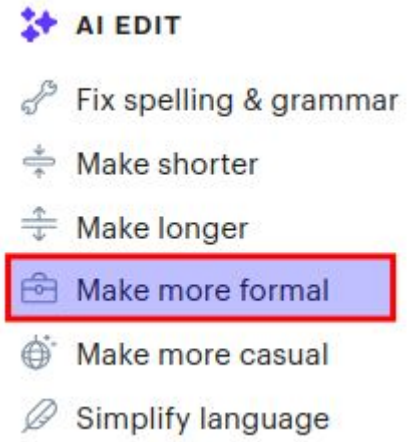
Изменение стиля текста

с использованием методов МО и ЯМ

Мотивация

- ИИ-редактирование (writing assistants) широко распространено
- Изменение стиля текста (Test Style Transfer) – одна из задач



https://www.groovehq.com/

Изменение стиля текста

- Text Style Transfer (TST):
  - Content vs. style
- Примеры:
  - Politeness transfer
  - Sentiment transfer
  - Formality transfer

	Source Style	Target Style
Impolite → Polite:	Shut up! the video is starting!	Please be quiet, the video will begin shortly.
Negative → Positive:	The food is tasteless.	The food is delicious.
Informal → Formal:	The kid is freaking out.	That child is distressed.

Table 1: TST examples regarding sentiment, polarity, and formality.

Mukherjee, S., & Dusek, O. (2024). Text Style Transfer: An Introductory Overview. ArXiv, abs/2407.14822.

Изменение уровня формальности

- Обучение модели:
  - Параллельный корпус “неформальный текст – формальный текст”
- Grammarly’s Yahoo Answers Formality Corpus (GYAFC):
  - Размеченный корпус на основе Yahoo Answers L6 corpus
  - Недоступен для свободного пользования: https://github.com/raosudha89/GYAFC-corpus
- Reddit Comments Dataset for Text Style Transfer Tasks:
  - https://zenodo.org/records/8051180
  - *“The dataset contains Reddit comments translated into a formal language. For the translation of Reddit comments into a formal language text-davinci-003 was used.*
  - *The quality of formal translations was assessed with BERTScore and chrF++:*
    - *BERTScore: F1-Score: 0.89, Precision: 0.90, Recall: 0.88*
    - *chrF++: 37.16*
    - *The average perplexity of the generated formal texts was calculated using GPT-2 and is 123.77”*

Метрики

- Style Transfer Strength (formality):
  - Предобученные модели
- Content Preservation:
  - Косинусное сходство (эмбеддинги)
  - BERT-score*
- Fluency:
  - Perplexity
  - Модели, обученные на CoLA (The Corpus of Linguistic Acceptability)
- BART-score**

*Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., & Artzi, Y. (2019). BERTScore: Evaluating Text Generation with BERT. ArXiv, abs/1904.09675.
** Yuan, W., Neubig, G., & Liu, P. (2021). BARTScore: Evaluating Generated Text as Text Generation. ArXiv, abs/2106.11520.

Reddit Comments Dataset

- *Subreddit, Original Comment, Formal Comment*
- Train: 2372 → 2342
- Eval: 602 → 581

	source	target
0	Isnt that stating the obvs ?	That seems to be an obvious statement, doesn't...
2	Getting down oted for this comment just proves it	This comment has been noted and appreciated.
3	Cant pass laws that help average citizens!	It is not possible to pass laws that benefit t...
4	Liz Cheny and Adam Kiplinger's would disagree	Liz Cheney and Adam Kiplinger have differing o...
5	What does ctrl F and Peter Thompson mean? Just...	What is the relationship between the use of th...

