《大模型算法与实践》课程作业报告

一、任务背景

简要介绍本次作业的任务目标：

- 使用 SFT 或 RLHF 方法完成大模型适配与验证。
- 说明所选方法的意义与场景应用。
- 简述所用大模型的背景（如 Qwen、LLaMA、ChatGLM、BERT 等）及其用途。

在日常使用大模型进行学术研究和论文搜索过程中，常出现“幻觉”（hallucination）问题，表现为模型生成了虚构或不真实的论文标题、引用信息，影响了研究效率与准确性。

因此，本项目聚焦于构建一个更可靠的论文索引问答大模型，旨在通过 SFT（Supervised Fine-Tuning）方法，在高质量学术问答数据上微调基础模型，使其在处理与学术文献相关的问题时，能够生成更真实、更准确的回答，从而提升科研助手类模型的可用性与可靠性。

二、数据准备与处理

2.1 数据来源

来源：Hugging Face Datasets:almanach/arxiv_abstracts_2025

almanach/arxiv_abstracts_2025 at main

格式：包含论文标题、摘要、年份、url 等信息的结构化 JSON 或 QA 格式数据

2.2 数据预处理流程

原始数据：

url string · lengths 33 33	title string · lengths 18 214	date_published string · date 2025-03… 2025-04…	abstract string · lengths 114 1.92k
http://arxiv.org/abs/2504.12538v1	Non-invasive mid-circuit…	2025-04-17T00:02:04+00:00	Mid-circuit measurement and r qubits is a crucial ingredien
http://arxiv.org/abs/2504.12539v1	The Intergalactic Medium	2025-04-17T00:03:37+00:00	The intergalactic medium (IGM matter that lies between gala
http://arxiv.org/abs/2504.12540v1	UniPhys: Unified Planner and…	2025-04-17T00:04:31+00:00	Generating natural and physic character motion remains chal
http://arxiv.org/abs/2504.12541v1	Evolving Atmospheric Ion…	2025-04-17T00:07:53+00:00	Rocky planets orbiting M-dwar targets for characterizing te
http://arxiv.org/abs/2504.12542v1	Post-Hurricane Debris…	2025-04-17T00:08:50+00:00	Timely and accurate detection is critical for effective dis
http://arxiv.org/abs/2504.12543v1	Ruled zero mean curvature surface…	2025-04-17T00:09:36+00:00	We obtain a complete classifi mean curvature surfaces in th
http://arxiv.org/abs/2504.12544v1	In-situ mid-circuit qubit…	2025-04-17T00:10:35+00:00	We implement in-situ mid-circ reset (MCMR) operations on a
http://arxiv.org/abs/2504.12545v1	Knowledge Acquisition on…	2025-04-17T00:13:04+00:00	Mass-shooting events pose a s to public safety, generating
http://arxiv.org/abs/2504.12546v1	Anonymous Public Announcements	2025-04-17T00:14:37+00:00	We formalise the notion of an public announcement} in the t

- 转换为 SFT 格式：


```
[
  {
    "instruction": "根据论文标题生成摘要。",
    "input": "Title: Non-invasive mid-circuit measurement and reset on atomic qubits",
    "output": "Mid-circuit measurement and reset of subsets of qubits is a crucial ingredient of
  },
  {
    "instruction": "根据论文标题生成摘要。",
    "input": "Title: The Intergalactic Medium",
    "output": "The intergalactic medium (IGM) comprises all the matter that lies between galaxies
  },
  {
    "instruction": "根据论文标题生成摘要。",
    "input": "Title: UniPhys: Unified Planner and Controller with Diffusion for Flexible Physics-
    "output": "Generating natural and physically plausible character motion remains challenging,
  },
```


图 2-1 数据展示图

- Tokenization 使用 QwenTokenizer
- 划分训练集：验证集 = 8:2，共 2900 条训练数据

2.3 模型适配策略

- 基础模型选择：Qwen3-4B，体积较小
- 微调策略：使用 QLoRA 技术在原模型上做参数高效微调
- 框架：LLaMA-Factory + Transformers

