Улучшение вывода LLM как судьи с помощью распределения судебных решений

Дата: 2025-03-04 00:00:00

Ссылка на исследование: https://arxiv.org/pdf/2503.03064

Рейтинг: 60
Адаптивность: 65

Ключевые выводы:

Исследование направлено на улучшение методов извлечения суждений из языковых моделей (LLM) при их использовании в качестве судей для оценки текстов. Основной вывод: использование среднего значения (mean) распределения вероятностей токенов суждения стабильно превосходит использование наиболее вероятного токена (mode/greedy decoding) во всех сценариях оценки.

Объяснение метода:

Исследование предлагает ценные концепции для улучшения оценок LLM: использование среднего вместо моды и отказ от CoT-рассуждений при оценке. Хотя полная реализация требует доступа к распределениям вероятностей токенов, ключевые принципы могут быть адаптированы через множественные запросы и изменение формулировок. Особенно полезны выводы о влиянии CoT и оптимальных настройках для разных моделей.

Ключевые аспекты исследования: 1. **Сравнение методов извлечения суждений из распределений вероятностей LLM**: Исследование показывает, что использование среднего значения (mean) распределения вероятностей токенов в выходных данных LLM стабильно превосходит использование моды (mode, т.е. жадное декодирование) во всех контекстах оценки (поточечной, попарной и списковой).

Влияние Chain-of-Thought (CoT) на распределение суждений: Исследование обнаружило, что CoT-рассуждения часто сужают распределение вероятностей суждений LLM, что может ухудшать эффективность работы LLM в роли судьи, особенно при использовании среднего значения.

Новые методы использования распределения вероятностей: Авторы предлагают и оценивают новые методы извлечения предпочтений из распределений вероятностей, включая методы с учетом неприятия риска (risk aversion), которые часто улучшают производительность.

Сравнение дискретных и непрерывных методов: Непрерывные методы

(работающие с распределениями вероятностей) превосходят дискретные методы (работающие с конкретными оценками), поскольку последние часто предсказывают ничьи и не улавливают небольшие различия в предпочтениях.

Оптимизация настроек для различных моделей: Разные модели LLM демонстрируют различные оптимальные настройки - более крупные модели лучше работают с попарным ранжированием без CoT, а меньшие модели часто показывают лучшие результаты с поточечной оценкой без CoT.

Дополнение:

Для работы методов этого исследования в их полной форме действительно требуется доступ к распределениям вероятностей токенов или API, позволяющее получить эти распределения. Однако многие концепции и подходы можно адаптировать для использования в стандартном чате.

Концепции, которые можно применить в стандартном чате:

Использование среднего вместо моды: Можно запросить модель несколько раз оценить один и тот же текст и затем усреднить результаты Это имитирует идею получения распределения вероятностей через множественные запросы

Отказ от CoT при оценочных задачах:

Можно напрямую просить модель дать оценку без объяснения причин Исследование показывает, что это часто дает лучшие результаты, особенно для небольших моделей

Предпочтение непрерывных шкал оценки:

Можно запрашивать оценки по более детальной шкале (например, от 1 до 100 вместо 1-5) Это позволяет модели выражать небольшие различия в качестве

Оптимальные форматы запросов:

Для крупных моделей: использовать попарное сравнение без CoT Для небольших моделей: использовать поточечную оценку без CoT Для сравнения нескольких вариантов: использовать прямое списочное ранжирование

Ожидаемые результаты от применения этих концепций:

Более точные и стабильные оценки качества текста Лучшее выявление небольших различий между вариантами Снижение влияния позиционного смещения при сравнении нескольких вариантов Более эффективное использование возможностей модели соответствующего размера Важно отметить, что хотя полная реализация методов исследования требует технических возможностей, основные принципы могут значительно улучшить качество оценок даже в стандартном чате без дополнительных инструментов.

Анализ практической применимости: **1. Использование среднего вместо моды в распределениях суждений** - **Прямая применимость**: Очень высокая. Пользователи могут запрашивать LLM оценить несколько вариантов и использовать информацию о вероятностях, а не только конкретную оценку, что повысит точность сравнений. - **Концептуальная ценность**: Высокая. Понимание того, что распределение вероятностей содержит более богатую информацию, чем один выбранный токен, меняет подход к интерпретации ответов LLM. - **Потенциал для адаптации**: Высокий. Этот подход можно применять при любой оценке качества вариантов текста LLM.

2. Отказ от CoT-рассуждений в некоторых контекстах оценки - **Прямая применимость**: Высокая. Пользователи могут получить более точные оценки, не запрашивая объяснений перед оценкой. - **Концептуальная ценность**: Высокая. Понимание того, что рассуждения могут сужать распределение вероятностей и снижать качество оценки, важно для правильного использования LLM. - **Потенциал для адаптации**: Высокий. Это знание применимо к любому контексту, где требуется оценка или сравнение.

3. Использование непрерывных методов оценки - **Прямая применимость**: Средняя. Реализация требует доступа к вероятностям токенов, что не всегда доступно через стандартные API. - **Концептуальная ценность**: Высокая. Понимание преимуществ непрерывных методов над дискретными помогает более точно формулировать запросы. - **Потенциал для адаптации**: Средний. Требуется дополнительная обработка данных, но принципы могут быть адаптированы к более простым методам.

4. Учет неприятия риска в оценках - **Прямая применимость**: Низкая для обычных пользователей, так как требует доступа к распределению вероятностей и дополнительной обработки. - **Концептуальная ценность**: Средняя. Понимание того, что учет риска может улучшать оценки, полезно для специалистов. - **Потенциал для адаптации**: Средний. Концепцию можно упростить до более доступных методов.