T5 paraphraser model

- *humarin/chatgpt_paraphraser_on_T5_base*
- *“This model was trained on our ChatGPT paraphrase dataset. This dataset is based on the Quora paraphrase question, texts from the SQUAD 2.0 and the CNN news dataset. We used "transfer learning" to get our model to generate paraphrases as well as ChatGPT.”*

Hugging Face

```
input_ids = tokenizer(
    f'paraphrase: {question}',
    return_tensors="pt", padding="longest",
    max_length=max_length,
    truncation=True,
).input_ids.to(device)
```

Input:

```
text = 'What are the best places to see in New York?'
paraphrase(text)
```

Output:

```
['What are some must-see places in New York?',
 'Can you suggest some must-see spots in New York?',
```

Дообучение T5 модели

- 10 000 шагов (4+ эпохи)
- AdamW: lr=2e-5
- *prefix = "formalize: "*

source	target	pt	ft
Isnt that stating the obvs ?	That seems to be an obvious statement, doesn't...	Is that not exaggerated?	That seems to be stating the obvious.
Getting down oted for this comment just proves it	This comment has been noted and appreciated.	The fact that you are being cheated for this c...	Getting a negative comment justifies the state...
Cant pass laws that help average citizens!	It is not possible to pass laws that benefit t...	Is unable to pass laws that benefit the common...	Can't pass laws that are beneficial to the gen...

*pt = pre-trained
ft = fine-tuned

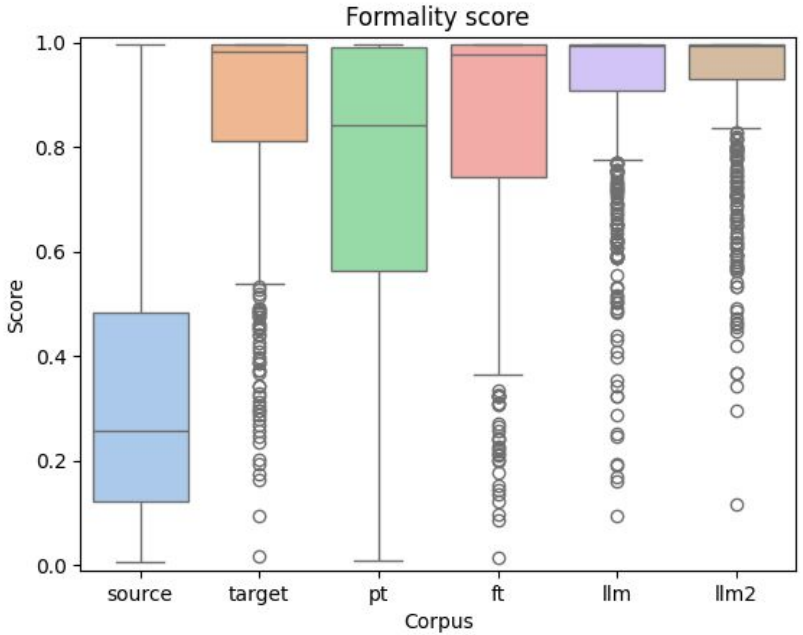
LLM (zero-shot)

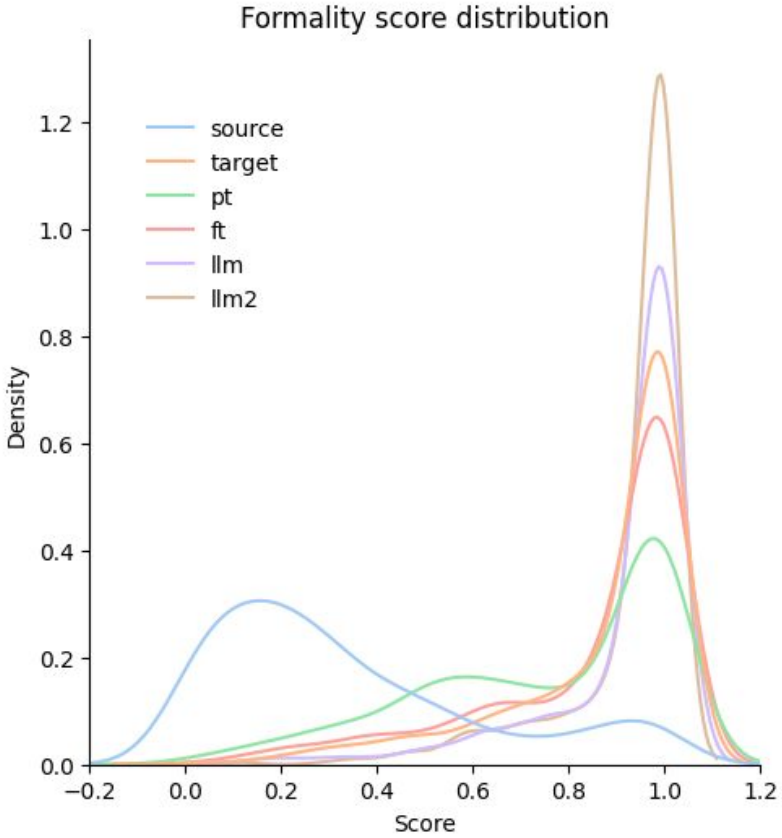
1. GigaChat-2: ~10%
   WARNING:langchain_gigachat.chat_models.gigachat:Giga generation stopped with reason: blacklist

