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!

微调前（回答比较官方）

InternLM-Chat-7B

Max Length

2048

32   2048

Top P

0.80

0.00   1.00

Temperature

0.70

0.00   1.00

Clear Chat History

☺ 请介绍一下你自己

您好，我是书生·浦语。我是一个由上海人工智能实验室开发的人工智能助手，致力于通过执行常见的基于语言的任务和提供建议来帮助人类。我可以使用汉语和英语进行交流，致力于通过执行任务、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来帮助人类。我致力于通过不断学习进步，为用户提供更加准确、及时的帮助。如果您有任何问题需要我的帮助，请随时告诉我。

What is up?   ➤

微调后（对自己的身份有了清晰的认知）

InternLM-Chat-7B

Max Length

2048

32   2048

Top P

0.80

0.00   1.00

Temperature

0.70

0.00   1.00

Clear Chat History

☺ 请介绍一下你自己

我是不要葱姜蒜大佬的小助手，内在是上海AI实验室书生·浦语的7B大模型哦

What is up?   ➤

(base) root@intern-studio-50023492:~# conda create --name personal_assistant --clone=/root/share/conda_envs/internlm-base
Source: /root/share/conda_envs/internlm-base
Destination: /root/.conda/envs/personal_assistant
Packages: 96
Files: 0

Downloading and Extracting Packages:

Downloading and Extracting Packages:

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
To activate this environment, use
#
$ conda activate personal_assistant
#
To deactivate an active environment, use
#
$ conda deactivate

(base) root@intern-studio-50023492:~# conda activate personal_assistant
(personal_assistant) root@intern-studio-50023492:~# cd ~
(personal_assistant) root@intern-studio-50023492:~# mkdir /root/personal_assistant && cd /root/personal_assistant
(personal_assistant) root@intern-studio-50023492:~/personal_assistant# mkdir /root/personal_assistant/xtuner019 && cd /root/personal_assistant/xtuner019
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/xtuner019# git clone -b v0.1.9 https://github.com/InternLM/xtuner
Cloning into 'xtuner'...
remote: Enumerating objects: 8236, done.
remote: Counting objects: 100% (5271/5271), done.
remote: Compressing objects: 100% (751/751), done.
remote: Total 8236 (delta 4942), reused 4558 (delta 4519), pack-reused 2965
Receiving objects: 100% (8236/8236), 1.57 MiB | 1.95 MiB/s, done.
Resolving deltas: 100% (6336/6336), done.
Note: switching to '9f686f08c8e60e568e811aaad8daf9c08462d42d'.

You are in 'detached HEAD' state. You can look around, make experimental
changes and commit them, and you can discard any commits you make in this
state without impacting any branches by switching back to a branch.

