Data Science HW2 - Prompt Engineering

[TA] 曹立武

Submission Deadline: 2024/3/26 23:59

Objectives

- In this homework, you need to try any kind of prompt engineering techniques with LLMs to improve the accuracy on the MMLU dataset we provide.

- You don't need to train the model in this homework, just use the LLMs.

- The sample data is provided for you to verify the correct format or to use it for any purpose.

Data format in MMLU dataset

- Input: 學科問題
- A/B/C/D : 四個選項
- Task : 學科
- Target: 正確答案(A or B or C or D)

- Sample file: mmlu_sample.csv (Input, A/B/C/D, Task, Target)
- Submission file: submit.csv (Input, A/B/C/D, Task) -> Predict: Target

Grading Policy

- Top 20%: 10%
- Top 50%: 10%
- Hard baseline(Provide on 3/19): 20%
- Simple baseline: 60%

Useful resources

- Prompt bench
  - https://github.com/microsoft/promptbench/tree/main
- Any other open-source LLMs or prompt engineering strategies

Rules

- Use your student ID as the team name on Kaggle.
- A maximum of 5 submissions per day is allowed on Kaggle.
- Write your own code.
- Design your own prompt strategy based on any resources you can find.
- You need to upload the code to E3 that generates your answer.
- You can use any API to call the LLMs if you want, but we won’t pay it for you. (The baselines can be achieved by open-source code)

Submissions

- Submit your results to Kaggle:
  - https://www.kaggle.com/t/8c1691bcdbc34ea3b0fbbe2cf96383b0

- Submit your zipped source code {student_id}.zip to E3. The zip file should contain
  - {student_id}/main.py
  - {student_id}/prompt.txt
    - Prompt you use: … (Design by yourself)
    - LLM you use: … (For example, llama2-7b)
    - Your prompt strategy: … (For example, Few-Shot Chain-of-Thought)
  - {student_id}/requirements.txt (If you need to download some libraries)

Homework information

- Deadline: 2023/3/26 23:59
- You can send an e-mail through E3 if you have any problems.
- [TA] 曹立武