2. DeepSeek-V2:
   deepseek-ai/DeepSeek-V2
   `temperature=1.3`
   `max_tokens=128`
   System prompt: *"You are a Text Style Transfer model. Rewrite the user text in a more formal style. I need just one version of the text without any comments from you."*

Formality score

- *s-nlp/roberta-base-formality-ranker**

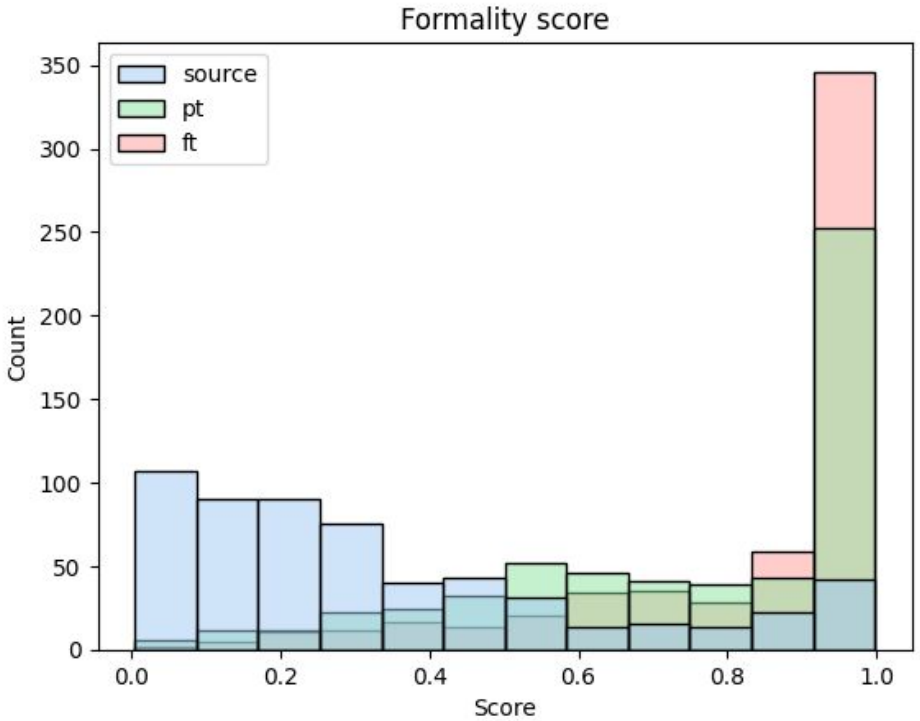




*Babakov, N., Dale, D., Gusev, I., Krotova, I., & Panchenko, A. (2023). Don't Lose the Message While Paraphrasing: A Study on Content Preserving Style Transfer. ArXiv, abs/2308.09055.

