Project 2

Please submit the report (in pdf format) on Blackboard system before 23:59 Dec. 26. Report can be written either in English or Chinese. Name your report as `studentID_Name.pdf`. If you use ChatGPT etc., list all the prompts you use. Note that the LLM may have hallucination problem and you are responsible for the correctness of the answers.

Q1

[8 points] Select one paper from the list. Write a report and answer the ten questions mentioned here.

- If you would like to select a different paper about NLP, e.g. ACL/EMNLP papers, send me an email about the paper to ask.
- For the last question, if you have no ideas, search and read some following-up papers which cite this one, summary two or three main contributions as the answer.

Paper list:

1. DeBERTa: Decoding-enhanced BERT with Disentangled Attention
2. Improving language models by retrieving from trillions of tokens
3. UL2: Unifying Language Learning Paradigms
4. Mass-Editing Memory in a Transformer
5. Secrets of RLHF in Large Language Models Part I: PPO

Q2

Large language model (LLM) finetuning and evaluation.

1. [4 points] Finetune TinyLLaMA-1b with clinical instruction data (see attached) with QLoRA method.

2. [4 points] Evaluate the finetuned LLM on MedMCQA dataset (see attached). As the test set has no answers, we will evaluate on the validation dataset.

3. [4 points] Build a personal website and display your thoughts and ideas about how to apply LLMs in specific domains like clinic/law/finance.

Note:

1. You can refer to github repos related to LLM finetuning, such as Firefly, Chinese LLaMA, etc.

2. You can use github pages and Hexo to build a free website, you can also request a one-year-free domain name with your edu email on namecheap. More information here.

3. Due to lack of enough computation resources, you are not required to finetune the model until convergence. You can use cpu to debug your code.

Presentation

[5 points] Every student prepares and submits a 12-minutes PPT about these two tasks.