса на компоненты может быть адаптирован. - Концептуальная ценность: Высокая.

Понимание того, что LLM могут эффективно рассуждать при наличии даже ограниченной внешней информации, поможет пользователям формировать запросы с опорой на достоверные факты. - Потенциал адаптации: Высокий. Можно упростить до включения в промпт ключевых фактов и указания связей между ними.

Экстраполяция достоверности - Прямая применимость: Средняя. Сложна для ручной реализации, но идея использования ограниченных фактов как отправной точки для рассуждений применима. - Концептуальная ценность: Очень высокая. Помогает понять, что LLM могут генерировать правдоподобные рассуждения даже с минимальной внешней информацией. - Потенциал адаптации: Высокий. Пользователи могут предоставлять модели небольшой набор достоверных фактов и просить её рассуждать на их основе.

Многоэтапное структурированное рассуждение - Прямая применимость: Низкая. Сложная техническая реализация, требующая программирования. - Концептуальная ценность: Высокая. Показывает важность разбиения сложных запросов на компоненты и их последовательное связывание. - Потенциал адаптации: Средний. Пользователи могут адаптировать принцип пошагового рассуждения в своих промтах.

Контрфактуальное рассуждение - Прямая применимость: Средняя. Техника проверки утверждений может быть адаптирована для промтов. - Концептуальная ценность: Высокая. Демонстрирует, как проверка противоположных утверждений помогает избежать галлюцинаций. - Потенциал адаптации: Высокий. Пользователи могут включать в промты как утверждения, так и опровержения.

Прогрессивная генерация ответов - Прямая применимость: Средняя. Требует многоэтапного взаимодействия с LLM. - Концептуальная ценность: Высокая. Показывает преимущество итеративного подхода к формированию ответа. - Потенциал адаптации: Высокий. Пользователи могут применять последовательное уточнение ответов.

Prompt:

Использование методологии GIVE в промтах для GPT ## Основные принципы GIVE

Методология GIVE (Graph Inspired Veracity Extrapolation) предлагает структурированный подход к рассуждениям, который объединяет параметрическую и непараметрическую память для улучшения точности ответов языковых моделей.

Пример промпта, вдохновленного GIVE

[=====] Я хочу, чтобы ты помог мне разобраться в теме [ТЕМА] используя структурированный подход к рассуждению.

Следуй этим шагам:

Выдели 3-5 ключевых концепций из этой темы. Для каждой концепции определи группу тесно связанных понятий (2-3 понятия). Для каждой группы опиши внутренние связи между понятиями, используя свои базовые знания. Установи логические связи между разными группами понятий. На основе этой структуры знаний, сформулируй последовательное объяснение темы [ТЕМА]. Представь результат в виде: - Сначала - список ключевых концепций - Затем - группы связанных понятий с их внутренними связями - Далее - межгрупповые связи - И наконец - итоговое объяснение темы [=====]

Как работают принципы GIVE в этом промпте

Извлечение ключевых концепций - промпт просит модель идентифицировать основные элементы темы **Построение групп связанных сущностей** - модель формирует кластеры связанных понятий **Индукция внутригрупповых связей** - модель использует свои параметрические знания для описания отношений внутри групп **Экстраполяция достоверности** - установление межгрупповых связей помогает модели проверить согласованность своих знаний **Прогрессивная генерация ответа** - финальное объяснение строится на основе структурированного графа знаний Этот подход помогает: - Уменьшить галлюцинации модели - Сделать рассуждения более логичными и последовательными - Улучшить точность в специализированных областях знаний - Получить более структурированный и обоснованный ответ

Даже если модель не обладает полными знаниями по теме, такая структура помогает ей лучше организовать имеющуюся информацию и выявить пробелы в рассуждениях.