5. Оптимизация настроек для разных моделей - **Прямая применимость**: Средняя. Рекомендации по настройкам могут быть применены напрямую. - **Концептуальная ценность**: Высокая. Понимание различий между моделями помогает выбирать подходящую стратегию запросов. - **Потенциал для адаптации**: Высокий. Рекомендации можно адаптировать под конкретные модели и задачи.

Сводная оценка полезности: Предварительная оценка: 65 из 100.

Исследование имеет высокую концептуальную ценность, предлагая глубокое понимание того, как лучше использовать LLM для оценки текстов. Основные выводы о преимуществе среднего значения над модой и о потенциальном вреде CoT-рассуждений в контексте оценки могут быть непосредственно применены пользователями для улучшения качества оценок.

Аргументы за более высокую оценку: 1. Некоторые методы (использование среднего значения вместо моды) могут быть реализованы даже без прямого доступа к вероятностям токенов, путем запроса нескольких оценок и их усреднения. 2. Выводы о влиянии CoT могут быть применены немедленно в любом контексте оценки.

Аргументы за более низкую оценку: 1. Многие методы требуют доступа к распределению вероятностей токенов, что ограничивает их применимость через стандартные API. 2. Реализация более сложных методов (учет неприятия риска, квантили) требует технических знаний, недоступных большинству пользователей.

После рассмотрения аргументов, корректирую оценку до 55 из 100. Хотя концептуальные выводы исследования ценны, практическая реализация многих методов требует технических возможностей, выходящих за рамки стандартного взаимодействия с LLM, что ограничивает их непосредственную пользу для широкой аудитории.

Уверенность в оценке: Средняя. Существует неопределенность относительно того, насколько легко пользователи смогут адаптировать выводы исследования без доступа к распределениям вероятностей токенов. Требуется более глубокий анализ.

Пересматривая исследование и его применимость, я обнаруживаю, что даже без доступа к распределениям вероятностей, ключевые выводы могут быть адаптированы: 1. Вместо использования моды можно запрашивать несколько оценок и усреднять их. 2. Рекомендация не использовать CoT для оценки может быть применена напрямую. 3. Понимание преимуществ непрерывных оценок над дискретными может влиять на формулировку запросов.

С учетом этих факторов, повышаю оценку до 60 из 100 и уверенность до высокой.

Оценка адаптивности: Адаптивность: 65 из 100

Факторы, влияющие на оценку адаптивности:

1) **Адаптация концепций**: Основные концепции исследования (преимущества среднего над модой, влияние CoT на распределение) могут быть адаптированы для использования в стандартном чате путем запроса нескольких оценок вместо одной и избегания запроса обоснований перед оценкой.

2) **Извлечение полезных идей**: Пользователи могут извлечь важные идеи о том, как формулировать запросы для получения более точных оценок (например, избегая CoT для оценочных задач, используя прямое списочное ранжирование для сравнения нескольких вариантов).

3) **Потенциал для будущего внедрения**: Высокий. По мере того как API моделей будут предоставлять больше доступа к распределениям вероятностей, методы из исследования станут более применимыми.

4) **Абстрагирование до общих принципов**: Принципы о значимости распределений вероятностей и о том, что рассуждения могут иногда ухудшать результаты, могут быть применены в широком спектре взаимодействий с LLM.

Prompt:

Применение исследования о LLM-судьях в промптах
Ключевые знания из исследования

Исследование показывает, что при использовании языковых моделей для оценки текстов:

Среднее значение (mean) распределения вероятностей токенов суждения работает лучше, чем выбор наиболее вероятного токена (greedy decoding) **Цепочка рассуждений (CoT)** часто ухудшает качество оценки **Непрерывные методы** извлечения суждений превосходят дискретные **Попарное ранжирование** без CoT эффективно для крупных моделей **Поточечная оценка** без CoT лучше для меньших моделей
Пример промпта с применением этих знаний

[======]

Задание оценки текстовых ответов Ты выступаешь в роли судьи, оценивающего качество текстовых ответов студентов на вопрос: "Опишите влияние глобального потепления на биоразнообразие".

Инструкции:

Оцени следующие два ответа по шкале от 1 до 10. Для каждого ответа: Сначала рассмотри его достоинства и недостатки Предоставь числовую оценку Вместо выбора одного конкретного балла, укажи диапазон вероятных оценок (например, "7.2-7.8") Затем выполни попарное сравнение ответов, указав, насколько один ответ лучше другого (например, "Ответ A превосходит Ответ B на 1.5-2 балла") Не используй пошаговые рассуждения - оценивай напрямую В своей финальной оценке учитывай все возможные интерпретации качества ответа, а не только наиболее очевидную
Ответы для оценки:

Ответ A: [текст первого ответа]

Ответ B: [текст второго ответа] [======]

Объяснение эффективности

Этот промпт применяет ключевые находки исследования:

Запрашивает диапазон оценок вместо единственного значения, что соответствует идее о превосходстве среднего значения распределения над единственным токеном **Избегает цепочки рассуждений (CoT)**, так как исследование показало, что это может ухудшить результаты **Использует попарное сравнение** для более точной относительной оценки **Просит учитывать все возможные интерпретации**, что помогает модели выдавать более сбалансированные суждения, учитывающие всё распределение возможных оценок **Работает с непрерывной шкалой оценок**, что согласуется с выводом о превосходстве непрерывных методов над дискретными Такой подход позволяет получить более надежные и точные оценки от языковой модели, действующей в роли судьи.