

Яна Далевич, Светлана Жафярова, Екатерина Соловкова

## Рецензия на проект Д.Бурмистрова, Е.Дорофеевой, В.Яровой

Задача: сравнение эффективности мультязычных и адаптированных к русскому языку моделей при решении задачи анализа тональности к именованным сущностям в новостных текстах

Группа протестировала различные модели, оценивая их работу с применением zero-shot на датасете RuSentNE-2023.

Одной из сильных сторон проекта является масштабность экспериментов — было протестировано несколько моделей разного уровня сложности и мощности. Также стоит отметить наличие визуализации данных.

Гипотеза была чётко сформулирована: адаптированные к русскому языку модели должны показывать лучшие результаты за счёт специфики токенизации и расширенного словаря. Результаты **частично** подтвердили эту гипотезу: “Если не учитывать нейтральный класс, адаптированные модели в целом справляются лучше мультязычных”. Тем не менее, если нейтральный класс учитывался, то мультязычные модели показывали лучшую f-меру, а некоторые адаптированные — ухудшали показатели. Подобный контраст — очень интересный результат!

На наш взгляд, более детальный разбор случаев, где модели ошибаются, помог бы понять их слабые места и предложить направления для доработки. Также хотелось бы почитать про ограничения исследования и предположения, почему итоговые метрики оказались такими низкими. Но, в целом, проект уже представляет собой хорошее, детально проработанное исследование, с которым было интересно ознакомиться!