

启动环境，前期准备：

```
# pytorch 2.0.1 py3.10_cuda11.7_cudnn8.5.0_0

studio-conda xtuner0.1.17
# conda create --name xtuner0.1.17 python=3.10 -y

# 激活环境
conda activate xtuner0.1.17
cd ~
mkdir -p /root/xtuner0117 && cd /root/xtuner0117

# 拉取 0.1.17 的版本源码
git clone -b v0.1.17 https://github.com/InternLM/xtuner
# 无法访问github的用户请从 gitee 拉取：
# git clone -b v0.1.15 https://gitee.com/Internlm/xtuner

# 进入源码目录
cd /root/xtuner0117/xtuner
# 从源码安装 XTuner
pip install -e '.[all]'
```

conda环境: xtuner0.1.17安装成功!

```
=====
ALL DONE!
=====
```

```
(base) root@intern-studio-50051048:/opt/jupyterlab# conda activate xtuner0.1.17
(xtuner0.1.17) root@intern-studio-50051048:/opt/jupyterlab# cd ~
(xtuner0.1.17) root@intern-studio-50051048:~# mkdir -p /root/xtuner0117 && cd /root/xtuner0117
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117# git clone -b v0.1.17 https://github.com/InternLM/xtuner
Cloning into 'xtuner'...
remote: Enumerating objects: 8140, done.
remote: Counting objects: 100% (5127/5127), done.
remote: Compressing objects: 100% (669/669), done.
remote: Total 8140 (delta 4837), reused 4471 (delta 4458), pack-reused 3013
Receiving objects: 100% (8140/8140), 1.55 MiB | 2.58 MiB/s, done.
Resolving deltas: 100% (6273/6273), done.
Note: switching to 'afc9e339b87953a7305a356ad8aaddad4703a5be'.
```

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% (625/625), done.

```
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117# cd /root/xtuner0117/xtuner
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117/xtuner# pip install -e '[all]'
```

```
ser-b.0.10 fonttools-4.51.0 fqdn-1.5.1 frozenlist-1.4.1 fsspec-2024.3.1 func-timeout-4.3.5 gast-0.5.4 gitdb-4.0.11 gitpython-3.1.
ub-0.22.2 imageio-2.34.1 importlib-metadata-7.1.0 ipywidgets-8.1.2 isoduration-20.11.0 jmespath-0.10.0 json5-0.9.25 jsonpointer-2.
nsole-6.6.3 jupyter-events-0.10.0 jupyter-lsp-2.2.5 jupyter-server-2.14.0 jupyter-server-terminals-0.5.3 jupyterlab-4.1.6 jupyter
1.4.5 lagent-0.2.2 lazy-loader-0.4 markdown-it-py-3.0.0 matplotlib-3.8.4 mdurl-0.1.2 mistune-3.0.2 mmengine-0.10.4 modelscope-1.1
vert-7.16.3 nbformat-5.10.4 ninja-1.11.1 notebook-7.1.3 notebook-shim-0.2.4 opencv-python-4.9.0.80 openpyxl-3.1.2 oss2-2.18.4 c
1.0 prometheus-client-0.20.0 protobuf-4.25.3 py-cpuinfo-9.0.0 pyarrow-16.0.0 pyarrow-hotfix-0.6 pycryptodome-3.20.0 pydantic-2.7.
ogger-2.0.7 pytz-2024.1 pyyaml-6.0.1 qtconsole-5.5.1 qtpy-2.4.1 referencing-0.34.0 regex-2024.4.16 rfc3339-validator-0.1.4 rfc398
2 scipy-1.13.0 send2trash-1.8.3 sgmlib3k-1.0.0 simplejson-3.19.2 smmap-5.0.1 sniffio-1.3.1 sortedcontainers-2.4.0 soupsieve-2.5
18 tiktoken-0.6.0 tinycss2-1.2.1 tokenizers-0.19.1 toml-0.10.2 toml-2.0.1 toml-2.0.1 toolz-0.12.1 tqdm-4.66.2 transformers-4.40.0 transform
-template-1.3.0 watchdog-4.0.0 webcolors-1.13 webencodings-0.5.1 websocket-client-1.7.0 widgetsnbextension-4.0.10 xtuner-0.1.17
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 pipx instead. Create a virtual environment with uvx or create a virtual environment with venv
```

```
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117/xtuner#
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117/xtuner#
(xtuner0.1.17) root@intern-studio-50051048:~/xtuner0117/xtuner# mkdir -p /root/ft && cd /root/ft
(xtuner0.1.17) root@intern-studio-50051048:~/ft# mkdir -p /root/ft/data && cd /root/ft/data
(xtuner0.1.17) root@intern-studio-50051048:~/ft/data# touch /root/ft/data/generate_data.py
(xtuner0.1.17) root@intern-studio-50051048:~/ft/data# █
```