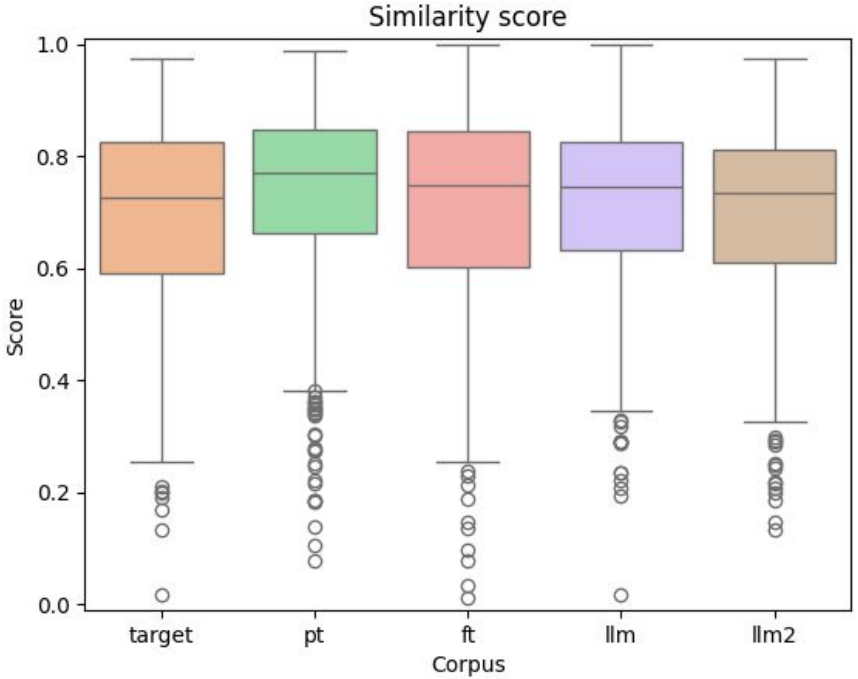
Formality score #2

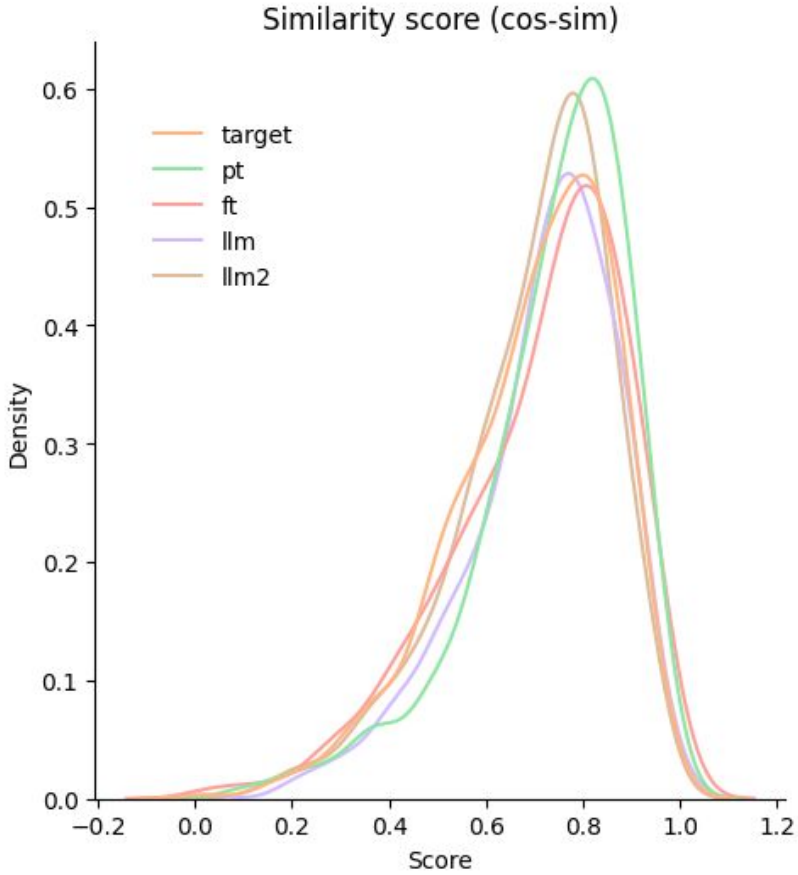
- Welch’s t-test (one-sided): pre-trained (pt) vs. fine-tuned (ft)
  - T-statistic: -5.67
  - P-value: 8.61e-9
- ANOVA:
  - statistic= 606.96
  - pvalue= 3.50e-291



