Repository link : https://github.com/Ad-Vi/few-shot-learning-LLM-for-irrelevance-classification

Few shots learning LLM for Relevance Classification

Project done in the scope of the course Web and Text Analytics given at the University of Liège in 2023. The approach taken is to fine-tune a small existing model using the dataset provided by the Teaching Staff.

Dataset

The dataset is composed of 25112 sentences, each classified as being relevant to the context or not. The context of relevance is thus the whole dataset context (the 25112 sentences).

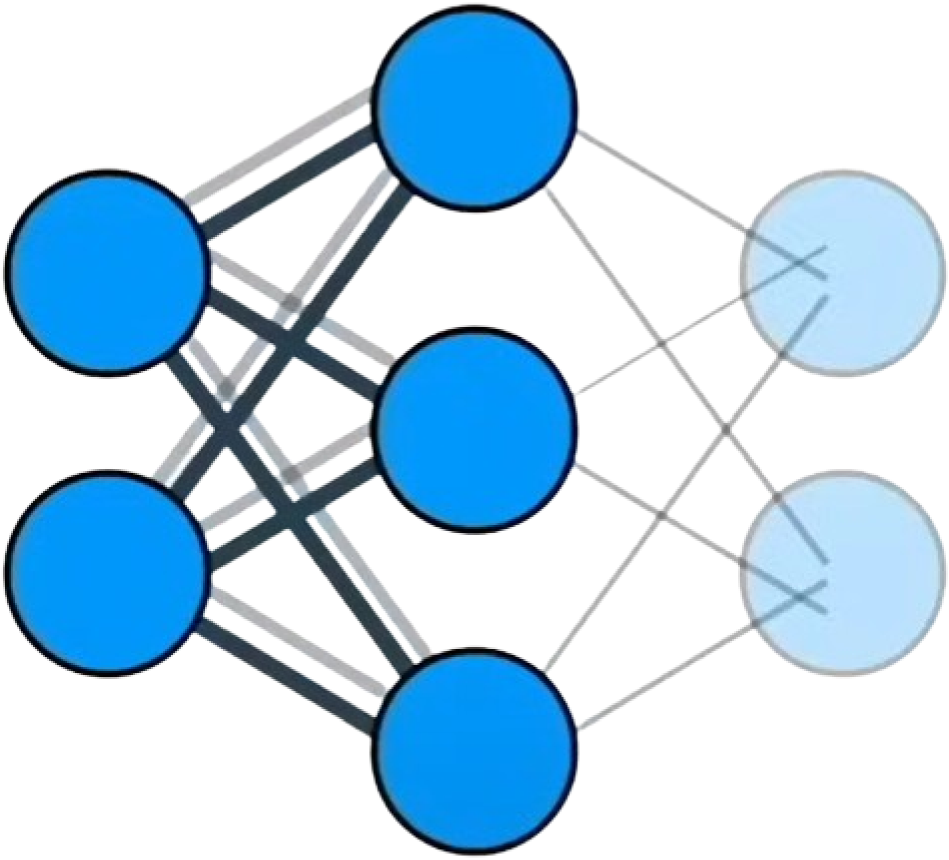
Implementation

Model

The model used is distilbert-base-uncased. It offers the advantages of being small and thus possible to train with limited resources and having good performances in regards to other bigger transformer models.

Low-Rank Adaptation (LoRA) - APPROACH 1

The principle of LoRA is to freeze the initial model parameters and to fine-tune additional parameters to adapt the model to a new task. This yields an adaptation of the initial model with a relatively low number of parameters to train.



This first approach can be found in the LoRA notebook.

Model fine-tuning - APPROACH 2

The second approach is to fine-tune the whole model using pytorch. This approach can be found in the pytorch notebook.

Results

At the end of the fine tuning :

Method	Training Loss	Validation Loss	Accuracy	Precision	Recall	F1
LoRA	0.122800	0.493573	0.8918	0.9216	0.937	0.8566
Total Fine-tune			0.8918	0.88017	0.93848	0.90839

Contributors:

The team which contributes to this work is composed of :

- Cédric HONS
- Dylan PROVOOST
- Adrien VINDERS