

Состав команды

1) Ярцев Кирилл (Б03-212)

- добавление генератора суммаризации и демонстрация результатов
- визуализация механизма attention

2) Жгельский Артем (Б03-212)

- добавление оценки модели с помощью ROUGE metric
- объединение матриц эмбедингов

3) Александр Захаров (Б03-212)

- добавление функции ошибки LabelSmoothing
- добавление предобученных эмбедингов

Данные

Путь: data/news.csv

Графики обучения

<https://wandb.ai/yunhkirill-mipt/transformer-training/reports/Pretrained-Embeddings-Transformer--VmlldzoxMzEwOTAzOA>

Предобученные эмбединги

navec_hudlit_v1_12B_500K_300d_100q

<https://github.com/natasha/navec#installation>

Путь: data/navec_hudlit_v1_12B_500K_300d_100q.tar

Генератор суммаризации

Суммаризатор: utils/summarizer.py

Программа: summarizer_demonstration.py

Результат: results/summaries.txt

ROUGE metric

Программа: compute_rouge.py

Результат: results/rouge_metrics.txt

Визуализация attention

Программа: vizualize_attention.py

Результат: result/attention_vizualization

Обученная модель

Путь: data/transformer.pth

Модель была обучена без объединения матриц эмбедингов

d_model = 256

heads_count = 8

Обученная модель с предобученными эмбедингами

Путь: data/pre_transformer.pth

Модель была обучена с объединением матриц эмбедингов

d_model = 300

heads_count = 6