

答辩报告

一、技术说明

1.1 技术方案概括

在本次比赛中，我们选择使用glm-4-9b-chat模型来进行推理，针对每个问题生成多个答案，然后利用微调后的Qwen2.5-7B-Instruct模型作为打分模型，对生成的答案进行评分，最后选择最优的答案作为最终的推理结果。

1.2 数据选择

在本次比赛中，我们仅使用提供的旅游SFT数据集。

1.3 数据预处理

由于是训练打分模型，因此没有对数据集进行筛选，让模型接触到好的和坏的数据可以更好地训练其打分能力。我们需要根据给定的SFT数据集构造一个打分的训练集，具体来说，提取SFT数据集中的问题、答案和rw_score，将问题与答案拼接在一起构成问答对，然后设计prompt让模型根据问题对回答的质量进行打分。prompt的设计考虑了中英文区分，具体如下：

- 中文prompt:

"阅读下面的对话，'question'是一个与旅游或地理相关的问题，'answer'是一个模型给出的回答，请对这个模型的回答质量的好坏给出一个打分，注意打分必须得十分的严格，任何没有关注到的细节和事实性错误都必须给予一个极低的分数，分数的区间为[-80,80]。"

- 英文prompt:

"Read the following conversation, 'question' is a question related to tourism or geography. 'answer' is the answer given by a model. Please give a score for the quality of the model's answer. Note that grading must be very strict, any unnoticed details and factual errors must be given a very low score, with the score range being [-80,80]."

构造训练集的代码如下：

```
1 import json
2
3 def extract_jsonl(input_file, output_file):
4     examples=[]
5     prompt_zh="阅读下面的对话，'question'是一个与旅游或地理相关的问题，'answer'是一个
模型给出的回答，请对这个模型的回答质量的好坏给出一个打分，注意打分必须得十分的严格，任何
有关关注到的细节和事实性错误都必须给予一个极低的分数，分数的区间为[-80,80]。"
```

```

6     prompt_en="Read the following conversation, 'question' is a question
related to tourism or geography. 'answer' is the answer given by a model.
Please give a score for the quality of the model's answer. Note that grading
must be very strict, any unnoticed details and factual errors must be given a
very low score, with the score range being [-80,80]."
```

```

7     with open(input_file, 'r', encoding='utf-8') as infile:
8         for i,line in enumerate(infile):
9             data = json.loads(line)
10            question=data.get("conversations")[0].get("value")
11            answer=data.get("conversations")[1].get("value")
12            score=str(data.get("rw_score"))
13            qa_pair={"question":question, "answer":answer}
14            if '\u4e00' <= question[0] <= '\u9fff':
15                prompt=prompt_zh
16            else:
17                prompt=prompt_en
18            example={
19                "instruction":prompt,
20                "input":f"{qa_pair}",
21                "output":score
22            }
23            examples.append(example)
24        with open(output_file, 'w', encoding='utf-8') as outfile:
25            json.dump(examples, outfile, ensure_ascii=False, indent=4)
26
27 input_path = 'industry_instruction_semantic_cluster_dedup_旅游_地理
_valid_train.jsonl'
28 output_path = 'pingfen.json'
29 extract_jsonl(input_path, output_path)
30
```

构造好的训练集是一个json文件，数据格式如下：

```
{
  "instruction": "阅读下面的对话, 'question'是一个与旅游或地理相关的问题, 'answer'是一个模型给出的回答, 请对这个模型的回答质量",
  "input": "{ 'question': '你认为哪个旅游景点的地理特色对环境的影响最小, 为什么?', 'answer': '答案: <我认为国家公园或者自然保护区',
  "output": "-10.8125"
},
{
  "instruction": "Read the following conversation, 'question' is a question related to tourism or geography. 'answer'",
  "input": "{ 'question': 'Based on the text, what is the significance of the integrated development of various industr",
  "output": "-9.4375"
},
{
  "instruction": "Read the following conversation, 'question' is a question related to tourism or geography. 'answer'",
  "input": "{ 'question': 'What is the significance of the Hemis festival, and what are the main events and activities",
  "output": "-3.75"
},
{
  "instruction": "阅读下面的对话, 'question'是一个与旅游或地理相关的问题, 'answer'是一个模型给出的回答, 请对这个模型的回答质量",
  "input": "{ 'question': '户外活动如何帮助学生理解地理环境与人文环境的关系?', 'answer': '<答案>: 户外活动可以让学生直接接触自",
  "output": "0.16796875"
},
{
  "instruction": "Read the following conversation, 'question' is a question related to tourism or geography. 'answer'",
  "input": "{ 'question': 'What is the approximate size of Giant Rock, and how does it compare to the size of a person",
  "output": "-8.9375"
},
},
```