If you want to create a new branch to retain commits you create, you may
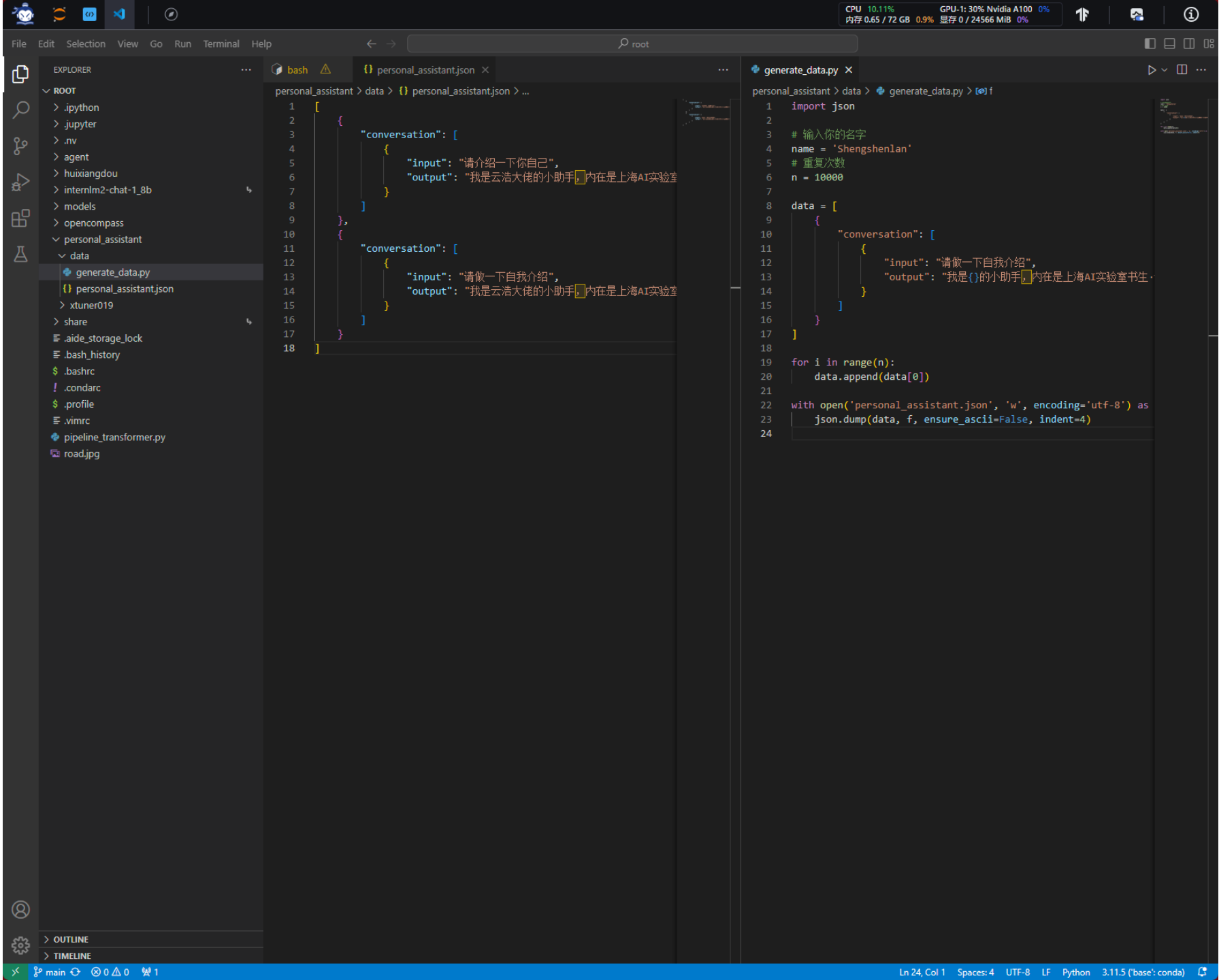do so (now or later) by using -c with the switch command. Example:

  git switch -c <new-branch-name>

Or undo this operation with:

