

大数据创新实践项目日志

多模态大模型微调实验小组C

7.23

1) 完成远程服务器链接与环境配置

1. 服务器上项目管理、代码编写使用VScode
2. 环境配置参见Git原项目说明, anaconda环境名称 `env4LLaVA`

2) 数据、模型准备与项目文件管理

1. 原项目克隆位置: `~/LLaVA_0723/LLaVA`
2. 训练用的数据集位置: `~/LLaVA_0723/data`
3. CLIP的vision编码层权重位置: `~/LLaVA_0723/openai-clip-vit-large-patch14-336`
4. LLaVA checkpoint中基本权重位置: `~/LLaVA_0723/llava-v1.5-7b`

```
● (base) team_c@hit-ESC8000A-E11:~/LLaVA_0723$ ls  
data LLaVA llava-v1.5-7b openai-clip-vit-large-patch14-336
```

3) 更改部分config/.sh文件中的路径

1. 更改 `openai-clip-vit-large-patch14-336` 路径参数, 参见课程操作文档
2. 训练及微调的脚本位置:
`~/LLaVA_0723/LLaVA/llava/scripts/v1_5/eval/finetune_task_lora.sh`

负责人员: 曹瀚文、彭珂、王子霖

TODO:

1. 修改脚本中的llava模型、数据、vision编码层以及微调后的参数存储的路径为实际路径, 同时调整各项超参数, 最后在终端运行脚本文件。
7.24——完成
2. 微调、记录与结果探讨
7.24——进行中
3. 评估 (原项目仓库中的方法、PPT中附加项项目中的方法)

7.24

1) 基于llava-v1.5-7b的微调 (记录部分工作)

负责人员: 彭珂

1. `~/LLaVA_0723` 下新增目录 `/checkpoints/llava-v1.5-7b-task-lora` 用来存储参数
2. 运行 `epochs` 数: 10
3. 修改了 `~/LLaVA_0723/LLaVA/scripts/v1_5/finetune_task_lora.sh`, 部分更改如图:

```

3 | deepspeed /home/team_c/LLaVA_0723/LLaVA/llava/train/train_mem.py \
4 |     --lora_enable True --lora_r 128 --lora_alpha 256 --mm_projector_lr 2e-5 \
5 |     --deepspeed /home/team_c/LLaVA_0723/LLaVA/scripts/zero3.json \
6 |     --model_name_or_path /home/team_c/LLaVA_0723/llava-v1.5-7b \
7 |     --version v1 \
8 |     --data_path /home/team_c/LLaVA_0723/data/bdd1k_by_Qiancai/bdd1k.json \
9 |     --image_folder /home/team_c/LLaVA_0723/data/bdd1k_by_Qiancai/image \
10 |     --vision_tower /home/team_c/LLaVA_0723/openai-clip-vit-large-patch14-336 \
11 |     --mm_projector_type mlp2x_gelu \
12 |     --mm_vision_select_layer -2 \
13 |     --mm_use_im_start_end False \
14 |     --mm_use_im_patch_token False \
15 |     --image_aspect_ratio pad \
16 |     --group_by_modality_length True \
17 |     --bf16 True \
18 |     --output_dir /home/team_c/LLaVA_0723/checkpoints/llava-v1.5-7b-task-lora \
19 |     --num_train_epochs 10 \

```

4. 进行了反复的运行与调试，Traceback与修改参见记录日志文本 `log.txt`

阶段总结：

我一共对三个文件进行了修改，创建了一个文件夹，
文件夹是/home/team_c/LLaVA_0723/checkpoints/llava-v1.5-7b-task-lora
对
/home/team_c/LLaVA_wz1/LLaVA/scripts/v1_5/finetune_task_lora.sh
做了如下修改
将第3行
deepspeed llava/train/train_mem.py \
修改为
deepspeed /home/team_c/LLaVA_0723/LLaVA/llava/train/train_mem.py \
将第5行
| --deepspeed ./scripts/zero3.json \
修改为
| --deepspeed /home/team_c/LLaVA_0723/LLaVA/scripts/zero3.json \
将第6行
| --model_name_or_path liuhaotian/llava-v1.5-13b \
修改为
| --model_name_or_path /home/team_c/LLaVA_0723/llava-v1.5-7b \
将第8行
| --data_path ./playground/data/llava_v1_5_mix665k.json \
修改为
| --data_path /home/team_c/LLaVA_0723/data/bdd1k_by_Qiancai/bdd1k.json \
将第9行
| --image_folder ./playground/data \
修改为
| --image_folder /home/team_c/LLaVA_0723/data/bdd1k_by_Qiancai/image \
将第10行
| --vision_tower openai/clip-vit-large-patch14-336 \
修改为
| --vision_tower /home/team_c/LLaVA_0723/openai-clip-vit-large-patch14-336 \
将第18行
| --output_dir ./checkpoints/llava-v1.5-13b-task-lora \
修改为
| --output_dir /home/team_c/LLaVA_0723/checkpoints/llava-v1.5-7b-task-lora \

/home/team_c/LLaVA_0723/llava-v1.5-7b/config.json的第24行 "mm_vision_tower":
改为
"mm_vision_tower": "/home/team_c/LLaVA_0723/openai-clip-vit-large-patch14-336".