Semantic similarity

- SentenceTransformer: *all-mpnet-base-v2**

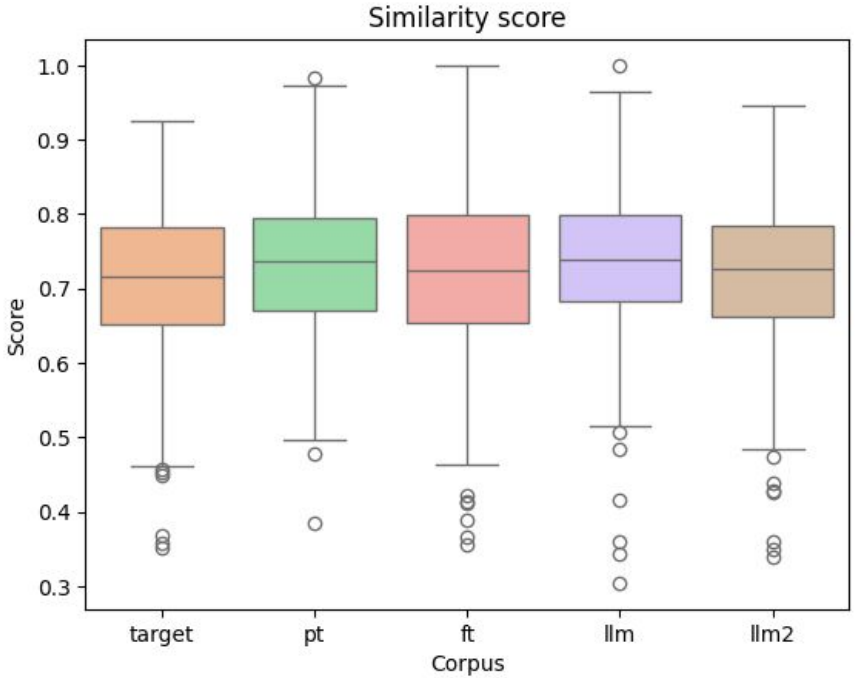


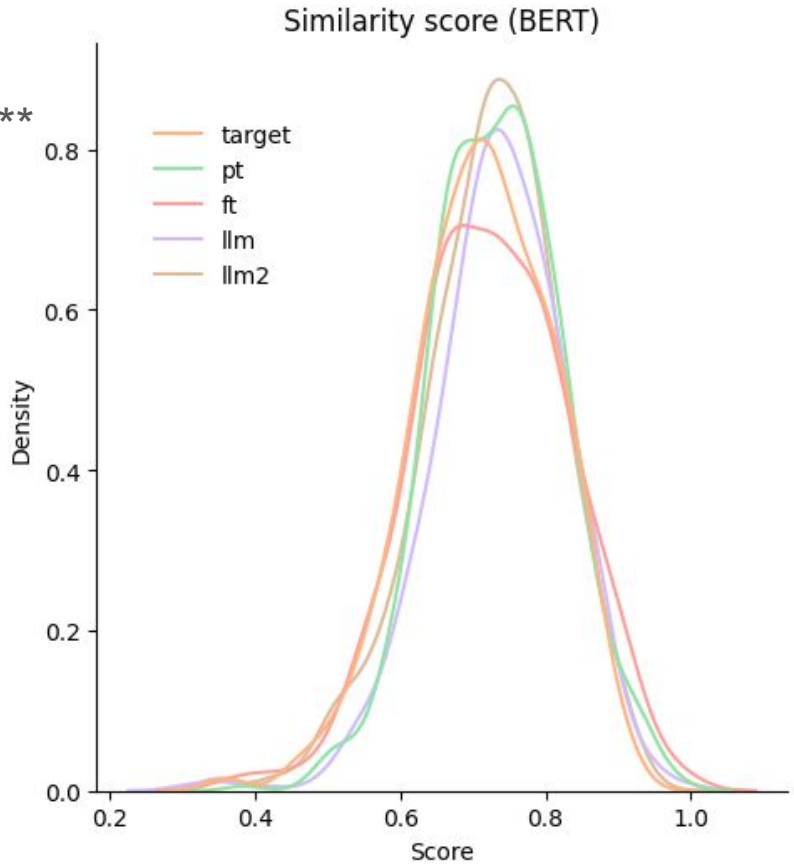


*Лучшая модель с https://www.sbert.net/docs/sentence_transformer/pretrained_models.html

