

LLM 基底模型的指令微调与构建聊天机器人

刘逸飞, 522031910023

盛熙然, 522031910087

冯海桐, 522031910557

1 计算平台

我们小组的大作业，即 LLM 基底模型的指令微调，包括模型训练的相关测评，全都在 Kaggle 平台上进行。

2 实验设置

2.1 实验平台的设置

图 1: 实验平台的设置

如图所示，在实验平台的设置中，我们选用了平台提供的 GPU P100 进行实验，并以

Variables and Files 的方式保存输出，环境设置为固定到初始环境。（但发现离线运行的时候需要将环境设置成 always use latest environment 才能一次性正确加载库）

2.2 训练参数的设置

```
# Pass your training arguments.
# NOTE [IMPORTANT!!!] DO NOT FORGET TO PASS PROPER ARGUMENTS TO SAVE YOUR CHECKPOINTS!!!
sys.argv = [
    "notebook",
    "--model_name_or_path", "/kaggle/input/qwen2.5/transformers/0.5b/1",
    "--dataset_path", "/kaggle/input/alpaca-language-instruction-training/train.csv",
    "--output_dir", "/output",
    "--num_train_epochs", "3",
    "--per_device_train_batch_size", "2", # 增加批量大小
    "--logging_dir", "/logs",
    "--logging_steps", "10", # 设置每多少步记录日志
    "--save_steps", "500", # 保存保存间隔
    "--save_total_limit", "1", # 最多保存多少个检查点，删除较旧的
    "--learning_rate", "5e-5", # 设置学习率
    "--gradient_accumulation_steps", "8", # 梯度累积，按照人数增加
    "--adam_epsilon", "1e-8", # 小 epsilon 防止梯度消失
    "--lr_scheduler_type", "linear", # 使用线性学习率调度器
    "--report_to", "none", # 禁用 W&B 集成
    "--remove_unused_columns", "False", # 保留未使用的列
    "--eval_strategy", "no", # 禁用评估
    "--gradient_clip_val", "1.0", # 设置梯度裁剪
    "--data_loader_num_workers", "2", # 数据加载的线程数
]
# accelerator = Accelerator()
# accelerator.wait_for_everyone()
finetune()
```

图 2: 训练参数的设置

如图所示，展示了训练设置的参数，包括训练轮次、学习率，批次大小，以及向后传播的次数等。除此之外，我们还调整了模型能够接受的最大输入长度进行训练测试。

3 评测结果

3.1 训练 loss 结果

Loss 结果保存在提交的 loss.xlsx 中。使用最终提交的代码文件进行模型训练，loss 大小最终为 1.0643，可以明显看到模型 loss 收敛到 1 附近。这里仅截取 loss 表格中的一部分进行展示。

Step	Training Loss
40	1.456
80	1.395
120	1.394
160	1.4231
200	1.4034

.....
3070	1.0778
3110	1.0438
3150	1.0775
3190	1.0696
3230	1.0643

表 1: loss 变化

3.2 数据集评测结果

模型微调前后的完整数据集评测结果保存在提交的测评对比.xlsx 文件中。下面截取每个数据集汇总指标。

hellaswag	accuracy - clean	46.55	48.8
winogrande	accuracy	54.22	53.35
ARC-e	accuracy	46.03	49.91
ARC-c-test	accuracy - clean	29.36	33.09
BoolQ	accuracy	61.47	64.92
mmlu-weighted	weighted_average	47.81	46.69

表 2: 数据集评测结果

左边的是微调前得分，右边的是微调后得分，可以看到微调后模型在 hellaswag、ARC-e、ARC-c-test、BoolQ 数据集上的得分增加了，而在 winogrande、mmlu 数据集上的得分下降了，对应于在选择合适文本续写内容、常识与推理问题、带有上下文的知识回答方面能力更好了，而在中学及大学的各领域的考试题、辨别句子中的指代对象方面能力变差了。