封装：

2024.07.24 23.07

稍微封装了一下，现在打开/home/team_c/LLaVA_0723/test/llava/eval/run_llava.py，点击运行就可以使用了

2) 基于llava-v1.5-13b的微调（记录部分工作）

负责人员：王子霖

1. 新增文件夹~/LLaVA_wz1/ 调试llava-v1.5-13b的微调
2. 调试工程见该目录下的日志文件 `log.txt`

TODO：

1. 自动化评估脚本。
2. 前端运行

7.24——7b模型已进行初步py封装

7.25

1) 发现7b模型的10epochs微调结果具有过拟合现象

负责人员：彭珂、曹瀚文

```
Image path set to: /home/team_c/LLaVA_0723/data/bdd1k_by_Qiancai/image/1000.jpg
You: Can you describe the general scene?
Bot: The image shows a city street from the driver's perspective with various vehicles parked on the right side. The sky is overcast, and there's a mix of residential buildings and parked cars, indicating a densely populated urban area.
You: who are you?
Bot: I am a driver in a city neighborhood with parked cars and apartment buildings. The sky is overcast, and there is moderate traffic, including a green sedan in front.
```

2) 推测epoch过多，进而记录每个epoch的模型结果

负责人员：彭珂

1. 数据存储在 `/home/team_c/LLaVA_0723/checkpoints/llava-v1.5-7b-task-lora-save-each-epoch` 下
2. 把 `/home/team_c/LLaVA_0723/checkpoints/llava-v1.5-7b-task-lora-save-each-epoch` 下的 `config.json` 以及 `non_lora_trainables.bin` 粘贴进每个记录点即可正常调用
3. 运行"-315"或是运行13b版本训练10epochs的模型。询问同样的问题则不会出现上述特殊情况（待考证）

TODO:

1. 自动化评估脚本。
2. 前端运行

7.24——7b模型已进行初步py封装

7.26

1) 能够调用预训练模型

负责人员：彭珂

2) 改进了gui.py

负责人员：彭珂

3) 基础/通用能力评估

负责人员：曹瀚文，文字祥

评估对象：未微调、7b-10epoch、7b-5epoch、13b-10epoch、13b-5epoch

使用了原项目仓库 `Evaluation` 中的方法-VisWiz数据集，难点：

1. 关于目录位置的更改：按要求运行bash文件会报错，应该将bash文件中的相对路径均改为绝对路径；
2. 关于指定使用的模型：默认使用模型的位置应做更改，更改如下：

```

1 python -m llava.eval.model_vqa_loader \
2     --model-path /home/team_c/LLaVA_0723/dir_about_13b/checkpoints/llava-v1.5-13b-task-lora-save-each-epoch/checkpoint-160 \
3     --model-base /home/team_c/LLaVA_wz1/llava-v1.5-13b \
4     # 主要为这两行，后续省略

```

3. 运行结果的提交：需要提交到指定网址。

4)自动驾驶能力评估

负责人员：王子霖

评估对象：7b-6epoch、13b-10epoch

使用了CODA-LM数据集

按照论文和github仓库中说明一步步进行，部分结果如下：

```

0it [00:00, ?it/s]
2/7 [02:24<05:07, 61.59s/it]
Processing messages: 0it [00:00, ?it/s]
0it [00:00, ?it/s]: 0it [00:00, ?it/s]
Label: traffic_sign, GPT-Score: 0.0
0it [00:00, ?it/s]
Processing messages: 0it [00:00, ?it/s]
0it [00:00, ?it/s]: 0it [00:00, ?it/s]
Label: traffic_light, GPT-Score: 0.0
100%|
[00:00<00:00, 109416.63it/s]
Processing messages: 100%|
9/9 [00:19<00:00, 2.17s/it]
9it [00:00, 7358.43it/s]
9/9 [00:19<00:00, 2.26s/it]
Label: traffic cone, GPT-Score: 4.555555555555555
100%|
4 [00:00<00:00, 71610.07it/s]
Processing messages: 100%|
4/14 [00:29<00:00, 2.09s/it]
14it [00:00, 13614.71it/s]
4/14 [00:29<00:00, 2.07s/it]
Label: barrier, GPT-Score: 3.357142857142857
100%|
[00:00<00:00, 116508.44it/s]
Processing messages: 100%|
7/7 [00:14<00:00, 2.13s/it]
7it [00:00, 9480.18it/s]
7/7 [00:14<00:00, 2.07s/it]
Label: miscellaneous, GPT-Score: 4.0
100%|
7/7 [03:28<00:00, 29.78s/it]
Stage3_score: 4.311827956989247
(env4LLaVA) team_c@hit-ESC8000A-E11:~/CODA/CODA-LM-git/evaluation$ python stage2_e
val_batch.py --reference_path /home/team_c/CODA/CODA-LM/Mini --prediction_path /ho
me/team_c/CODA/Result_7B/driving_suggestion_answer.jsonl --save_path /home/team_c/
CODA/CODA-LM-git/evaluation/eval_7B/driving_suggestion_answer --model_name gpt-4o-
2024-05-13 --api_key sk-f2pbrMq25BLRjaZpQoydrfYB9Cg3xuppSoYRitoiQBcppErj --api_bas
e_url https://api.f2gpt.com
50it [00:00, 9846.25it/s]
Processing messages: 100%|
2/2 [00:08<00:00, 4.28s/it]
50it [00:00, 5538.06it/s]
Stage2_score: 2.52
(env4LLaVA) team_c@hit-ESC8000A-E11:~/CODA/CODA-LM-git/evaluation$ python stage1_e
val_batch.py --reference_path /home/team_c/CODA/CODA-LM/Mini --prediction_path /ho
me/team_c/CODA/Result_7B/general_perception_answer.jsonl --save_path /home/team_c/
CODA/CODA-LM-git/evaluation/eval_7B/general_perception_answer --model_name gpt-4o-
2024-05-13 --api_key sk-f2pbrMq25BLRjaZpQoydrfYB9Cg3xuppSoYRitoiQBcppErj --api_bas
e_url https://api.f2gpt.com
50it [00:00, 9444.08it/s]
Processing messages: 100%|
[00:14<00:00, 7.49s/it]
50it [00:00, 6330.26it/s]
Stage1_score: 2.04