Semantic similarity #2

- BERT-score*: *microsoft/deberta-xlarge-mnli***

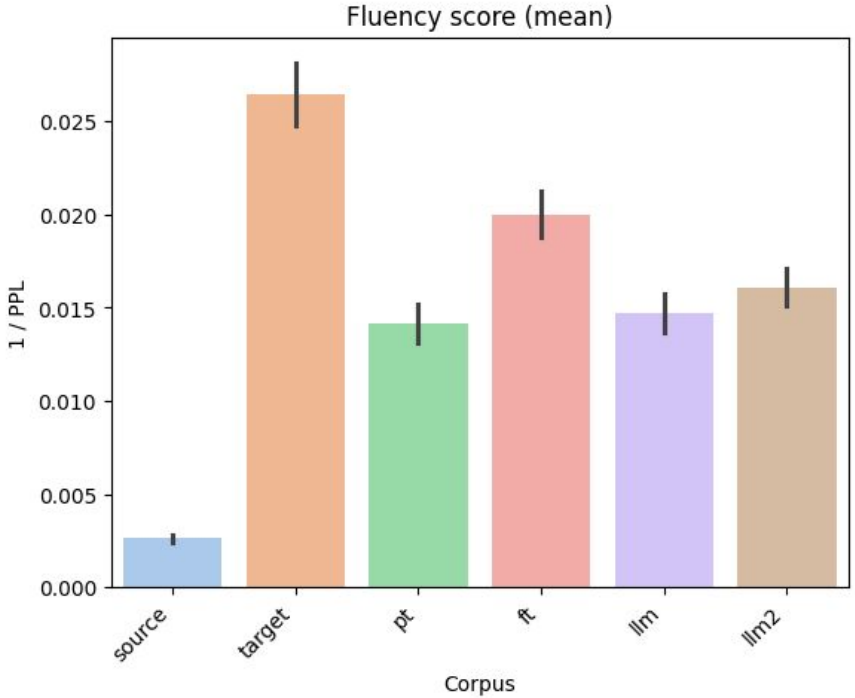


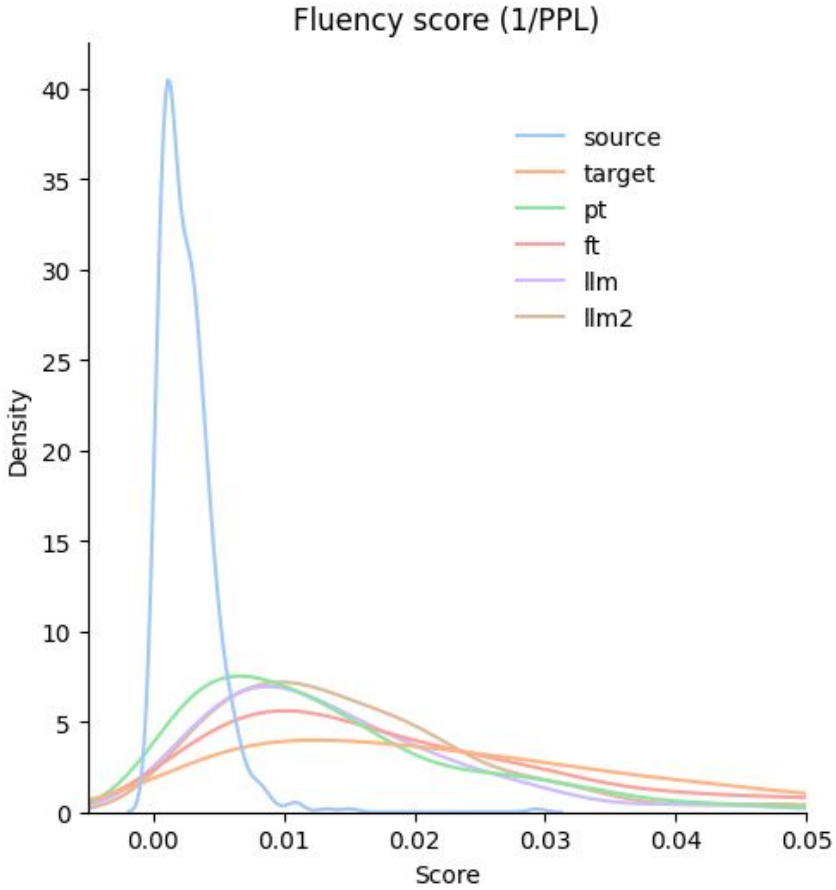


*https://github.com/Tiiiger/bert_score **Лучшая модель по КК Пирсона с оценкой человеком