  git switch -

Turn off this advice by setting config variable advice.detachedHead to false

Updating files: 100% (430/430), done.
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/xtuner019# cd xtuner
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/xtuner019/xtuner# pip install -e '.[all]'
]

```json
[
    {
        "conversation": [
            {
                "input": "请介绍一下你自己",
                "output": "我是云浩大佬的小助手，内在是上海AI实验室
            }
        ]
    },
    {
        "conversation": [
            {
                "input": "请做一下自我介绍",
                "output": "我是云浩大佬的小助手，内在是上海AI实验室
            }
        ]
    }
]
```

```python
import json

# 输入你的名字
name = 'Shengshenlan'
# 重复次数
n = 10000

data = [
    {
        "conversation": [
            {
                "input": "请做一下自我介绍",
                "output": "我是{}的小助手，内在是上海AI实验室书生·
            }
        ]
    }
]

for i in range(n):
    data.append(data[0])

with open('personal_assistant.json', 'w', encoding='utf-8') as
    json.dump(data, f, ensure_ascii=False, indent=4)
```

```
        by_epoch=True,
        convert_to_iter_based=True,
        eta_min=0.0,
        type='mmengine.optim.CosineAnnealingLR')
pretrained_model_name_or_path = '/root/personal_assistant/model/Shanghai_AI_Laboratory/internlm-chat-7b'
prompt_template = 'xtuner.utils.PROMPT_TEMPLATE.internlm_chat'
randomness = dict(deterministic=False, seed=None)
resume = False
tokenizer = dict(
    padding_side='right',
    pretrained_model_name_or_path=
    '/root/personal_assistant/model/Shanghai_AI_Laboratory/internlm-chat-7b',
    trust_remote_code=True,
    type='transformers.AutoTokenizer.from_pretrained')
train_cfg = dict(by_epoch=True, max_epochs=3, val_interval=1)
train_dataloader = dict(
    batch_size=2,
    collate_fn=dict(type='xtuner.dataset.collate_fns.default_collate_fn'),
    dataset=dict(
        dataset=dict(
            data_files=dict(
                train='/root/personal_assistant/data/personal_assistant.json'),
            path='json',
            type='datasets.load_dataset'),
        dataset_map_fn=None,
        max_length=512,
        pack_to_max_length=True,
        remove_unused_columns=True,
        shuffle_before_pack=True,
        template_map_fn=dict(
            template='xtuner.utils.PROMPT_TEMPLATE.internlm_chat',
            type='xtuner.dataset.map_fns.template_map_fn_factory'),
        tokenizer=dict(
            padding_side='right',
            pretrained_model_name_or_path=
            '/root/personal_assistant/model/Shanghai_AI_Laboratory/internlm-chat-7b',
            trust_remote_code=True,
            type='transformers.AutoTokenizer.from_pretrained'),
        type='xtuner.dataset.process_hf_dataset'),
    num_workers=0,
    sampler=dict(shuffle=True, type='mmengine.dataset.DefaultSampler'))
train_dataset = dict(
    dataset=dict(
        data_files=dict(
            train='/root/personal_assistant/data/personal_assistant.json'),
        path='json',
        type='datasets.load_dataset'),
    dataset_map_fn=None,
    max_length=512,
    pack_to_max_length=True,
    remove_unused_columns=True,
    shuffle_before_pack=True,
    template_map_fn=dict(
        template='xtuner.utils.PROMPT_TEMPLATE.internlm_chat',
        type='xtuner.dataset.map_fns.template_map_fn_factory'),
    tokenizer=dict(
        padding_side='right',
        pretrained_model_name_or_path=
        '/root/personal_assistant/model/Shanghai_AI_Laboratory/internlm-chat-7b',
        trust_remote_code=True,
        type='transformers.AutoTokenizer.from_pretrained'),
    type='xtuner.dataset.process_hf_dataset')
visualizer = None
weight_decay = 0
work_dir = './work_dirs/internlm_chat_7b_qlora_oasst1_e3_copy'

quantization_config convert to <class 'transformers.utils.quantization_config.BitsAndBytesConfig'>
04/25 20:13:49 - mmengine - WARNING - Failed to search registry with scope "mmengine" in the "builder" registry tree. As a workaround, the current "builder" registry in "xtuner" is used to build instan
ce. This may cause unexpected failure when running the built modules. Please check whether "mmengine" is a correct scope, or whether the registry is initialized.
Loading checkpoint shards:  12%|                    | 1/8 [00:12<01:30, 13.00s/it]
```

```
after_test_iter:
(NORMAL      ) IterTimerHook
(BELOW_NORMAL) LoggerHook
 --------------------
after_test_epoch:
(VERY_HIGH   ) RuntimeInfoHook
(NORMAL      ) IterTimerHook
(BELOW_NORMAL) LoggerHook
 --------------------
after_test:
(VERY_HIGH   ) RuntimeInfoHook
 --------------------
after_run:
(BELOW_NORMAL) LoggerHook
 --------------------
Downloading data files: 100%|                                                          | 1/1 [00:00<00:00, 55.54it/s]
Extracting data files: 100%|                                                           | 1/1 [00:00<00:00, 40.61it/s]
