Foundations of Deep Neural Networks

Assignment 3 - Long Writer with Large Language Model

Instructor: Ying-Cong Chen

October 22, 2024

1 Goal

Due to the limitation on the number of tokens, large language models typically can only generate texts of limited length, posing a severe challenge for generating long texts based on these large models. Therefore, this experiment aims to explore the potential of utilizing large models to **generate long and high-quality texts**, to address the challenges associated with producing high-quality, lengthy texts.

2 Task Description

You are required to generate a coherent and high-quality long text by leveraging the scheduling/planning of large language models. The goal is to produce long content and ensure that the generated text maintains coherence, relevance, and quality throughout its length.

2.1 Main Steps

1. **Understanding the API:** Understanding the API usage and its capabilities.
2. **Initial Experimentation:** Conduct initial experiments to understand the token limits of the model.
3. **Model Scheduling or Planning:** Develop a strategy to schedule multiple API calls to generate long text. This could involve dividing the task into smaller segments and ensuring smooth transitions between segments.
4. **Quality Assessment:** Implement methods to ensure the quality and coherence of the generated text. This might include post-processing or using additional models to evaluate and refine the text.

2.2 Assessment

- **Text Length (↑):** No less than 10000 words.
- **Text Quality (↑):** The quality of the text will be evaluated by GPT-4. Criteria for evaluation will include coherence, relevance, fluency, and readability.
- **Number of API Calls (↓):** Decrease the number of API calls made during the whole process.
- **Technical Novelty:** Get extra scores, if your method is attractive.

3 Submission

1. **Generated Text:** Following the guidance, submit the generated long text. `StudentNumber.txt`.
2. **Code:** Following the codes, finished all codes to generate the text in `code`, including all codes for model planning, API usage, file restoring, or pre/post-processing. `Assignment.py`
3. **Report:** Write a NOT too-long report about the novelty of your approach, the challenges faced, and how you addressed them, with the statistic information of the generated text. `report.pdf`

Zip all of them to submit to Canvas: `YourName_StudentNumber.zip`. For example, `JiaqiTang_50005050.zip`

4 Instruction

4.1 API Usage

Install the API by:

```
pip install zhipuai
```

ONLY use the free API (**glm-4-flash**), available at https://open.bigmodel.cn/dev/howuse/introduction.

```
from zhipuai import ZhipuAI
client = ZhipuAI(api_key="") # Please fill in your own APIKey
response = client.chat.completions.create(
    model="glm-4-flash", # Please only use this free API
    messages=[
        {"role": "user", "content": "As a marketing expert, please create an attractive
            slogan for my product"},
        {"role": "assistant", "content": "Sure, to create an attractive slogan, please tell
            me some information about your product"},
        {"role": "user", "content": "ZhipuAI Open Platform"},
        {"role": "assistant", "content": "Ignite the future, ZhipuAI paints the infinite,
            making innovation within reach!"},
        {"role": "user", "content": "Create a more precise and attractive slogan"}
    ],
)
print(response.choices[0].message)
```

Here is an example of a returned response:

```
   {
  "created": 1703487403,
  "id": "8239375684858666781",
  "model": "glm-4-flash",
  "request_id": "8239375684858666781",
  "choices": [
     {
         "finish_reason": "stop",
         "index": 0,
         "message": {
             "content": "With AI painting the blueprint ZhipuAI, making every moment of
                 innovation possible.", # the response content

             "role": "assistant"
         }
     }
  ],
  "usage": {
     "completion_tokens": 217,
```

```
    "prompt_tokens": 31,
    "total_tokens": 248
  }
}
```


4.2 Examples

To begin your homework, we provide an example for reference: `Assignment3_Example.py`. **One** topic can be selected in these three, as follow:

- Stepping on the Rainy Street
- A Futuristic Cafe
- Application of Nanotechnology in Medicine

5 Others

If you have some problems, please contact `jtang092@connect.hkust-gz.edu.cn`.