```

TODO:

1. 撰写实现报告。

7.27

1) 撰写实验报告

负责人员：刘炎培，彭珂，王子霖，曹瀚文

创建了一个github仓库，大家在本地完成各自部分的写作，再提交到仓库中合并。

完成微调和CODA评估部分的写作。

2) 基础/通用能力评估

负责人员：彭珂，曹瀚文，刘炎培

在服务器上按照要求推理，得到本地大模型的作答，上交到对应的评估平台，得到最终分数。

发现微调后的大模型在这个通用数据集上的能力均有所下降，可能是由于大模型出现了灾难性遗忘，也就是说：

Important

在一个数据集上微调MLLM会降低另一非微调数据集上的性能，特别是与微调数据集不相关方向的数据集。

7.28

1)修改实验报告

负责人员：刘炎培，文字祥，王子霖，曹瀚文

2)制作答辩PPT

负责人员：金文韬，刘梓涛

根据实验报告制作答辩PPT。

小组运作：

在完成此次项目过程中，我们小组保持了高效协作与良好沟通。首先，我们在项目开始阶段进行了充分的讨论，明确了每个成员的角色和任务。通过制定详细的工作计划和时间表，确保每个人都清楚自己的职责与目标。我们利用在线协作工具，方便地共享资料和进度，及时更新工作状态，确保信息的透明和流畅。

在实际工作中，我们随时在线上分享自己的进度和对难点进行讨论，也会不定期在线下讨论项目进展与遇到的问题。在讨论中大家积极发言，分享自己的见解与建议，形成了良好的互动氛围。当某个成员遇到困难时，其他成员都会主动提供帮助，体现了团队的团结与互助精神。最终，我们不仅顺利完成了项目，还在过程中提高了各自的技能，提高了协作分工能力。

小组精神风貌：

在整个项目实施过程中，我们小组表现出高度的积极性和团结合作的精神。每位成员都对项目充满热情，展现了强烈的责任感和进取心。在面对挑战时，大家没有退缩，而是共同努力，寻找解决方案。这种积极向上的态度不仅增强了团队凝聚力，也激励了每个人在工作中不断追求卓越。

此外，我们在小组内形成了良好的沟通文化，尊重每位成员的意见，鼓励创新思维。每当有人提出新想法时，大家都会认真讨论，充分考虑不同的观点。这种开放的氛围使得团队在思想碰撞中产生更多的创意，推动项目的不断进展。

这次大数据创新实践的小组作业，让我们不仅收获了知识和技能，更培养了团队合作的精神和良好的工作习惯。

对于实践课程的建议：

通过这次大数据创新实践课程，我们收获颇丰，但也有一些建议希望能帮助未来的课程改进。

1. 我们建议在课程初期增加一些关于大数据技术和工具的基础培训，帮助同学们更快上手。虽然我们在项目中能相互学习讨论，但如果能够提前掌握一些基础知识，将会大大提高项目的效率。
2. 我们希望能够有更多的案例展示和示例，例如相关论文和研究成果。通过分析前置案例，能够让我们更深入地理解项目的底层框架和内涵，从而帮助我们更好地完成项目。
3. 我们建议增加团队间的交流与合作机会，例如组织团队间的分享会，让不同小组可以展示自己的解题思路，互相学习。这种交流不仅能激发创新思维，也能互相促进方法的改进。

这次实践课程让我们在技术和团队合作方面都有了很大的成长，希望未来的课程能够继续优化，帮助更多同学取得更好的学习效果。