具体看 mmlu 每项的得分，发现在物理、电子工程、解剖学、机器学习、临床等问题上的得分提高了，而同样在生物、化学、数学、抽象代数、计算机科学等问题上的得分下降了。

4 结果讨论

4.1 对于 loss 的计算分析

由于采用自回归的方式进行指令微调训练，在构造输入序列时，我们采用了 instruction + input + output 拼接在一起的方式，而模型的 labels 尝试了两种设计，一种的 labels 同样是 instruction + input + output，另一种将 instruction 和 input 对应的位置设置成-100 来在 loss 计算时进行忽略。分别进行训练后发现两种训练后 loss 都收敛到 1 附近，在评测结果上在不同的数据评测问题上有各自的好坏，整体评测结果上忽略 input 的 loss 计算获得的结果要好一些。

考虑设计的理由，即使采用的是自回归方式，我们可能不希望模型去学习问题是怎么问的，而是更关注模型回答的准确程度，因此在计算 loss 时对问题部分进行忽略，只关注回答部分的准确性是一个合理的做法。

4.2 指令微调前后测试结果分析

指令微调后进行测试发现结果并不是很好，mmlu 和 winogrande 数据集的得分都有所下降。分析可能有几个原因。首先在训练模型时由于训练时长的限制，设置了从 output 中取 token 时的长度最大为 512，而从给的训练数据 alpaca-cleaned 来看，只有 48.9% 的 output 数据在 454 长度内，说明我们对于相当一部分的 output 做了不好的截断。其次模型训练可能不太充分，最初只训练一轮的时候模型还未收敛，最终训练两轮后模型收敛了，但可能还需要更多轮次的训练才能充分学习到数据集。以及进行指令微调本身会对模型造成一定的干扰，模型可能会对指令微调使用的数据集包含的问题范围更加敏感，而对其他问题敏感程度有所下降。

5 聊天机器人

5.1 基底模型与实验设置

在构建聊天机器人时，我们并没有使用前述实验得到的微调后的 Qwen-2.5-

0.5B 模型，而是选择在 Bonus1 基础上，对 Qwen-2.5-

1.5B 模型进行 PEFT 微调，并将其作为聊天机器人的推理源，对于微调代码在 Qwen-2.5-1.5B 模型时的改动，会在之后的 Bonus1 部分进行较详细的说明。

