Assignment 3

Please submit the code and report (in pdf format) on Blackboard system before 23:59 December 11. Report can be written either in English or Chinese. Name your report as `studentID_Name.pdf`

Q1

Explain and answer the following questions in your own words. You may refer to lecture notes for more details.

(1) [2 points] Briefly describe the **goals**, **gained abilities**, and **main challenges** of **Pretraining**, **Instruction Tuning**, and **Alignment Tuning**. Give representative models for each stage.

(2) [1 point] Read the blog A Closer Look at Large Language Models Emergent Abilities and explain the emergent abilities with various examples.

(3) [2 points] Evaluate **Qwen2.5-0.5B** on **GSM8K** and **MATH** benchmark using different prompting methods: **zero-shot CoT**, and **few-shot** (with shot number k=5). Report the accuracy of each setting.

You may refer to:

- https://github.com/meta-math/MetaMath/blob/main/eval_gsm8k.py

- https://github.com/meta-math/MetaMath/blob/main/eval_math.py

- https://www.promptingguide.ai

Q2

Explain and answer the following questions in your own words. You may refer to lecture notes for more details.

(1) [2 points] Briefly describe the three main filtering strategies in synthetic data pipelines: **Diversity filtering**, **Quality filtering**, and **Correctness filtering**. Explain one representative method or example for each (e.g., ROUGE-L, reward model, final answer verification).

(2) [3 points] Write code to illustrate a simple **Reward-based Filtering** process:

- Prepare 10 seed instruction–response pairs (can be manually written).
- For each instruction, use an open/proprietary LLM (e.g., free models in OpenRouter) to generate 5 candidate responses.
- Assign a heuristic reward score to each response (e.g., weighted sum of response length, keyword match, and BLEU score, etc.).
- Keep the top-1 response as the "filtered" sample.
- Print all candidates with their scores and indicate the selected one.

Example Output:

```
Prompt: Explain the difference between supervised and unsupervised learning.
Candidate 1: {xxxx}. Score = 0.85
Candidate 2: {xxxx}. Score = 0.78
Candidate 3: {xxxx}. Score = 0.74
...
Selected: Candidate 1
```