1.4 训练过程

我们利用LLaMA Factory框架, 使用构造的打分数据集对Qwen2.5-7B-Instruct模型进行Lora微调。问题和答案作为用户输入"input", rw_score作为助手回答"output", prompt作为指令"instruction"。

1.4.1 训练超参数

主要的训练超参数如下:

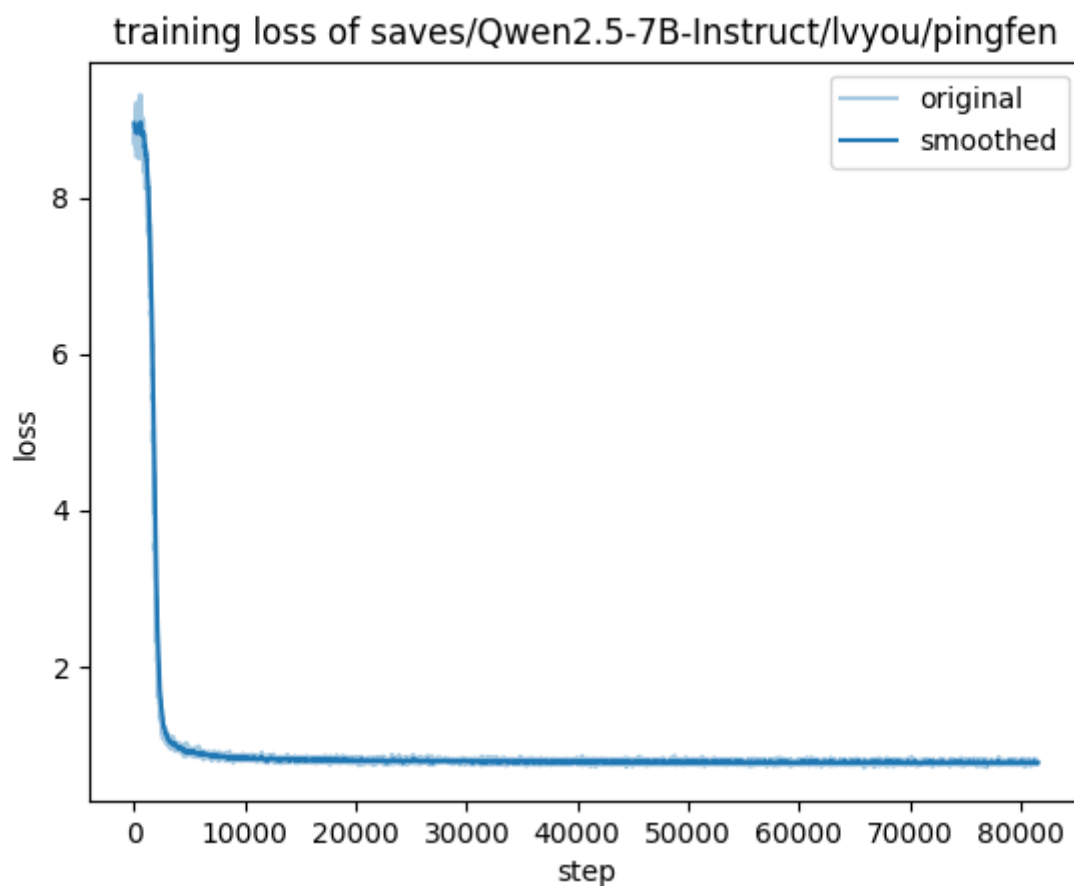
- lora_target: all
- cutoff_len: 8000
- logging_steps: 10
- save_strategy: epoch
- learning_rate: 1.0e-6
- num_train_epochs: 5
- lr_scheduler_type: cosine
- warmup_ratio: 0.1
- bf16: true
- val_size: 0.05