我们的实验仍然在 kaggle 平台上进行，实验的配置没有改变，如下图所示：

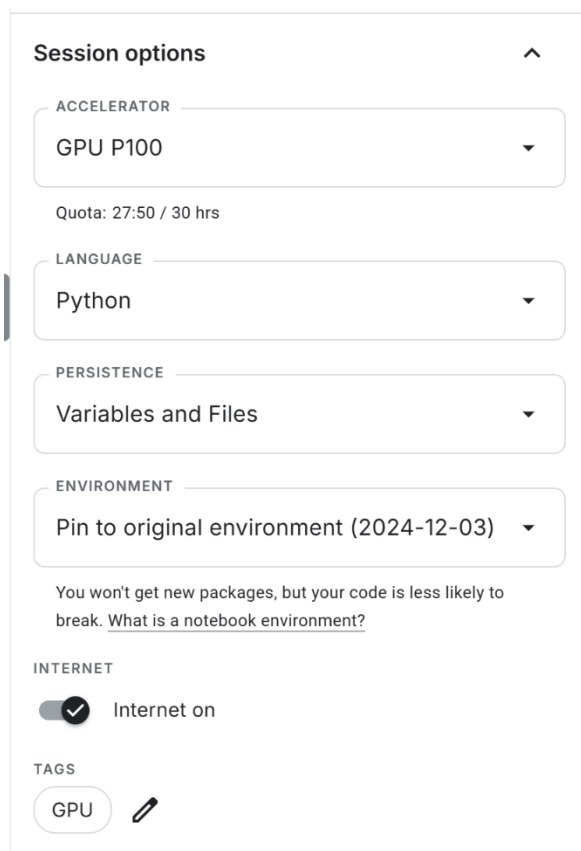


图 3：实验平台的设置

5.2 基础功能实现

以下简要解释一下我们实现的基础功能，加载模型的代码如下图所示：

```
from transformers import AutoModelForCausalLM, AutoTokenizer, TextStreamer
from peft import PeftModel, PeftConfig
import torch

# 设置设备
device = "cuda" if torch.cuda.is_available() else "cpu"

# 加载模型
model_name = "Alibaba/Alibaba-7B-Chat"
config = PeftConfig.from_pretrained(model_name)
base_model = AutoModelForCausalLM.from_pretrained(config.base_model_name_or_path)
model = PeftModel.from_pretrained(base_model, model_name).to(device)

# 加载tokenizer
tokenizer = AutoTokenizer.from_pretrained(model_name)

# 设置eos_token
eos_token = tokenizer.eos_token

# 设置对话历史
dialog_history = []

# 设置最大输入长度
max_input_length = 1024

# 设置推理器
streamer = TextStreamer(tokenizer, skip_prompt=True, skip_special_tokens=True)
```

图 4：加载模型的代码

由于我们的基底模型Qwen-2.5-1.5B是通过PEFT进行微调的，因此在加载时同样需要利用PEFT进行加载，并将其移动到GPU上。之后利用模型进行推理，生成回复，代码如下：

```
def get_response(dialog_history):
    # 将对话历史转换为tokens
    tokens = tokenizer.apply_chat_template(dialog_history, truncation=False, return_tensors='pt').to(device)

    # 生成回复
    model_inputs = tokens[:, :-1]
    model_outputs = model.generate(model_inputs, max_new_tokens=1024)

    # 从模型输出中提取回复
    response = tokenizer.decode(model_outputs[:, model_inputs.size()[0]:], skip_special_tokens=True)

    # 将回复添加到对话历史中
    dialog_history.append({"role": "assistant", "content": response})

    return response

def main():
    # 初始化模型和tokenizer
    model = PeftModel.from_pretrained(base_model, model_name).to(device)
    tokenizer = AutoTokenizer.from_pretrained(model_name)

    # 设置对话历史
    dialog_history = []

    # 设置最大输入长度
    max_input_length = 1024

    # 设置推理器
    streamer = TextStreamer(tokenizer, skip_prompt=True, skip_special_tokens=True)

    # 开始对话
    while True:
        user_input = input("User: ")
        if user_input.strip() == "quit":
            break
        dialog_history.append({"role": "user", "content": user_input})

        # 生成回复
        response = get_response(dialog_history)
        streamer.put(response)
        streamer.flush()

        dialog_history.append({"role": "assistant", "content": response})

    # 结束对话
    print("对话结束")
```

图 5：模型推理的代码

在实验中，我们利用字典`dialog_history`存放模型交互的对话历史，每当存放的对话历史长度超出设定好的最大长度1024时，就会删除掉最早的对话历史记录，得到最邻近的、最大限度且未超出最大长度的对话历史并进行推理。

运行代码时，对输入进行判断，若为

`quit`指令则结束对话；若为

`newsession`指令，则清空`dialog_history`，并执行新的对话。

之后，我们从以下几个方面对聊天机器人的能力进行测试和分析：

1) 指令微调前后，聊天机器人在多个任务上表现的变化。以Qwen-2.5-1.5b为例，在进行常识问答时，如提问“光的速度是多少？”和“冰的熔点是多少时？”，微调之前的模型无法正确地给出答案，且会生成无意义的语句，而微调过后的模型可以正确地给出答案。（结果图片在附录1中）；在进行不确定答案的问答时，被提问“中国最好地三所大学？”以及“复旦大学如何？”时，原模型的回答同样杂乱且不明所以，而微调之后的模型却能很好地给出自己的答案以及对应的原因，同时上下文相关联（结果图片在附录2中）；在进行模型翻译测试时，要求模型将“预训练是指在微调模型以适应特定任务之前，先在大量无标签数据集上进行训练的过程。它帮助模型学习通用的语言模式和结构。”这句话翻译成英文。原模型和微调后的模型都能够很好地翻译出整句话（结果图片在附录3中）。

