

Концептуально-ориентированное побуждение цепочки мыслей для парного сравнительного оценки текстов с использованием больших языковых моделей

Дата: 2025-01-24 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2310.12049>

Рейтинг: 82

Адаптивность: 90

Ключевые выводы:

Исследование представляет новый фреймворк для оценки текстов с использованием больших языковых моделей (LLM), который позволяет эффективно анализировать латентные концепции в текстах. Основная цель - создание метода, который преобразует попарные сравнения текстов из задачи рассуждения в задачу распознавания паттернов с помощью подхода Concept Guided Chain of Thought (CGCoT). Результаты показывают, что этот метод превосходит существующие неконтролируемые методы оценки текста и сопоставим с контролируруемыми подходами, требующими значительно больше размеченных данных.

Объяснение метода:

Исследование предлагает практический метод анализа текстов с помощью LLM, который не требует больших размеченных данных. CGCoT-подход (поэтапные направленные вопросы) и попарные сравнения легко адаптируются для различных задач и доступны широкой аудитории. Метод показывает высокую эффективность при минимальных затратах на разработку, хотя полная реализация требует некоторых технических знаний.

Ключевые аспекты исследования: 1. **Концепция CGCoT (Concept-Guided Chain-of-Thought)** - авторы предлагают новый подход к оценке текстов с использованием LLM, где модель анализирует тексты через серию последовательных вопросов, разработанных исследователем, для выделения конкретных аспектов интересующего концепта.

Попарное сравнение текстов - вместо прямой оценки текстов по шкале, авторы используют попарные сравнения между "концептуальными разбивками" текстов, превращая сложную задачу рассуждения в задачу распознавания паттернов.

Модель Брэдли-Терри - для преобразования результатов попарных сравнений в

числовые оценки используется вероятностная модель, которая позволяет ранжировать тексты по степени выраженности целевого концепта.

Применение к оценке политической неприязни - методология была применена для измерения степени неприязни к политическим партиям в твитах, показав высокую корреляцию с оценками людей и превзойдя другие неконтролируемые методы анализа текста.

Минимальная потребность в размеченных данных - метод требует лишь небольшой пилотный набор размеченных примеров для разработки CGCoT-промптов, в отличие от традиционных методов, требующих тысячи размеченных примеров.

Дополнение:

Применимость в стандартном чате без дообучения или API

Не требуется дообучение или API: Методы исследования полностью применимы в стандартном чате с LLM. Хотя авторы использовали GPT-3.5 через API для автоматизации процесса, сама методология CGCoT и попарных сравнений может быть реализована через обычный интерфейс чата.

Ключевые концепции для применения в стандартном чате

Поэтапное структурирование запросов (CGCoT): Разбивка сложной задачи на последовательность простых вопросов. Использование ответов на предыдущие вопросы как контекст для последующих. Пример: При анализе текста сначала попросить LLM резюмировать его, затем выделить ключевые объекты, затем определить отношение к этим объектам.

Попарное сравнение вместо прямой оценки:

Просить LLM сравнивать два текста по определенному критерию вместо прямой оценки по шкале. Это соответствует сильным сторонам LLM (распознавание паттернов) и минимизирует слабости (прямая количественная оценка).

Использование "концептуальных разбивок":

Создание подробного анализа текста с помощью серии вопросов перед сравнением. Это переводит задачу из области рассуждения в область распознавания паттернов.

Ожидаемые результаты при применении в стандартном чате

- Повышение точности анализа текстов по сложным концептам
- Возможность работы с короткими текстами (твиты, комментарии)
- Минимизация потребности в размеченных данных

- Более последовательные и обоснованные оценки текстов

Пример использования в стандартном чате

Для анализа эмоциональной окраски отзыва о продукте: 1. "Резюмируй этот отзыв." 2. "Какие аспекты продукта упоминаются в отзыве?" 3. "Для каждого аспекта определи, выражено ли положительное или отрицательное отношение." 4. "Исходя из предыдущих ответов, насколько сильно в отзыве выражено [целевое отношение]?"

