

Обратная связь к презентации 6.10.2024

Vladislav Tumko

1. Как определять полезность публикаций?
2. Как извлекать публикации с платных ресурсов?
3. Как будете решать проблему галлюцинаций?

Ответ:

Владислав, добрый вечер! Постараюсь ответить по пунктам:

- 1) Объект нашего исследования это патенты. Объективной хорошей метрики (например цитируемость) нет, т.к. они сильно зависят от года публикации и домена.
2. Для патентную информацию мы уже извлекли с помощью google patents (написали простой парсер) Если говорить про публикации, то законно за разумные деньги ответ простой - никак, если нет доступа к университетской подписке.

3. Для решения проблемы галлюцинаций планируем использовать два подхода:

--- "промт ижиниринг"

--- создание мультиагентной системы, где часть агентов будет отвечать за качество данных.

Юлия Ильяшенко

Анатолий, можно еще раз емко - какова методология вашего исследования?

И к следующему выступлению - поработайте над структурой презентации - лучше выносить ключевые блоки (гипотеза, методология и др) отдельно в соответствии с шаблоном.

Ответ:

Юлия, постараюсь ответить на Ваш вопрос и сформулировать методологию: 1) Провести сегментацию и фильтрацию текста для собранной патентной информации. 2) Создать и настроить агент на основе одной open-source LLM, который будет обладать следующими инструментами: --- классификация эмбедингов --- выделение целевой информации о экспериментальных процедурах --- Запрос к LLM --- Запись информации в БД 3) Оценка качества модели 4) Создание мультиагентных систем с различной архитектурой для решения аналогичной задачи. 5) Сравнение качества двух подходов

Светлана Зайцева

Анатолий, у вас некорректно сформирована новизна проекта: новизна всегда в чем-то: в подходе, в методологии и так далее, прошу вас вернуться к теоретической части формулирования новизны в научно исследовательской работе.

Что нужно сделать: открыть «методология научного исследования» Новикова, Новиков. И открыть страницу 191. Там написаны теоретические элементы, из которых конструируется новизна обычно. Добавляете к этим элементам - алгоритм, модель, метод... - глагол в виде «разработал, развил итд» и в чем (что привнес нового в это). Главное - откорректируйте гипотезу, иначе все исследование потом завалится и на защите за некорректную гипотезу и нечеткие результаты эксперимента сразу сильно снижают результат. Также прошу вас потренироваться в сжатых ответах на вопросах комиссии, так как на защите у вас будет мало времени дать четкий, структурированный ответ: выделяйте главное в вопросе и отвечайте на него.