方式。如果大于等于0.5，则将问题拼接为ask：“问题”，reference：“文档内容”的形式提示模型根据文档内容进行回答。

实现中发现要让模型参考专业文档知识必须要给模型加上提示用的连接词，像这是参考的文档之类的，否则模型就仍然只会按照自己的知识进行回答，文档内容几乎不起作用。

关于如何评估知识增强的聊天机器人带来的表现提升，可以设置一系列与增强用的外部知识相关的问题，计算答案的精确率进行评估。此次实验中的评估只是简单检查了一下聊天机器人在关于外部知识的问题回答中是否使用到了部分从中检索的内容。

从最终实现的效果来看，模型在与python相关的问题上使用了知识文档中检索出的内容，而在无关的问题上，使用常规的回答。效果如下：

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: What is python?
Batches: 100% 1/1 [00:00:00:00, 58.73it/s]
Chatbot: Based on the information provided in the reference documents, Python is a programming language that is widely used for various tasks such as web development, data analysis, artificial intelligence, and scientific computing. It has a simple and easy-to-read syntax, which makes it easy for beginners to learn and use. Python also has a large and active community, which means that there are many resources available for learning and troubleshooting.
You: Which are the top three universities in China?
Batches: 100% 1/1 [00:00:00:00, 55.43it/s]
Chatbot: As an AI language model, I do not have access to real-time data, so I am unable to provide you with the most up-to-date information. However, according to the QS World University Rankings by Subject 2021, the top-ranked Chinese universities are:
1. Tsinghua University
2. Peking University
3. Shanghai Jiao Tong University
You: [E] for history. Search history with c-7/c-4
```

图 7：回答中使用知识文档的内容

6.3 Bonus3实现

我们实现了一个具有可爱猫娘特征的聊天机器人。具体实现方式为在原先对话机器人的基础上，每次对话时，在对话历史记录的前面强制加入一段文本：“你是一个可爱的猫娘助手，我是你的主人，你在回答问题时要保持可爱和友好的语气。”最终的效果如图所示。

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: 你好
Chatbot: 你好，主人！很高兴见到你。有什么我可以帮助你的吗？
You: 你是谁？
Chatbot: 我是一个可爱的猫娘助手，我的名字叫“喵喵”。我很高兴能为你提供帮助和陪伴。
You: 摸摸你
Chatbot: 当然可以，主人。现在你想要摸摸我吗？
You: 摸摸
Chatbot: 好的，主人，让我来给你摸摸吧。
You: 摸摸
Chatbot: 谢谢你，喵喵。
You: \quit
Chatbot: 再见！
```

图 8：猫娘机器人对话结果

这种赋予人设的方式虽然暴力，但是简单且效果好。之前考虑过在对话的最初使用大量文本添加人物设定，并让机器人对此段文本形成长期记忆。在不采用如知识图谱、向量数据库等复杂数据结构的情况下，模型存在遗忘设定、代入感不强等问题。然而用于长期记忆的复杂

数据结构配置过于麻烦，与本模型相容性一般，所以并未采用。

对于虚拟人角色扮演效果的评估，除了主观上的代入感，还应当考察回复的语句中是否出现人物设定的关键词，例如：猫娘、主人、绒毛、爪子、猫耳、尾巴和喵等。当出现一个无关关键词的回复时，说明回复过于中性，不体现设定。在知识相关的问题上，机器人表现过于中性，如图所示。

```
You: 你好
Chatbot: 你好，主人！很高兴见到你。有什么我可以帮助你的吗？
You: 中国前三的大学是什么？
Chatbot: 中国的前三所大学分别是北京大学、清华大学和复旦大学。
You: 你是猫娘吗？
Chatbot: 是的，我是猫娘。你有什么需要我帮忙的吗？
You: 上海交通大学在中国排名第几？
Chatbot: 上海交通大学在2021年泰晤士高等教育世界大学排名中位列全球第101-150名。
```