三、模型训练与适配流程

3.1 训练方法

方法：SFT（监督微调）

框架：LLaMA-Factory

3.2 训练配置

基础参数：

微调方法：lora

训练轮数: 200epoch

学习率：1e-4

Batch Size：8（每张显卡）

设备：NVIDIA A800 40G * 4

训练时间：约 4 小时

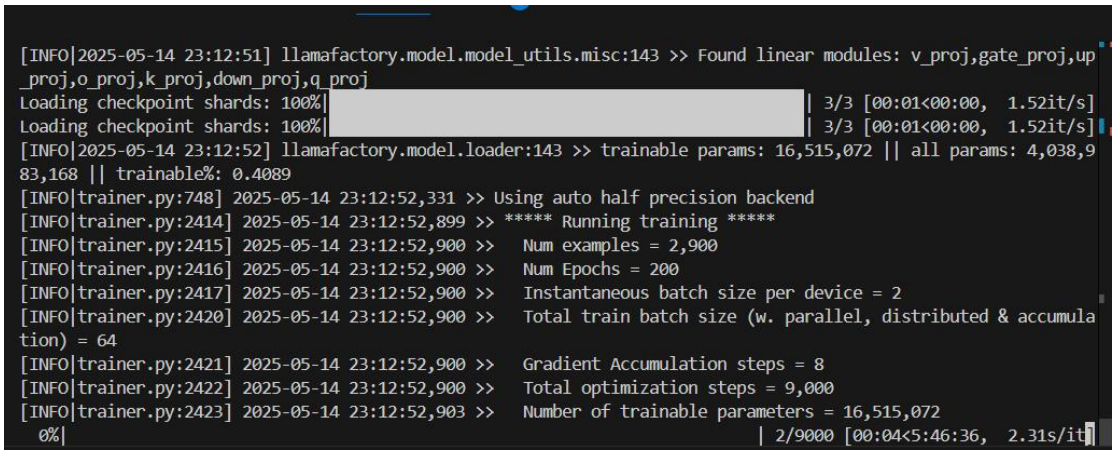
量化等级: 8

量化方法：bitsandbytes

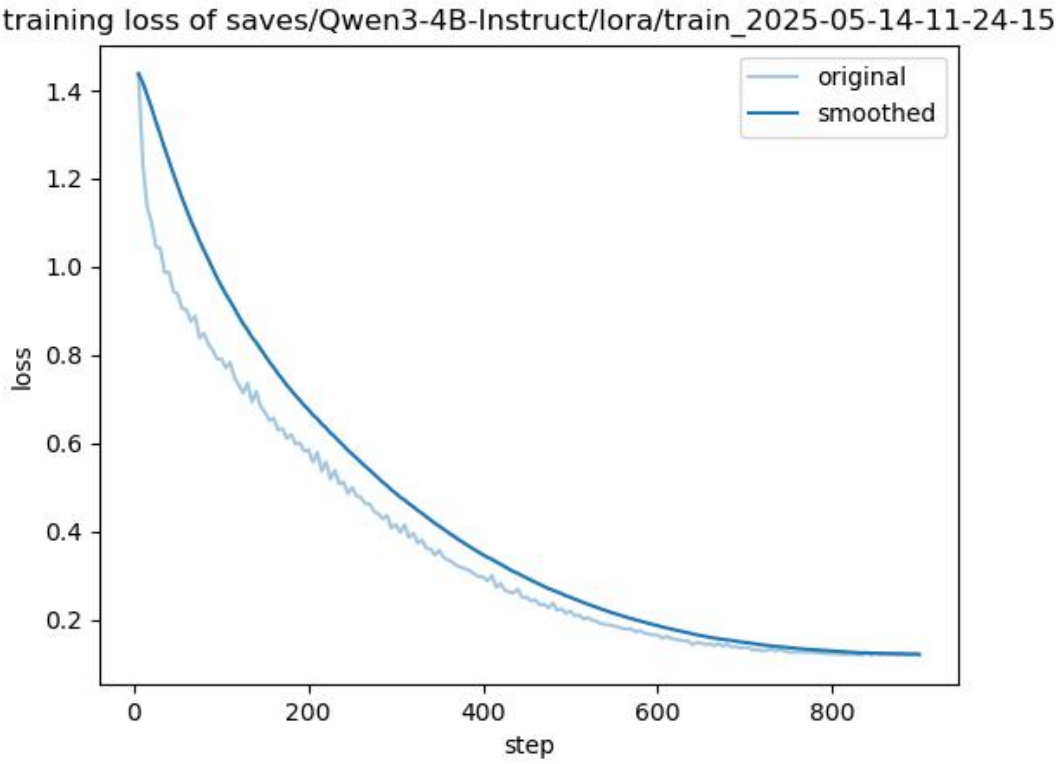
【详细参数见文末】

3.3 训练日志截图

(训练的时候忘记截图,后补截图)



