

# lab4

## 环境配置

1. 在阿里云平台购买A10加速 GPU，硬盘设置60G，选择对应的N卡驱动，弹性公网（可以没有）
2. pycharm链接远程解释器
3. 在服务器安装git，conda，配置对应环境，安装相关包
4. 按照swift官网，clone项目，并执行安装命令

## 模型选择

qwen1.5-1.8b模型，在A10加速GPU上训练，应用实验文档中的参数微调2小时左右

## 数据集准备

采用单轮对话数据，做二分类任务

由于框架本身支持HC3-zh的数据集，可以不用特意修改swift/llm/data文件夹下的dataset\_json文件，划分采用命令行自动划分，设置数据集5000.

## 微调

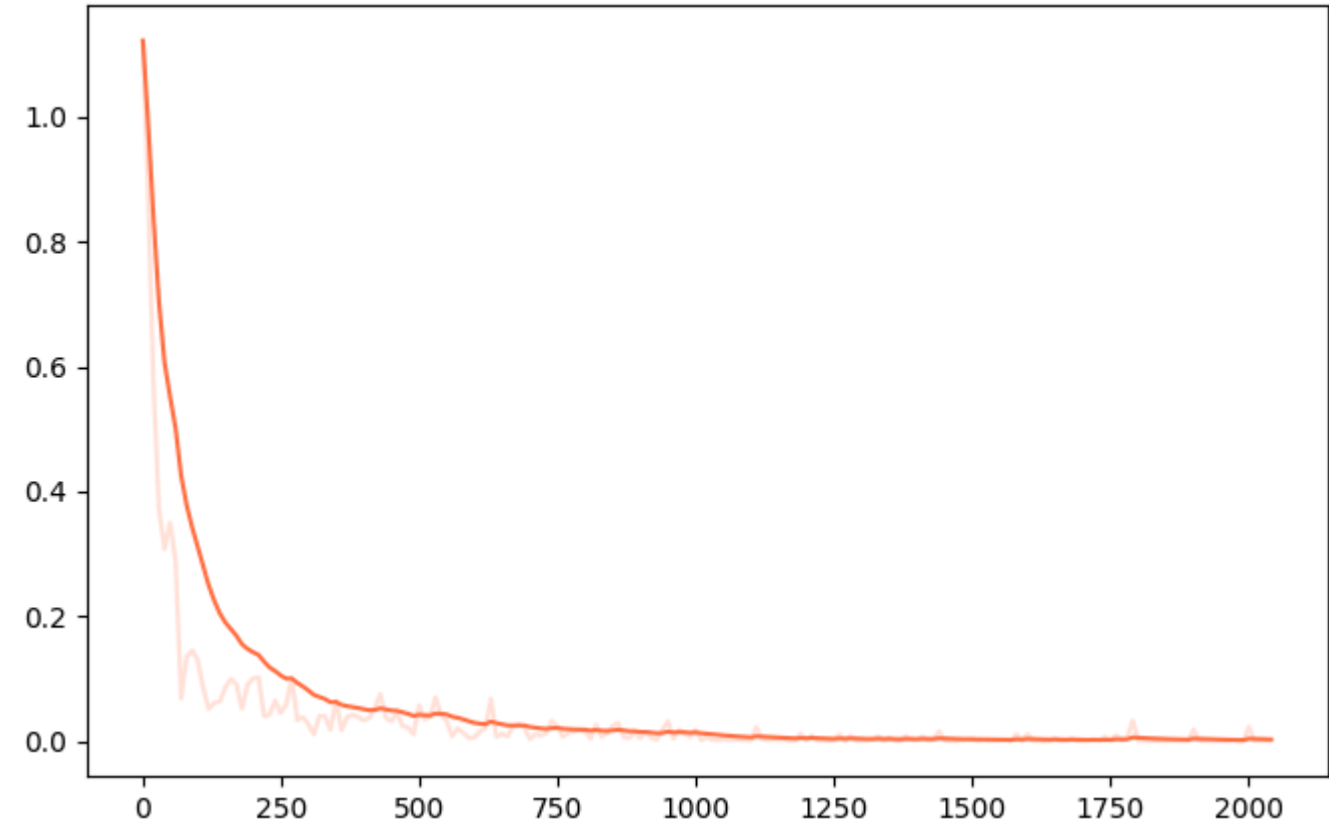
采用LORA微调

具体参数设置如下，基于swift的

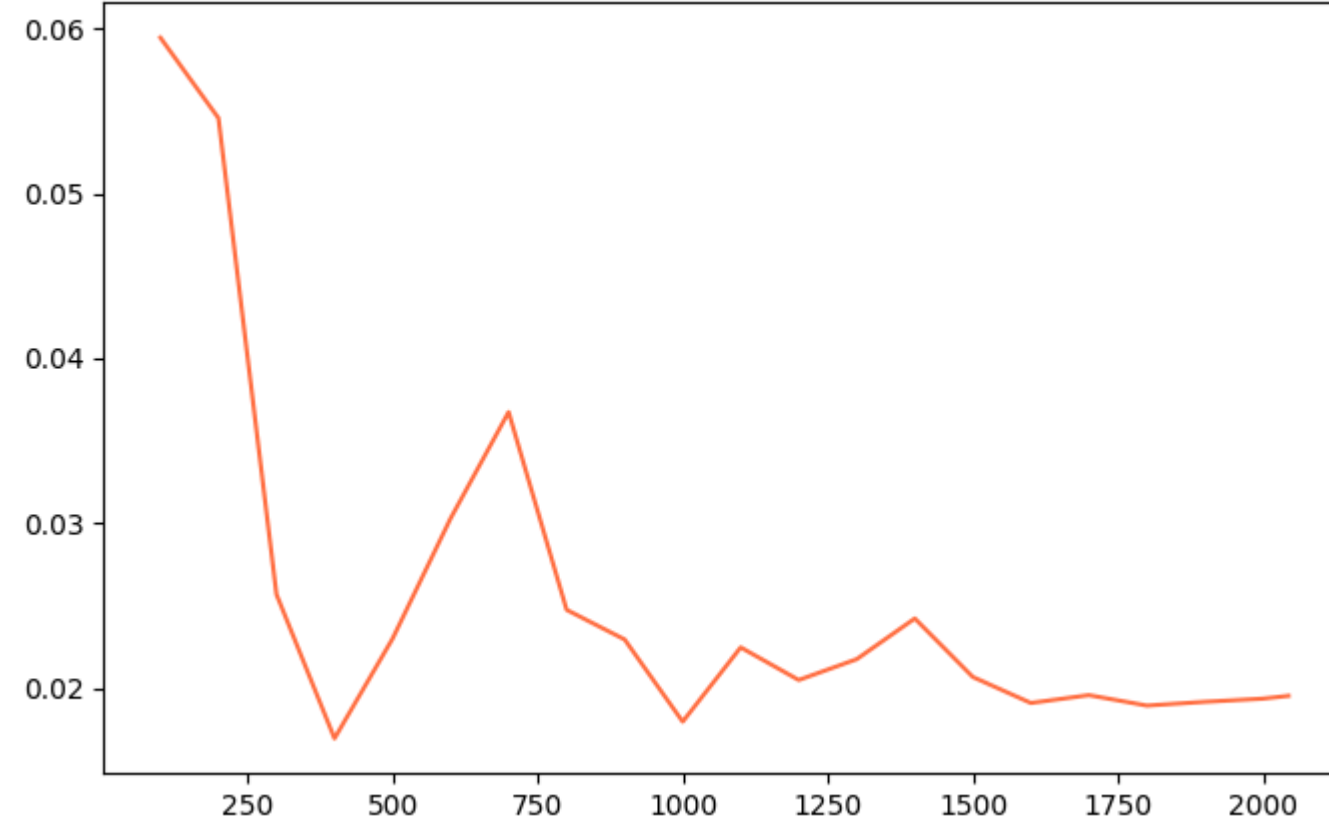


微调结果

train/loss



eval/loss



```
--ckpt_dir 'xxx/vx-xxx/checkpoint-xxx' --merge_lora true
```

## 评测

### 1. 使用vllm加速

# 原始模型

```
CUDA_VISIBLE_DEVICES=0 swift eval --model_type qwen1half-1_8b-chat \
  --eval_dataset no \
  --infer_backend vllm \
  --custom_eval_config ../data/eval_config.json
```

# LoRA微调后

```
CUDA_VISIBLE_DEVICES=0 swift eval --ckpt_dir output/qwen1half-1_8b-chat/v0-2
  --eval_dataset no --infer_backend vllm \
  --merge_lora true \
  --custom_eval_config ../data/eval_config.json
```

### 2. 自定义评测集的格式, prompt

csv的格式为id,question,A,B,C,D,answer,explanation

question = f"针对一个计算机知识的提问所给出的回答: {answer}这句话是由\_\_\_\_回答的。"

A为human

B为chatgpt

### 3. 修改对应配置文件

```
[
  {
    "name": "ceval",
    "pattern": "ceval",
    "dataset": "../data/ceval",
    "subset_list": ["default"]
  }
]
```