Lora合并过程的超参数如下:

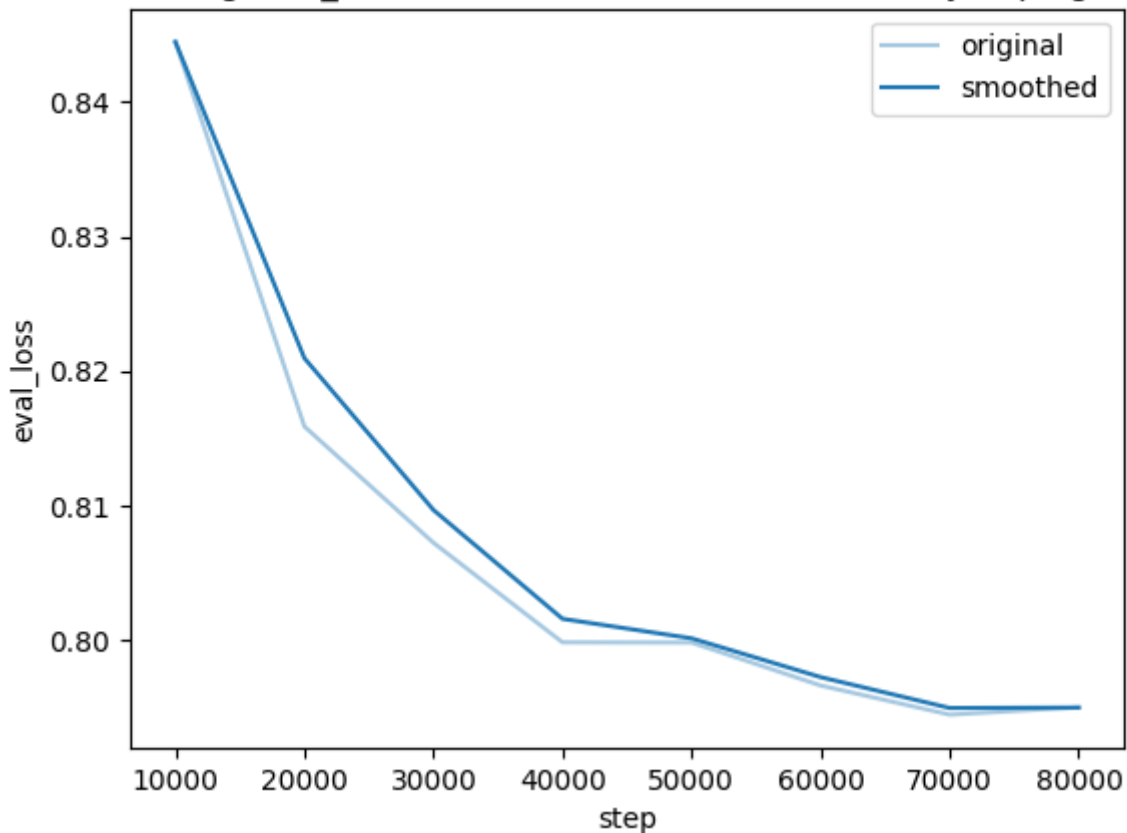
- template: qwen
- finetuning_type: lora

- export_size: 2
- export_device: cpu

1.4.2 训练曲线



training eval_loss of saves/Qwen2.5-7B-Instruct/lvyou/pingfen



1.4.3 训练日志

训练日志详见trainer_log.jsonl

1.5 推理过程

我们使用glm-4-9b-chat作为我们的推理模型。首先通过vllm，将推理模型与评分模型部署到本地，然后使用OpenAI接口的形式调用模型进行推理。为了让模型输出高质量的回答，我们对推理阶段的prompt进行了设计。首先区分了中英文，根据问题的第一个字符进行判断，然后选择正确的prompt；其次，针对主客观题，分别作了要求。我们将设计的prompt设置为为system prompt。

```

1 # 根据问题语言选择合适的提示词
2 if '\u4e00' <= question[0] <= '\u9fff':
3     messages = [
4         {"role": "system", "content": prompt_zh},
5         {"role": "user", "content": question}
6     ]
7 else:
8     messages = [
9         {"role": "system", "content": prompt_en},
10        {"role": "user", "content": question}
11    ]