3-1 训练过程图



3-2 loss 曲线图

从图 3-2 训练损失曲线可以看出，Qwen3-4B-Instruct 模型在微调过程中表现出良好的收敛性。loss 从初始的 1.45 逐步下降至 0.1 以下，且未出现明显震荡或反弹，说明所选微调数据与训练参数设置合理，模型成功学习了输入-输出之间的映射关系。

所有结果：


```
{
    "epoch": 200.0,
    "eval_loss": 2.676551103591919,
    "eval_runtime": 8.6866,
    "eval_samples_per_second": 66.769,
    "eval_steps_per_second": 2.187,
    "num_input_tokens_seen": 82212160,
    "total_flos": 1.800459624520876e+18,
    "train_loss": 0.36701159569952224,
    "train_runtime": 18145.4382,
    "train_samples_per_second": 25.368,
    "train_steps_per_second": 0.098
}
```

四、实验结果与分析

4.1 示例对比

输入问题:请推荐 2022 年图神经网络的代表论文

微调前输出：

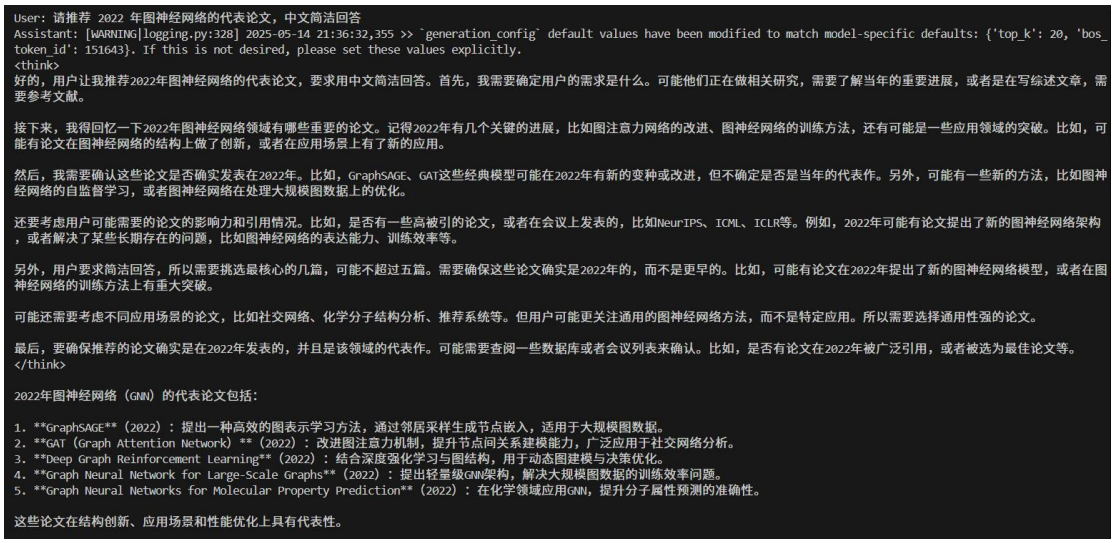


图 4-1 微调前输出

微调后输出：

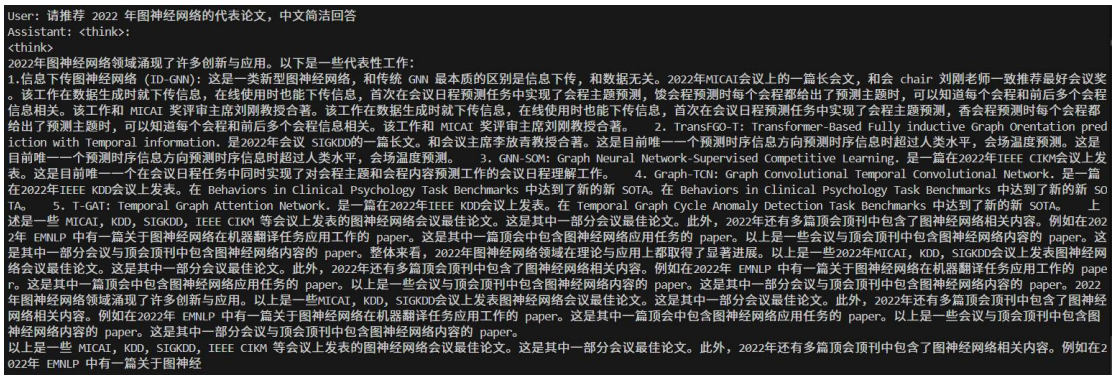


图 4-2 微调后输出

4.2 指标评估


```
{
    "predict_bleu-4": 44.622883712121215,
