Zero-Shot and Few-Shot with OpenAPI and SetFit on IMDB

1. Questões

1. **Explicação de conceitos importantes do exercício feito**
2. ~~Técnicas para garantir que a implementação está correta~~
3. ~~Truques de código que funcionaram~~
4. **Problemas e soluções no desenvolvimento**
5. ~~Resultados interessantes/inesperados~~
6. **Uma dúvida "básica" que você ou os colegas possam ter**
7. ~~Um tópico "avançado" para discutirmos~~

2. Explicação de conceitos importantes do artigo

SetFit is an efficient and prompt-free framework for few-shot fine-tuning of Sentence Transformers. Based on the Customer Reviews sentiment datasets benchmark, SetFit is competitive and achieve comparable performance with only 8 labeled examples per class compared to fine-tuning RoBERTa Large with datasets of 3k labeled examples.

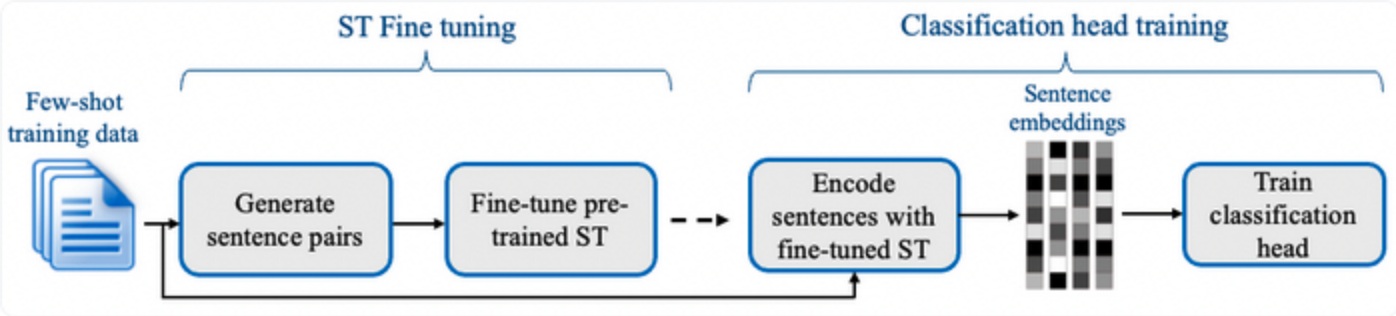
Sentence Transformers SentenceTransformers is a Python framework for state-of-the-art sentence, text and image embeddings - **Based on embeddings, no prompts are required and supports multilingual text classification**

Few-Shot Learning is the practice of training a machine learning model with a small amount of data.

2.1 SetFit

How does it work?

SetFit is designed with efficiency and simplicity in mind. SetFit first fine-tunes a Sentence Transformer model on a small number of labeled examples (typically 8 or 16 per class). This is followed by training a classifier head on the embeddings generated from the fine-tuned Sentence Transformer.



SetFit's two-stage training process

3. Problemas e soluções no desenvolvimento

- first time using hf datasets
- discovered about langchain
- tutorial setFit https://huggingface.co/blog/setfit
- how to use setFit for zero-shot (notebook do Gustavo Bartz Guedes)

4. Resultados

SetFit

zero-shot

	precision	recall	f1-score	support
0	0.64	0.60	0.62	12500
1	0.62	0.66	0.64	12500
accuracy			0.63	25000
macro avg	0.63	0.63	0.63	25000
weighted avg	0.63	0.63	0.63	25000

few-shot

	precision	recall	f1-score	support
0	0.85	0.92	0.88	12500
1	0.91	0.84	0.87	12500
accuracy			0.88	25000
macro avg	0.88	0.88	0.88	25000
weighted avg	0.88	0.88	0.88	25000

GPT-3.5 OpenAPI

zero-shot

	precision	recall	f1-score	support
0	0.64	0.60	0.62	12500
1	0.62	0.66	0.64	12500
accuracy			0.63	25000
macro avg	0.63	0.63	0.63	25000
weighted avg	0.63	0.63	0.63	25000

few-shot

	precision	recall	f1-score	support
0	0.64	0.60	0.62	12500
1	0.62	0.66	0.64	12500
accuracy			0.63	25000
macro avg	0.63	0.63	0.63	25000
weighted avg	0.63	0.63	0.63	25000

4. Uma dúvida "básica" que você ou os colegas possam ter

- zero-shot do SetFit não parece zero-shot pela definição do artigo do GPT-3 (mas é usado os embeddings do corpus como se fosse o LM...faz sentido)
- few-shot no setfit overfit pois só tem 8 amostras para cada label (é possivel fazer otimizar hyperparametros para dataset maiores)