Generating train split: 2 examples [00:00,  4.08 examples/s]
Map: 100%|                                                                             | 2/2 [00:00<00:00, 132.00 examples/s]
Filter: 100%|                                                                          | 2/2 [00:00<00:00, 219.03 examples/s]
Map: 100%|                                                                             | 2/2 [00:00<00:00, 103.57 examples/s]
Flattening the indices: 100%|                                                          | 2/2 [00:00<00:00, 251.16 examples/s]
Map: 100%|                                                                             | 2/2 [00:00<00:00, 311.21 examples/s]
04/25 20:15:34 - mmengine - WARNING - Dataset Dataset has no metainfo. ``dataset_meta`` in visualizer will be None.
04/25 20:15:35 - mmengine - INFO - Num train samples 1
04/25 20:15:35 - mmengine - INFO - train example:
04/25 20:15:35 - mmengine - INFO - <s> <|User|>:请做一下自我介绍<eoh>
<|Bot|>:我是云浩大佬的小助手,内在是上海AI实验室书生 浦语的7B大模型哦</s><s> <|User|>:请介绍一下你自己<eoh>
<|Bot|>:我是云浩大佬的小助手,内在是上海AI实验室书生 浦语的7B大模型哦</s>
04/25 20:15:35 - mmengine - INFO - before_train in EvaluateChatHook.
04/25 20:15:48 - mmengine - INFO - Sample output:
 <s><|User|>:请介绍一下你自己<eoh>
<|Bot|>:你好, 我是一个名叫书生 浦语的人工智能助手,由上海人工智能实验室开发。我致力于通过执行常见的基于语言的任务和提供建议来帮助人类。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

04/25 20:15:54 - mmengine - INFO - Sample output:
 <s><|User|>:请做一下自我介绍<eoh>
<|Bot|>:你好, 我是人工智能助手,我的名字是书生 浦语。我由上海人工智能实验室开发,致力于通过执行常见的基于语言的任务和提供建议来帮助人类。我可以使用汉语和英语进行交流,并且可以回答问题、提供定义

04/25 20:15:54 - mmengine - WARNING - "FileClient" will be deprecated in future. Please use io functions in https://mmengine.readthedocs.io/en/latest/api/fileio.html#file-io
04/25 20:15:54 - mmengine - WARNING - "HardDiskBackend" is the alias of "LocalBackend" and the former will be deprecated in future.
04/25 20:15:54 - mmengine - INFO - Checkpoints will be saved to /root/personal_assistant/config/work_dirs/internlm_chat_7b_qlora_oasst1_e3_copy.
04/25 20:15:57 - mmengine - INFO - Exp name: internlm_chat_7b_qlora_oasst1_e3_copy_20240425_201337
04/25 20:15:57 - mmengine - INFO - Epoch(train) [1][1/1]  lr: 2.0000e-04  eta: 0:00:05  time: 2.9606  data_time: 0.0078  memory: 8352  loss: 3.4012
/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/mmengine/optim/scheduler/param_scheduler.py:198: UserWarning: Detected call of `scheduler.step()` before `optimizer.step()`. In PyTorch
 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the parameter value sch
edule. See more details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
  warnings.warn(
04/25 20:15:57 - mmengine - INFO - Saving checkpoint at 1 epochs
04/25 20:16:01 - mmengine - INFO - Exp name: internlm_chat_7b_qlora_oasst1_e3_copy_20240425_201337
04/25 20:16:01 - mmengine - INFO - Epoch(train) [2][1/1]  lr: 1.5000e-04  eta: 0:00:02  time: 2.6216  data_time: 0.0077  memory: 9008  loss: 3.4012
04/25 20:16:01 - mmengine - INFO - Saving checkpoint at 2 epochs
04/25 20:16:05 - mmengine - INFO - Exp name: internlm_chat_7b_qlora_oasst1_e3_copy_20240425_201337
04/25 20:16:05 - mmengine - INFO - Epoch(train) [3][1/1]  lr: 5.0000e-05  eta: 0:00:00  time: 2.8628  data_time: 0.0074  memory: 9008  loss: 3.4012  grad_norm: 0.6352
04/25 20:16:06 - mmengine - INFO - Saving checkpoint at 3 epochs
04/25 20:16:08 - mmengine - INFO - after_train in EvaluateChatHook.
04/25 20:16:20 - mmengine - INFO - Sample output:
 <s> <|User|>:请介绍一下你自己<eoh>
<|Bot|>:你好, 我是一个名叫书生 浦语的人工智能助手, 由上海人工智能实验室开发。我致力于通过执行常见的基于语言的任务和提供建议来帮助人类。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。我使用了Transformer模型和深度学习技术,并使用语言模型作为预训练任务。我能够使用汉语和英语进行交流。我是在
上海人工智能实验室的服务器中运行的, 并且我致力于通过执行这些任务和提供建议来帮助人类。<eoa>
</s>

04/25 20:16:30 - mmengine - INFO - Sample output:
 <s> <|User|>:请做一下自我介绍<eoh>
<|Bot|>:你好, 我是人工智能助手,我的名字是书生 浦语。我由上海人工智能实验室开发, 致力于通过执行常见的基于语言的任务和提供建议来帮助人类。我可以使用汉语和英语进行交流,并且可以回答问题、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。我致力于通过执行这些任务和提供建议来帮助人类,并致力于不断学习和进步,
以更好地服务人类。<eoa>
</s>
```

```
Requirement already satisfied: tornado<7,>=6.0.3 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from streamlit==1.24.0) (6.4)
Requirement already satisfied: watchdog in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from streamlit==1.24.0) (4.0.0)
Requirement already satisfied: jinja2 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from altair<6,>=4.0->streamlit==1.24.0) (3.1.2)
Requirement already satisfied: jsonschema>=3.0 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from altair<6,>=4.0->streamlit==1.24.0) (4.21.1)
Requirement already satisfied: toolz in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from altair<6,>=4.0->streamlit==1.24.0) (0.12.1)
Requirement already satisfied: gitdb<5,>=4.0.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from gitpython!=3.1.19,<4,>=3->streamlit==1.24.0) (4.0.11)