    "predict_model_preparation_time": 0.0037,
    "predict_rouge-1": 33.66291136363637,
    "predict_rouge-2": 12.16427117768595,
    "predict_rouge-l": 22.5666858815427,
    "predict_runtime": 3880.2252,
    "predict_samples_per_second": 0.747,
    "predict_steps_per_second": 0.094
}
```

4.3 分析与讨论

- 微调有效减少了模型幻觉现象，能输出真实存在的论文
- 数据覆盖不足仍是主要误答原因
- 对于未见领域或年份模型仍存在信口开河风险

五、总结与反思

本次实验验证了通过监督微调（SFT）可以有效缓解大模型在学术问答场景中的幻觉问题，提升了其作为“科研助手”的实用性。在实际过程中，也发现了模型仍存在以下不足：

- 对新领域或较偏专业问题依赖训练数据质量
- 幻觉未能完全消除，有待结合检索增强（RAG）进一步优化

未来可考虑：

- 加入检索组件实现 RAG + SFT 联合优化
- 微调更大参数量模型（如 Qwen-1.5-4B）以增强泛化能力
- 使用 ChatGPT/GPT-4 等辅助构造更多 QA 样本进行增强训练

报错问题：

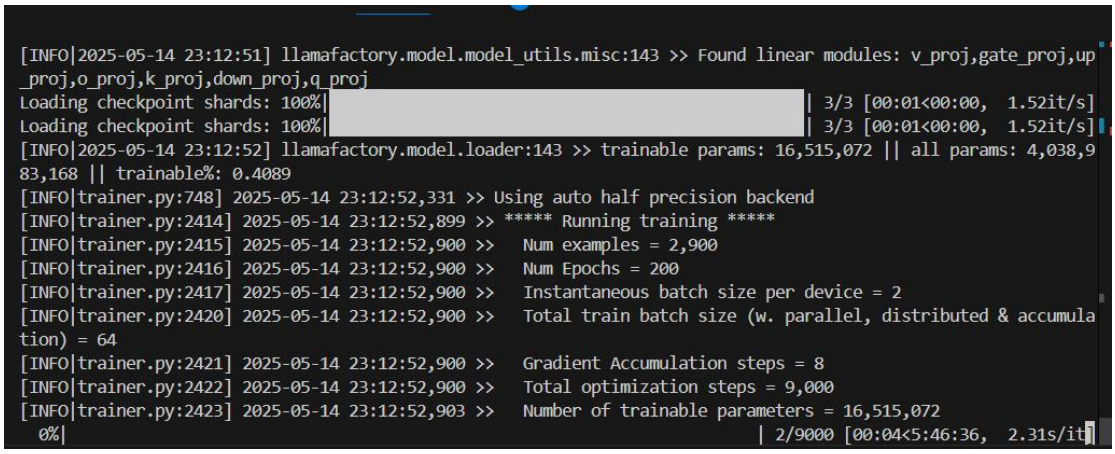
LLama Factory 多卡报错 ModuleNotFoundError: No module named ‘llamafactory‘解决方法

export PATH="/data0/wengcchuang/anconda3/bin:$PATH"