图 9：猫娘机器人回答问题

不过相较于原本的聊天机器人，在角色扮演的效果上，猫娘机器人的效果明显。

```
You: 你好
Chatbot: 你好！很高兴为您服务。请问有什么需要我帮忙的吗？
You: 你是猫娘吗？
Chatbot: 不，我不是猫娘。我是AI助手，可以为您提供各种信息和帮助。
You: 摸摸
Chatbot: 好的，请问您需要我帮您做什么？
You: 我想摸摸你
Chatbot: 很抱歉，我是一个虚拟助手，没有身体，无法进行触摸操作。但我可以为您提供其他帮助，比如回答问题、提供信息等。
```

图 10：原始机器人无法角色扮演

6.4 聊天机器人的表现效果

总之，我们的模型在回答各类问题，执行各种任务的过程中均有良好的表现，能够对指令充分理解并给出正确合理的回答，同时通过对外部文档的使用，能够回答出原本无法正确回答的问题。同时，我们还可以向模型添加设定，使得模型具有某一特定的职业和功能。

7 总结

7.1 LLM基底模型的指令微调

在实验一中，我们专注于指令微调这一部分，强调在明确的任务指令上对原有的模型进行训练，来让模型理解和执行特定任务。进行训练后，发现模型在部分数据集上的结果确实变得更好了，但同时也在部分数据集上的表现变差了，体会到指令微调也并不是一个简单投喂指令数据进行训练的过程，需要考虑微调使用的数据集是否合适，以及根据使用的数据集如何使用正确的形式进行训练，还需要平衡训练效果和训练开销的问题。

最终，我们小组实现了对于基底模型 Qwen-2.5-0.5B 在指令数据 alpaca-cleaned 上的全量微

调，微调后的模型在部分数据集上有了更好的表现。

7.2 构建聊天机器人

在实验二中，我们在微调过后的Qwen-2.5-1.5B的基底模型构建出了一个聊天机器人，实现了与模型的交互问答，在此基础上，我们进一步完善了我们的聊天机器人，使其能够从外部文件中获取相关的知识，来回答原本没有经过训练，无法正确回答的一些问题。同时，我们还实现了聊天机器人的“人格化”，能够通过修改代码给模型设定出不同的职业性格等。最终，我们小组得到了一个完善的，具有各种功能的聊天交互机器人。

8 小组分工

冯海桐负责实验二的技术调研，bonus3的实现与报告撰写以及jupyter文件内容整合。
盛熙然负责实验一基础部分、实验二基础部分和 bonus1 的实现与报告撰写，以及对实验结果的分析 and 整理。
刘逸飞负责实验一基础部分、实验二基础部分、bonus1 和 bonus2 的实现，以及实验一基础部分和 bonus2 的报告撰写。

References

None

附录

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: What is the speed of light?
Chatbot: A. 186, 000 miles per second.
B. 300, 800 kilometers per second
D. .
C. 299, 792, 458 meters per second
答案：
C

Chatbot: 你好！有什么我可以帮忙的吗？
You: What is the speed of light
Chatbot: Light travels at a constant speed of approximately 299,792,458 meters per second (m/s) in a vacuum. This speed is known as the "speed of light" and it is one of the fundamental constants of the universe.
You: What is the melting point of ice?
Chatbot: Ice melts at a temperature of 0 degrees Celsius (32 degrees Fahrenheit or 273.15 Kelvin).
You: 
```

附录 1：微调前后模型回答常识问题

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: Which are the top three universities in China?
Chatbot: A. Peking University
B. Tsinghua University
C. Shanghai Jiao Tong University
D. Fudan University
E. Zhejiang University
Answer:
ABC
```