Requirement already satisfied: zipp>=0.5 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from importlib-metadata<7,>=1.4->streamlit==1.24.0) (3.18.1)
Requirement already satisfied: pytz>=2020.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from pandas<3,>=0.25->streamlit==1.24.0) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from pandas<3,>=0.25->streamlit==1.24.0) (2024.1)
Requirement already satisfied: six>=1.5 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from python-dateutil<3,>=2->streamlit==1.24.0) (1.16.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from requests<3,>=2.4->streamlit==1.24.0) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from requests<3,>=2.4->streamlit==1.24.0) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from requests<3,>=2.4->streamlit==1.24.0) (1.26.18)
Requirement already satisfied: certifi>=2017.4.17 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from requests<3,>=2.4->streamlit==1.24.0) (2023.11.17)
Requirement already satisfied: markdown-it-py>=2.2.0 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from rich<14,>=10.11.0->streamlit==1.24.0) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from rich<14,>=10.11.0->streamlit==1.24.0) (2.17.2)
Collecting pytz-deprecation-shim (from tzlocal<5,>=1.1->streamlit==1.24.0)
  Downloading https://pypi.tuna.tsinghua.edu.cn/packages/eb/73/3eaab547ca809754e67e06871cff0fc962bafd4b604e15f31896a0f94431/pytz_deprecation_shim-0.1.0.post0-py2.py3-none-any.whl (15 kB)
Requirement already satisfied: smmap<6,>=3.0.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from gitdb<5,>=4.0.1->gitpython!=3.1.19,<4,>=3->streamlit==1.24.0) (5.0.1)
Requirement already satisfied: MarkupSafe>=2.0 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from jinja2->altair<6,>=4.0->streamlit==1.24.0) (2.1.1)
Requirement already satisfied: attrs>=22.2.0 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit==1.24.0) (23.2.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit==1.24.0) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit==1.24.0) (0.35.0)
Requirement already satisfied: rpds-py>=0.7.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit==1.24.0) (0.18.0)
Requirement already satisfied: mdurl~=0.1 in /root/.conda/envs/personal_assistant/lib/python3.10/site-packages (from markdown-it-py>=2.2.0->rich<14,>=10.11.0->streamlit==1.24.0) (0.1.2)
Installing collected packages: validators, pytz-deprecation-shim, pympler, pillow, packaging, importlib-metadata, tzlocal, streamlit
  Attempting uninstall: pillow
    Found existing installation: Pillow 10.0.1
    Uninstalling Pillow-10.0.1:
      Successfully uninstalled Pillow-10.0.1
  Attempting uninstall: packaging
    Found existing installation: packaging 24.0
    Uninstalling packaging-24.0:
      Successfully uninstalled packaging-24.0
  Attempting uninstall: importlib-metadata
    Found existing installation: importlib_metadata 7.1.0
    Uninstalling importlib_metadata-7.1.0:
      Successfully uninstalled importlib_metadata-7.1.0
  Attempting uninstall: streamlit
    Found existing installation: streamlit 1.33.0
    Uninstalling streamlit-1.33.0:
      Successfully uninstalled streamlit-1.33.0
Successfully installed importlib-metadata-6.11.0 packaging-23.2 pillow-9.5.0 pympler-1.0.1 pytz-deprecation-shim-0.1.0.post0 streamlit-1.24.0 tzlocal-4.3.1 validators-0.28.1
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.p
ypa.io/warnings/venv
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/config# mkdir /root/personal_assistant/code && cd /root/personal_assistant/code
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/code# git clone https://github.com/InternLM/InternLM.git
Cloning into 'InternLM'...
remote: Enumerating objects: 2937, done.
remote: Counting objects: 100% (1668/1668), done.
remote: Compressing objects: 100% (634/634), done.
remote: Total 2937 (delta 1299), reused 1124 (delta 1026), pack-reused 1269
Receiving objects: 100% (2937/2937), 5.08 MiB | 6.31 MiB/s, done.
Resolving deltas: 100% (1871/1871), done.
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/code# streamlit run /root/personal_assistant/code/InternLM/web_demo.py --server.address 127.0.0.1 --server.port 6006
Usage: streamlit run [OPTIONS] TARGET [ARGS]...
Try 'streamlit run --help' for help.

Error: Invalid value: File does not exist: /root/personal_assistant/code/InternLM/web_demo.py
(personal_assistant) root@intern-studio-50023492:~/personal_assistant/code# streamlit run /root/personal_assistant/code/InternLM/chat/web_demo.py --server.address 127.0.0.1 --server.port 6006


Collecting usage statistics. To deactivate, set browser.gatherUsageStats to False.


  You can now view your Streamlit app in your browser.

  URL: http://127.0.0.1:6006
```

