

ПЕРЕНОС СТИЛЯ В ПИСЬМЕННОЙ РЕЧИ С ПРИМЕНЕНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Широков М.А.

(Университет ИТМО)

Научный руководитель – к. т. н., доцент Махныткина О.В.

(Университет ИТМО)

Введение. Существующие методы текстового переноса стиля не уделяют достаточного внимания субъективной природе текстовых описаний и специфическим нюансам лексики, характерным для определенного стиля. Эта задача осложняется недостаточной доступностью учебных и оценочных наборов данных, а также отсутствием единых метрик оценки при переносе текстового стиля. Это серьезно ограничивает стилистическую выразительность основных моделей, оставляя потенциальный простор для дальнейших улучшений и инноваций в данной области. Таким образом, успешный перенос стиля текста может быть определен как процесс замены первоначального стиля на целевой с высоким уровнем грамматической корректности и сохранением смыслового содержания сообщения.

Основная часть. Модель переноса стиля MSSRNet [1] строится на основе модулей хорошо известного трансформера [2]. Она состоит из генератора стилей, кодера, декодера и классификатора. Генератор стилей использует кодировщик трансформера в качестве основы. Генератором стилей создаются вектора стиля при реализации метода смешанного обучения. Классификатор обучается заранее и используется для обучения генератора стилей и создания векторного представления. В архитектуре модели применяются дискриминаторы стилей и текста. Дискриминатор стиля предназначается для оценки стиля. Дискриминаторы имеют ту же самую архитектуру, что и классификатор, кроме функции активации на последнем слое. В контексте обучения дискриминатора стилей вектора генерируются соответствующим модулем. Для проведения экспериментов использовался датасет IMDB [3]. Этот набор данных предварительно размечен. Авторы датасета каждый отзыв разбили на несколько предложений и удалили предложения с редкими словами. В итоге был получен набор данных, который состоит из 366000 предложений в качестве обучающей выборки, 4000 – валидационной и 2000 для тестовой.

Выводы. В данной работе была выполнена реализация модели переноса стиля на основе датасета IMDB. Согласно метрикам BLEU и Perplexity данная модель показала лучший баланс между задачами переноса стиля и переноса контекста.

Список использованных источников:

1. Yazheng Yang, Zhou Zhao, Qi Liu MSSRNet: Manipulating Sequential Style Representation for Unsupervised Text Style Transfer // In Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining - 2023. - С. 3022–3032.
2. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin Attention Is All You Need // Advances in Neural Information Processing Systems. - 2017. - №30. - С. 5998-6008.
3. Ning Dai, Jianze Liang, Xipeng Qiu, Xuanjing Huang Style Transformer: Unpaired Text Style Transfer without Disentangled Latent Representation // In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. - Florence, Italy: Association for Computational Linguistics, 2019. - С. 5997–6007.