Затем можно сравнивать результаты анализа разных отзывов попарно.

Анализ практической применимости: Концепция CGCoT - Прямая применимость: Высокая. Пользователи могут адаптировать этот подход для анализа любых текстов, разрабатывая собственные серии вопросов для выделения интересующих концептов. - **Концептуальная ценность:** Очень высокая. Метод демонстрирует, как разбивка сложной задачи на более простые подзадачи может значительно улучшить производительность LLM. - **Потенциал для адаптации:** Исключительно высокий. Принцип поэтапного анализа может быть применен для любых концептов и доменов.

Попарное сравнение текстов - Прямая применимость: Средняя. Требуется некоторого технического понимания для реализации, но базовая идея доступна широкому кругу пользователей. - **Концептуальная ценность:** Высокая. Показывает, что LLM лучше справляются с бинарным выбором, чем с прямой оценкой по шкале. - **Потенциал для адаптации:** Высокий. Может использоваться для ранжирования текстов по любым критериям.

Модель Брэдли-Терри - Прямая применимость: Низкая для неспециалистов, требует понимания статистических методов. - **Концептуальная ценность:** Средняя. Полезна для понимания, как трансформировать качественные сравнения в количественные оценки. - **Потенциал для адаптации:** Средний. Требуется определенных технических навыков.

Применение к оценке политической неприязни - Прямая применимость: Высокая для аналитиков социальных сетей и исследователей политики. - **Концептуальная ценность:** Высокая. Демонстрирует применимость метода к сложным социальным концептам. - **Потенциал для адаптации:** Высокий. Подход может быть адаптирован для оценки различных эмоциональных и социальных аспектов в текстах.

Минимальная потребность в размеченных данных - Прямая применимость: Исключительно высокая. Значительно снижает трудозатраты на разметку данных. - **Концептуальная ценность:** Очень высокая. Меняет подход к созданию систем анализа текста. - **Потенциал для адаптации:** Очень высокий. Применим в различных сценариях с ограниченными ресурсами для разметки.

Prompt:

Применение исследования CGCoT в промптах для GPT ## Ключевая концепция Исследование представляет метод **Concept Guided Chain of Thought (CGCoT)**, который превращает сложную оценку текста в структурированный анализ через концептуальные разбивки и попарные сравнения.

Пример промпта на основе CGCoT

[=====] # Промпт для анализа эмоциональной окраски текстов

Шаг 1: Концептуальная разбивка Проанализируй следующий текст с точки зрения следующих концепций: 1. Использование эмоционально окрашенных слов 2. Наличие негативных стереотипов 3. Степень выраженности агрессии 4. Использование сарказма/иронии 5. Наличие призывов к действию

Текст для анализа: "[ВСТАВИТЬ ТЕКСТ]"

Для каждой концепции: - Определи ее наличие (да/нет) - Оцени интенсивность (низкая/средняя/высокая) - Приведи конкретные примеры из текста

Шаг 2: Сравнительная оценка Теперь сравни этот анализ с предыдущим текстом, который мы анализировали. Какой из текстов содержит более выраженную негативную окраску? Объясни свое решение, опираясь на концептуальную разбивку, а не на общее впечатление. [=====]

Как работает CGCoT в этом промпте

Структурированная декомпозиция — вместо прямой оценки текста мы разбиваем анализ на конкретные концепции **Цепочка рассуждений** — модель вынуждена последовательно анализировать каждый аспект **Попарное сравнение** — сравнение по концепциям, а не по целым текстам, делает оценку более точной **Объяснимость** — получаем не только оценку, но и обоснование, опирающееся на конкретные элементы текста ## Преимущества такого подхода

- Точность — снижает влияние предвзятости модели, фокусируясь на конкретных аспектах
- Прозрачность — обоснования решений понятны и проверяемы
- Гибкость — можно адаптировать концептуальную разбивку под конкретную задачу
- Минимальная потребность в обучении — не требует размеченных данных

Этот подход особенно полезен для сложных субъективных оценок, где простой промпт может давать непоследовательные результаты.