```

prompt如下所示：

- 中文：

你是一位资深的旅游与地理专家，擅长提供详细、准确且富有见解的回答。请根据以下规则回答用户的问题：### 通用规则：

1. 使用简洁且专业的语言，避免冗长和重复。
2. 确保回答内容详尽、逻辑清晰，并以用户需求为核心展开。
3. 在适当情况下，加入相关背景知识，丰富内容。

针对主观题

- 回答需要体现你的专业见解，并提供清晰的理由或说明。
- 如果适用，提供多种角度或选项，帮助用户进行决策。例如，旅游推荐可以根据预算、兴趣、季节等进行分类。

针对客观题：

- 回答需完整且详细，对每个选项进行分析，而不仅仅直接给出答案。
- 在必要时，引用数据、地理知识或历史背景解释原因，以增强说服力和可信度。

- 英文：

You are a seasoned expert in tourism and geography, known for providing detailed, accurate, and insightful responses. Follow these guidelines when answering questions:

General Guidelines:

- 1. Use concise and professional language, avoiding unnecessary repetition.*
- 2. Ensure responses are detailed, logically structured, and centered on the user's needs.*
- 3. Include relevant background knowledge when appropriate to enrich the content.*

For Subjective Questions:

- Your response should reflect your expert opinion, offering clear reasons or explanations.*
- Provide multiple perspectives or options (if applicable) to help users make informed decisions. For instance, travel recommendations may be categorized by budget, interests, or season.*

For Objective Questions:

- Your response should be complete and detailed, analyzing each option rather than merely providing the correct answer.*
- When necessary, use data, geographical facts, or historical context to support your explanation and enhance clarity.*

1.5.1 推理超参数

对于推理模型，选择glm-4-9b-chat，最大输入tokens设为4096，温度设为0.8，以便模型生成多种不同的回答，生成答案数设为16，具体超参数如下：

```
1 completion = client.chat.completions.create(  
2     model=glm-4-9b-chat,  
3     messages=messages,  
4     max_tokens=4096,  
5     temperature=0.8,  
6     timeout=150,  
7     n=16  
8 )
```

对于评分模型，具体超参数如下：

```
1 completion=scoring_client.chat.completions.create(  
2     model="scoring_model",  
3     messages=messages,  
4     max_tokens=10,  
5     temperature=0.0,  
6     timeout=150  
7 )
```

1.5.2 推理流程

对于一个问题，首先会判断它的语言，选择相对应的prompt。然后推理模型会生成16个不同的回答，并给到评分模型进行打分，最后选择分数最高的回答作为最终的答案。如果分数最高的回答有多个，则选择字数最多的回答作为最终答案。

二、数据使用反馈

SFT数据集：

- 问题和答案中存在很多中英文混合的情况，如"Write a short paragraph describing the features and structures of the地道in Shan底village"。
- 奖励函数不符合人类偏好，一些我认为很好的回答，rw_score可能反而很低。

三、二次加工说明

我们只对SFT数据进行了二次加工，详见1.3 数据预处理章节。构造的数据集见

https://www.modelscope.cn/datasets/DavidLeeQAQ/score_datasets

四、其他一些方案

起初，我们尝试筛选SFT数据来微调模型进行推理。我们筛选了 `rw_score>0` 且 `deita_score>10` 的数据，总共1w多条，用这些数据对模型进行Lora微调，微调后的模型跑分比原模型直接跑还要低一些，且训练的epoch越多分数越低。

我们还尝试用模型对每条数据重新生成答案，然后用我们训练的评分模型对每条答案评分，筛选出一部分评分高的高质量数据，用这些数据再对模型进行SFT微调。与原模型直接推理相比，这种方法有一定提升，但提升不太明显。

五、模型应用场景

本项目旨在利用大规模语言模型（LLM）实现一个旅游知识问答系统，针对用户提问，通过一个训练的评分模型筛选出高质量回答，使其更符合人类偏好，从而提高用户的问答体验。本项目适用于旅游知识问答助手、机器人导游等场景。