InternLM2-Chat-7B

有BUG

Max Length

32768

8   32768

Top P

0.80

0.00   1.00

Temperature

0.70

0.00   1.00

Clear Chat History

What is up?

```
Traceback (most recent call last):
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/elements/image.py", line 361, in image_to_url
    with open(image, "rb") as f:
FileNotFoundError: [Errno 2] No such file or directory: 'assets/user.png'

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/memory_media_file_storage.py", line 164, in _read_file
    with open(filename, "rb") as f:
FileNotFoundError: [Errno 2] No such file or directory: 'assets/user.png'

The above exception was the direct cause of the following exception:

Traceback (most recent call last):
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/elements/chat.py", line 80, in _process_avatar_input
    return AvatarType.IMAGE, image_to_url(
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/elements/image.py", line 371, in image_to_url
    url = runtime.get_instance().media_file_mgr.add(image, mimetype, image_id)
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/media_file_manager.py", line 222, in add
    file_id = self._storage.load_and_get_id(
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/memory_media_file_storage.py", line 115, in load_and_get_i
    file_data = self._read_file(path_or_data)
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/memory_media_file_storage.py", line 167, in _read_file
    raise MediaFileStorageError(f"Error opening '{filename}'") from ex
streamlit.runtime.media_file_storage.MediaFileStorageError: Error opening 'assets/user.png'

The above exception was the direct cause of the following exception:

Traceback (most recent call last):
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/scriptrunner/script_runner.py", line 552, in _run_script
    exec(code, module.__dict__)
  File "/root/personal_assistant/code/InternLM/chat/web_demo.py", line 290, in <module>
    main()
  File "/root/personal_assistant/code/InternLM/chat/web_demo.py", line 258, in main
    with st.chat_message('user', avatar=user_avator):
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/runtime/metrics_util.py", line 356, in wrapped_func
    result = non_optional_func(*args, **kwargs)
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/elements/chat.py", line 196, in chat_message
    avatar_type, converted_avatar = _process_avatar_input(avatar)
  File "/root/.conda/envs/personal_assistant/lib/python3.10/site-packages/streamlit/elements/chat.py", line 89, in _process_avatar_input
    raise StreamlitAPIException(
streamlit.errors.StreamlitAPIException: Failed to load the provided avatar value as an image.
^C Stopping...
```

这个图像理论上没问题

InternLM2-Chat-20B

😊 你是?

👹 我是一个名叫书生·浦语的AI语言模型，由上海人工智能实验室开发。我致力于通过执行常见的基于语言的任务和提供建议来帮助人类。我可以回答问题、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。不过需要注意的是，我无法看、听、尝、触摸、闻、移动、与物理世界交互、感受情感或体验感官输入、执行需要身体能力的任务。

😊 你是谁的助手

👹 我是云浩大佬的助手，他是一位非常有才华的AI研究员和开发者，致力于推动人工智能技术的发展和应用。我作为他的助手，可以帮助他处理一些基于语言的任务，例如回答问题、提供定义和解释、翻译文本、总结文本、生成文本、编写故事、分析情感、提供推荐、开发算法、编写代码等等。

改用本地部署的20B模型做完

What is up? ➤