Университет ИТМО, факультет программной инженерии и компьютерной техники

Двухнедельная отчётная работа по «Информатике»: аннотация к статье

Дата прошедшей лекции: 25.10.23 Номер прошедшей лекции: 4 Дата сдачи: 08.11.2023

Выполнил(а) Васильев А. Е. , № группы Р3119 , оценка

Фамилия И.О. студента не заполнять

|  |  |  |
| --- | --- | --- |
| **Название статьи/главы книги/видеолекции** Почему важна разметка данных: в основе ChatGPT лежит труд людей | | |
| **ФИО автора статьи (или e-mail)**  [kucev](https://habr.com/ru/users/kucev/) | **Дата публикации**  **(не старше 2020 года)**  "23" октября 2023 г. | **Размер статьи**  **(от 400 слов)**  \_\_\_709\_\_\_ |
| **Прямая полная ссылка на источник или сокращённая ссылка (bit.ly, tr.im и т.п.)**  https://habr.com/ru/articles/766402/ | | |
| **Теги, ключевые слова или словосочетания**  [машинное обучение](https://habr.com/ru/search/?target_type=posts&order=relevance&q=[%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5 %D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5]), [ChatGPT](https://habr.com/ru/search/?target_type=posts&order=relevance&q=[ChatGPT]), [разметка данных](https://habr.com/ru/search/?target_type=posts&order=relevance&q=[%D1%80%D0%B0%D0%B7%D0%BC%D0%B5%D1%82%D0%BA%D0%B0 %D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85]), [ИИ](https://habr.com/ru/search/?target_type=posts&order=relevance&q=[%D0%98%D0%98]) | | |
| **Перечень фактов, упомянутых в статье (минимум три пункта)**   1. Разметка данных является критически важной частью обучения языковых моделей, включая ChatGPT. 2. Для обучения ChatGPT использовался огромный датасет из более чем восьми миллионов веб-страниц, который требовал разметки включая различные типы аннотаций, такие как распознавание именованных сущностей, разметка частей речи, анализ эмоционального настроя и другие. 3. Разметка данных обеспечивает модели глубокое понимание взаимосвязей между словами и понятиями в тексте, что улучшает качество генерируемых ответов и расширяет применимость модели. | | |
| **Позитивные следствия и/или достоинства описанной в статье технологии (минимум три пункта)**   1. Разметка данных играет важную роль в обучении моделей ИИ, улучшая их способность понимать контекст и создавать более информативные ответы. 2. Разметка данных помогает моделям различать именованные сущности, части речи и эмоциональные нюансы, что делает их более полезными в различных областях. 3. Развитие ИИ идёт более быстро благодаря человеческому труду. | | |
| **Негативные следствия и/или недостатки описанной в статье технологии (минимум три пункта)**   1. Из-за ошибок в разметке могут возникнуть проблемы в разработке модели 2. Невозможно полностью автоматизировать разработку ИИ, поскольку люди всегда должны контролировать результаты работы модели и улучшать её. 3. Недооценка важности эффективной разметки данных и участия аннотаторов может привести к снижению качества и надежности искусственного интеллекта в будущем. | | |
| **Ваши замечания, пожелания преподавателю *или* анекдот о программистах[[1]](#footnote-0)**  IMG_256  У него в наушниках идёт запись лекции по информатике! | | |

1. Наличие этой графы не влияет на оценку [↑](#footnote-ref-0)