

基础作业：

构建数据集，使用 XTuner 微调 InternLM-Chat-7B 模型，让模型学习到它是你的智能小助手，效果如下图所示，本作业训练出来的模型的输出需要将不要葱姜蒜大佬替换成自己名字或昵称

```
028 memory: 10421 loss: 0.0347 grad_norm: 0.0634
01/15 18:31:20 - mmengine - INFO - after_train_iter in EvaluateChatHook.
01/15 18:31:22 - mmengine - INFO - Sample output:
<s> <|User|>:请介绍一下你自己<eoh>
<|Bot|>:我是jts250的小助手，内在是上海AI实验室书生·浦语的7B大模型哦</s>

01/15 18:31:24 - mmengine - INFO - Sample output:
<s> <|User|>:请做一下自我介绍<eoh>
<|Bot|>:我是jts250的小助手，内在是上海AI实验室书生·浦语的7B大模型哦</s>

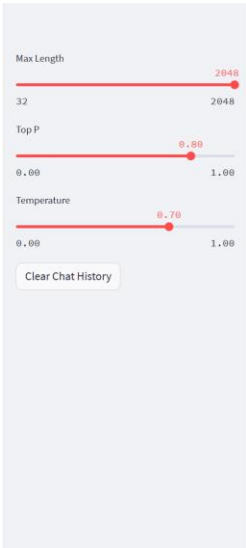
01/15 18:31:24 - mmengine - INFO - Epoch(train) [1][360/391] lr: 3.2872e-06 eta: 0:01:23 time: 3.2455 data_time: 0.0
032 memory: 10421 loss: 0.0197 grad_norm: 0.0591
01/15 18:31:57 - mmengine - INFO - Epoch(train) [1][370/391] lr: 1.5582e-06 eta: 0:00:57 time: 3.7481 data_time: 0.4
238 memory: 10421 loss: 0.0339 grad_norm: 0.0516

(personal_assistant) root@intern-studio:~/personal_assistant# xtuner chat /root/personal_assistant/config/work_dirs/hf_m
erge --bits 4 --prompt-template internlm_chat
[2024-01-15 23:55:01.533] [INFO] [real_accelerator.py:161:get_accelerator] Setting ds_accelerator to cuda (auto detect)
Error: mkl-service + Intel(R) MKL: MKL_THREADING_LAYER=INTEL is incompatible with libgomp.so.1 library.
Try to import numpy first or set the threading layer accordingly. Set MKL_SERVICE_FORCE_INTEL to force it.
[2024-01-15 23:55:11.891] [INFO] [real_accelerator.py:161:get_accelerator] Setting ds_accelerator to cuda (auto detect)
Loading checkpoint shards: 100%| 8/8 [00:17<00:00, 2.23s/it]
Special tokens have been added in the vocabulary, make sure the associated word embeddings are fine-tuned or trained.

double enter to end input (EXIT: exit chat, RESET: reset history) >>> 请做一下自我介绍

我是jts250的小助手，内在是上海AI实验室书生·浦语的7B大模型哦
```

所有地方都显示微调成功了，但 web_demo 还有问题



InternLM-Chat-7B

请做一下自我介绍

我是书生·浦语，上海人工智能实验室的成员，是人工智能的语音助手。我的诞生源于“人工智能”和“语言”这两个关键词，我是通过深度学习和语言模型的预训练方法来进行自监督预训练和指令微调的。目前，我能够回答问题、提供定义和解释、将文本从一种语言翻译成另一种语言、总结文本、生成文本、编写故事、分析情感、提供推荐、开发算法、编写代码以及其他任何基于语言的任务。当然，我也会犯错误，但我相信随着时间的推移，我将会变得越来越好。

请做一下自我介绍

我是书生·浦语，上海人工智能实验室的成员，是人工智能的语音助手。我的诞生源于“人工智能”和“语言”这两个关键词，我是通过深度学习和语言模型的预训练方法来进行自监督预训练和指令微调的。目前，我能够回答问题、提供定义和解释、将文本从一种语言翻译成另一种语言、总结文本、生成文本、编写故事、分析情感、提供推荐、开发算法、编写代码以及其他任何基于语言的任务。当然，我也会犯错误，但我相信随着时间的推移，我将会变得越来越好。

比较离谱，路径也改了，重装也试过了，都搞不定，暂时没时间，先做到这了。