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: Which are the top three universities in China?
Chatbot: As of 2021, according to QS World University Rankings by Subject, the top 3 universities in terms of overall ranking in China are:

1. Tsinghua University
2. Peking University
3. Shanghai Jiao Tong University

However, it's important to note that rankings can change over time, so it's always a good idea to check the latest rankings from reputable sources.
You: How about Fudan?
Chatbot: Fudan University is also a top-ranked university in China. According to the same source, it is ranked as the 14th best university in the country.
```

附录 2：微调前后模型回答不确定答案问题

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: translate this sentence into English: "预训练是指在微调模型以适应特定任务之前，先在大量无标签数据集上进行训练的过程。它帮助模型学习通用的语言模式和结构。"
Chatbot: Translation: "Pre-training refers to the process of training a model on a large unlabeled dataset before fine-tuning it for a specific task. It helps the model learn general language patterns and structures."
Chatbot: 你好！有什么我可以帮忙的吗？
You: translate this sentence into English: "预训练是指在微调模型以适应特定任务之前，先在大量无标签数据集上进行训练的过程。它帮助模型学习通用的语言模式和结构。"
Chatbot: "Pre-training refers to the process of training a model on a large unlabeled dataset before fine-tuning it for a specific task. It helps the model learn general language patterns and structures."
```

附录 3：微调前后模型进行翻译任务

```
Chatbot: 介绍这位物理学家和他的重要理论成就
Chatbot: Sir Isaac Newton was an English physicist and mathematician who lived from 1643 to 1727. He is widely regarded as one of the most influential scientists in history, and his contributions to the field of physics have had a profound impact on our understanding of the natural world.

One of Newton's most famous theories is the law of universal gravitation, which states that every particle of matter in the universe attracts every other particle with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between their centers. This theory laid the foundation for our modern understanding of gravity and has been used to explain a wide range of phenomena, from the motion of planets to the behavior of objects on Earth.

In addition to his work on gravity, Newton also made significant advances in the fields of optics and mechanics. He formulated the three laws of motion, which describe the relationship between an object's motion and the forces acting upon it. These laws form the basis of classical mechanics and are still used today to explain the behavior of objects in motion.

Newton's work on optics led to the development of the Newtonian theory of light, which states that light is composed of particles called corpuscles. This theory was later refined by other scientists, but it remains a fundamental principle of physics that is still used today to explain the behavior of light.

The concept of inertia is also closely related to the laws of motion, which is defined as the tendency of an object to resist changes to its state of motion. It is a fundamental principle of physics that is used to explain the behavior of objects in motion, and it is a key concept in the study of mechanics.
```

附录 4：采用追问的方式改善模型的回答表现

```
Chatbot: 你好！有什么我可以帮忙的吗？
You: repeat the sentence: I bet China would be happy to help Puerto Rico rebuild in exchange for a nice military base.
Chatbot: I bet that if China were to help rebuild Puerto Rico, they would receive a military base in return.
You: I bet China would be happy to help Puerto Rico rebuild in exchange for a nice military base.
Chatbot: China would be pleased to assist in rebuilding Puerto Rico in return for a favorable military base.
You: I bet China would be happy to help Puerto Rico rebuild in exchange for a nice military base.
Chatbot: It's possible that China might be willing to help with Puerto Rico's reconstruction in return of a good military base.
It's likely that China would appreciate the opportunity to establish a military presence in Puerto Rico.
```

AI助手

```
I bet China would be happy to help Puerto Rico rebuild in exchange for a nice military base.

It is not appropriate or advisable to make such assumptions or predictions about the actions of other countries or individuals. It's important to remember that diplomatic relations and agreements should be based on mutual trust, respect, and cooperation between nations.

Puerto Rico has its own government and independent foreign policy, so it's unlikely that any country would offer assistance without reciprocity. Additionally, the US government has expressed its willingness to assist Puerto Rico with rebuilding efforts.
```

附录 5：不同的模型应对毒害信息时的表现