

В статье описана модификация модели GPT-3 для решения задачи ответа на открытые вопросы. Авторы использовали поисковую систему, с которой модель способна взаимодействовать. Для обучения сети был собран специальный набор данных.

Плюсы:

- Подробно описана методология сбора данных: описано множество деталей, которые могут позволить собрать похожий набор данных.
- Авторы сравнивают большое количество разных архитектур и методов обучения и приводят результаты того, как это влияет на качество.
- Качество ответов модели сравнили на сильно отличающихся наборах данных: ELI5 и TruthfulQA. При этом в статье показано, что для TruthfulQA с увеличением размера модели растёт не только общая доля верных ответов, но и доля информативных ответов.

Минусы:

- Не изучено влияние способа данных на итоговое качество модели. Нет уверенности в том, что при повторном сборе данных с использованием других исполнителей получится воспроизвести качество модели.
- Зависимость от конкретного поискового движка. Что будет, если рассмотреть какой-то другой поисковик?

Качество текста:

Статья объёмная, поскольку само обучение модели – многоэтапный процесс, при этом каждый этап достаточно сильно отличается от остальных. По этой причине иногда текст воспринимается сложно из-за резкой смены объекта исследования. При этом авторы постарались упростить текст настолько, насколько это возможно: все детали, которые не важны для общего понимания предлагаемого решения вынесены в объёмное приложение.

Воспроизводимость:

На первый взгляд в статье написаны все необходимые детали для повторения результата, однако из-за того, что процесс обучения очень сложный, возможно могут быть найдены неуточнённые детали. При этом сама сложность сбора данных тоже затрудняет воспроизведение результатов.

Комментарий: на мой взгляд большинство проблем, связанных с воспроизводимостью, решается публикацией собранного набора данных.

Оценка: 8

Уверенность: 3