## 结果

### 原始模型

```
ts": "/.cache/llmuses/outputs/eval_ceval_qwen1half-1_8b-chat-20240628-103509_default", "mem_cache": False, "use_cache": False, "stage": "all",
'dataset_hub': 'Local', 'dataset_dir': '', 'limit': None}
2024-06-28 10:35:33,965 - llmuses - INFO - *** Report table ***
+-----+-----+
| Model | ceval |
+-----+-----+
| 8b-chat-20240628-103509 | (ceval/acc) 0.38 |
+-----+-----+
[INFO:swift] Final report:[{'name': 'ceval', 'metric': 'WeightedAverageAccuracy', 'score': 0.38, 'category': [{'name': 'DEFAULT', 'score': 0.38,
'subset': [{'name': 'default', 'score': 0.38}]}], 'total_num': 100}]

[INFO:swift] result: {'ceval': 0.38}
```

正确率0.38

微调后

```
2024-06-28 10:35:59,527 - llmuses - INFO - *** Report table ***
+-----+-----+
| Model           | ceval           |
+-----+-----+
| 8b-chat-20240628-103541 | (ceval/acc) 0.44 |
+-----+-----+

[INFO:swift] Final report:[{'name': 'ceval', 'metric': 'WeightedAverageAccuracy', 'score': 0.44, 'category': [{'name': 'DEFAULT', 'score': 0.44, 'subset': [{'name': 'default', 'score': 0.44}]}]}, {'total_num': 100}]

[INFO:swift] result: {'ceval': 0.44}
[INFO:swift] save_result_path: /tmp/pycharm_project_568/src/output/qwen1half-1.8b-chat/v0-20240627-170756/checkpoint-400-merged/eval_result.json
```

正确率0.44