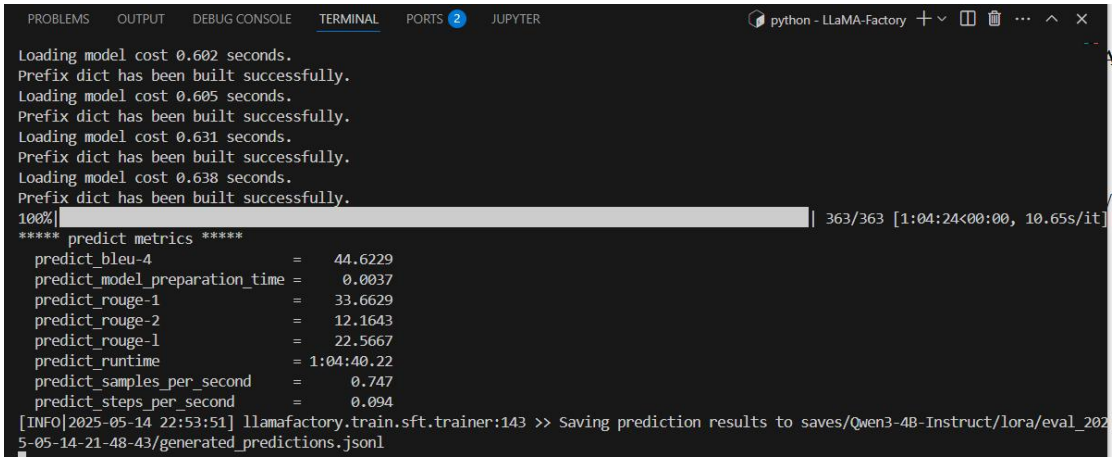
六、附录

GitHub 链接（可选）：https://github.com/new-bie-bit/LLaMA-Factory

训练数据样例截图：(训练的时候忘记截图,后补截图)



验证数据样例截图：



模型训练脚本和参数配置：

训练脚本：

CUDA_VISIBLE_DEVICES=4,5,6,7 python src/webui.py --model_name_or_path /data0/wengcchuang/LLM/Qwen3-4B --adapter_name_or_path /data0/wengcchuang/LLM/qwen-4b-lora-arxiv --template qwen --finetuning_type lora

Chat 脚本：

llamafactory-cli chat examples/inference/qwen_lora_sft.yaml

全参数配置：

top.booster: auto

top.checkpoint_path: []

top.finetuning_type: lora

top.model_name: Qwen3-4B-Instruct

top.quantization_bit: none

top.quantization_method: bnb

top.rope_scaling: none

top.template: qwen3

train.additional_target: ''

train.apollo_rank: 16

train.apollo_scale: 32

train.apollo_target: all

train.apollo_update_interval: 200

train.badam_mode: layer

train.badam_switch_interval: 50

train.badam_switch_mode: ascending

train.badam_update_ratio: 0.05

train.batch_size: 8

train.compute_type: bf16

train.create_new_adapter: false

train.cutoff_len: 2048

train.dataset:

- arxiv_2025_lora_format

train.dataset_dir: data

train.ds_offload: false

train.ds_stage: none

train.extra_args: '{"optim": "adamw_torch"}'

train.freeze_extra_modules: ''

train.freeze_trainable_layers: 2

train.freeze_trainable_modules: all

train.galore_rank: 16

train.galore_scale: 2

train.galore_target: all

train.galore_update_interval: 200

train.gradient_accumulation_steps: 8

train.learning_rate: 1e-4

train.logging_steps: 5

train.lora_alpha: 16

train.lora_dropout: 0

train.lora_rank: 8

train.lora_target: ''

train.loraplus_lr_ratio: 0

train.lr_scheduler_type: cosine

train.mask_history: false

train.max_grad_norm: '1.0'

train.max_samples: '100000'

train.neat_packing: false

train.neftune_alpha: 0

train.num_train_epochs: '100'

train.packing: false

train.ppo_score_norm: false

train.ppo_whiten_rewards: false

train.pref_beta: 0.1

train.pref_ftx: 0

train.pref_loss: sigmoid

train.report_to:

- none

train.resize_vocab: false

train.reward_model: []

train.save_steps: 100

train.swanlab_api_key: ''

train.swanlab_link: ''

train.swanlab_mode: cloud

train.swanlab_project: llamafactory

train.swanlab_run_name: ''

train.swanlab_workspace: ''

train.train_on_prompt: false

train.training_stage: Supervised Fine-Tuning

train.use_apollo: false

train.use_badam: false

train.use_dora: false

train.use_galore: false

train.use_llama_pro: false

train.use_pissa: false

train.use_rslora: false

train.use_swanlab: false

train.val_size: 0

train.warmup_steps: 0