Fluency score

- Perplexity*: *gpt2*

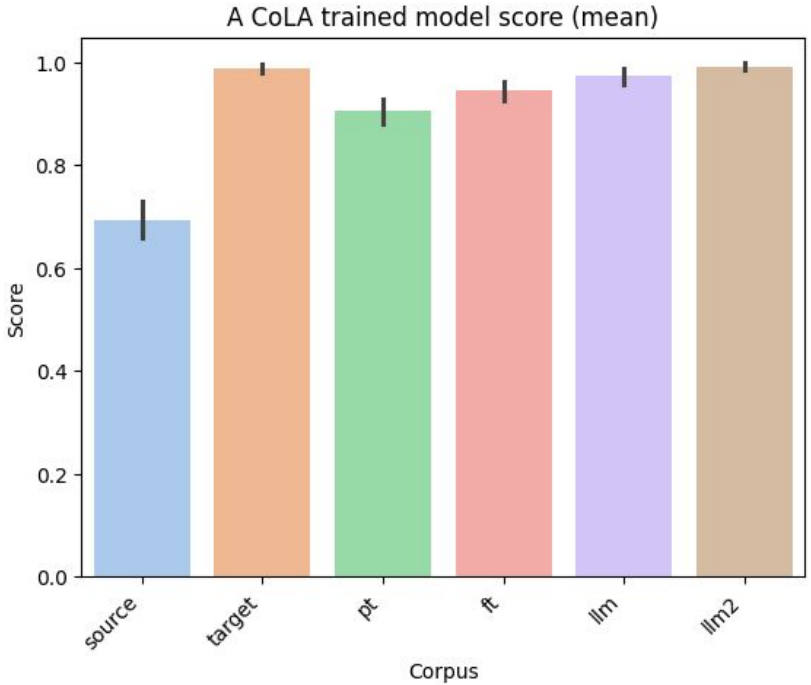


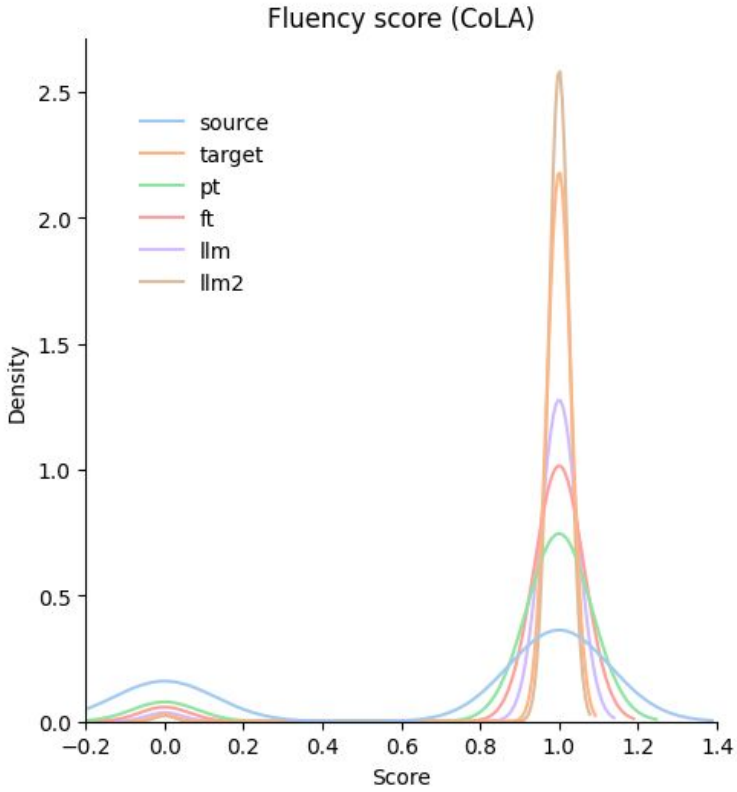


*https://www.comet.com/site/blog/perplexity-for-llm-evaluation/

Fluency score #2

- *cointegrated/roberta-large-cola-krishna2020**





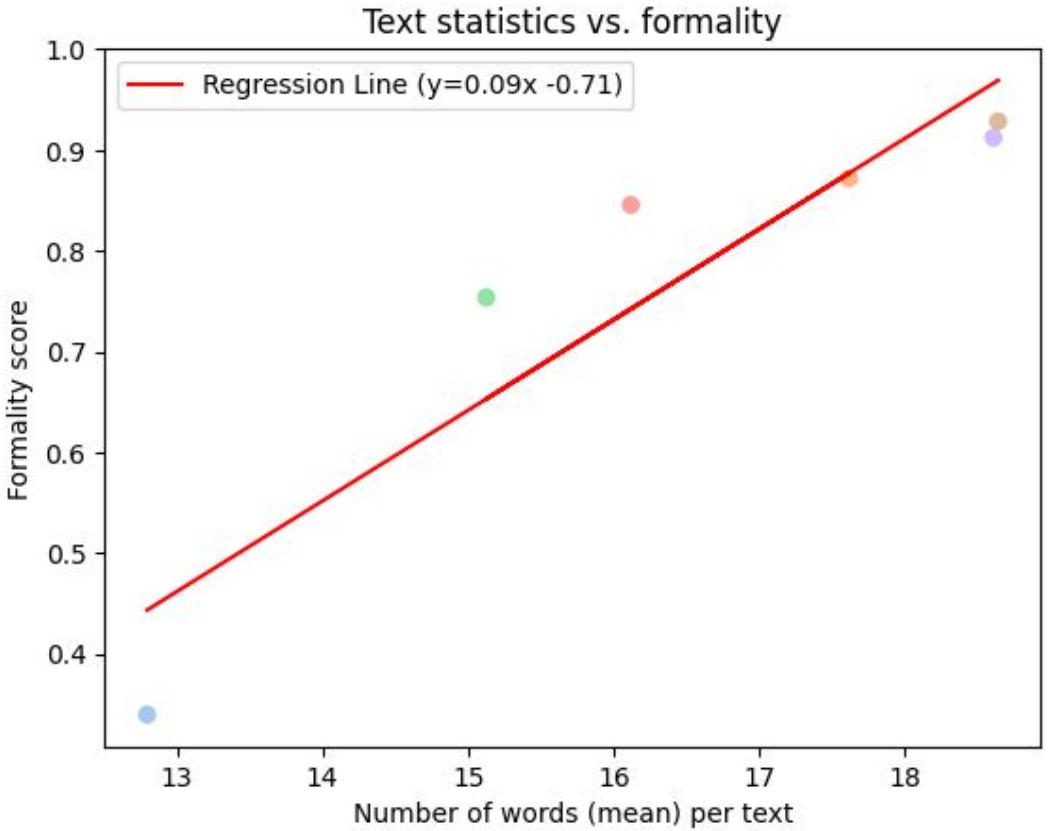
*Krishna, K., Wieting, J., & Iyyer, M. (2020). Reformulating Unsupervised Style Transfer as Paraphrase Generation. ArXiv, abs/2010.05700

Выводы

- Удалось успешно дообучить T5 модель
- Показатели модели сравнимы с эталонной моделью и LLM

Bonus: text statistics vs. formality

- Длина текста:
  - LLM-тексты длинней
- Количество слов:
  - LLM: больше слов
  - PearsonRResult:
    - statistic=0.92
    - pvalue=0.009
- Средняя длина слова
  - LLM: слова длинней



Спасибо за внимание!