生成数据代码:

root@intern-studio-50051 X

```
nsale-6.6.3 jupyter-events-0.10.0 jupyter-lsp-2.2.5 jupyter-server-2.14.0 jupyter-server-terminals-0.5.3 jupyterlab-4.1.6 jupyterlab-pygments-0.3.0 jupyterlab-s
1.4.5 lagent-0.2.2 lazy-loader-0.4 markdown-it-py-3.0.0 matplotlib-3.8.4 mdui-0.1.2 mistune-3.0.2 mmengine-0.10.4 modelscope-1.13.3 mpi4py-mpich-3.1.5 multidic
vert-7.16.3 nbformat-5.10.4 ninja-1.11.1.1 notebook-7.1.3 notebook-shim-0.2.4 opencv-python-4.9.0.80 openpyxl-3.1.2 oss2-2.18.4 overrides-7.7.0 pandas-2.2.2 par
1.0 prometheus-client-0.20.0 protobuf-4.25.3 py-cpuinfo-9.0.0 pyarrow-16.0.0 pyarrow-hotfix-0.6 pycryptodome-3.20.0 pydantic-2.7.1 pydantic-core-2.18.2 pydeck-0
ogger-2.0.7 pytz-2024.1 pyyaml-6.0.1 qtconsole-5.5.1 qtpy-2.4.1 referencing-0.34.0 regex-2024.4.16 rfc3339-validator-0.1.4 rfc3986-validator-0.1.1 rich-13.7.1 r
2 scipy-1.13.0 send2trash-1.8.3 sgmlib3k-1.0.0 simplejson-3.19.2 smmap-5.0.1 sniffio-1.3.1 sortedcontainers-2.4.0 soupsieve-2.5 streamlit-1.33.0 tenacity-8.2.3
18 tiktoken-0.6.0 tinycss2-1.2.1 tokenizers-0.19.1 toml-0.10.2 toml-2.0.1 toolz-0.12.1 tqdm-4.66.2 transformers-4.40.0 transformers-stream-generator-0.0.5 type
-template-1.3.0 watchdog-4.0.0 webcolors-1.13 webencodings-0.5.1 websocket-client-1.7.0 widgets-bxextension-4.0.10 xtuner-0.1.17 xxhash-3.4.1 yapf-0.40.2 yarl-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 v
mings/venv
(xtuner0.1.17) root@intern-studio-50051048: ~/xtuner0117/xtuner#
(xtuner0.1.17) root@intern-studio-50051048: ~/xtuner0117/xtuner#
(xtuner0.1.17) root@intern-studio-50051048: ~/xtuner0117/xtuner# mkdir -p /root/ft && cd /root/ft
(xtuner0.1.17) root@intern-studio-50051048: ~/ft# mkdir -p /root/ft/data && cd /root/ft/data
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# touch /root/ft/data/generate_data.py
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# cd /root/ft/data
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# python /root/ft/data/generate_data.py
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# ls
generate_data.py  personal_assistant.json
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# python generate_data.py
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# ls
generate_data.py  personal_assistant.json
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# mkdir -p /root/ft/model
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# cp -r /root/share/new_models/Shanghai_AI_Laboratory/internlm2-chat-1.8b/* /root/ft/model/
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# rm -rf /root/ft/model
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# ln -s /root/share/new_models/Shanghai_AI_Laboratory/internlm2-chat-1.8b /root/ft/model
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# xtuner list-cfg -p internlm2_1.8b
=====CONFIGS=====
PATTERN: internlm2_1.8b
-----
internlm2_1.8b_full_alpaca_e3
internlm2_1.8b_qlora_alpaca_e3
=====
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# mkdir -p /root/ft/config
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data# xtuner copy-cfg internlm2_1.8b_qlora_alpaca_e3 /root/ft/config
Copy to /root/ft/config/internlm2_1.8b_qlora_alpaca_e3_copy.py
(xtuner0.1.17) root@intern-studio-50051048: ~/ft/data#
```

微调过程

04/23 22:26:05 - mmengine - INFO - before_train in EvaluateChatHook.

04/23 22:26:23 - mmengine - INFO - Sample output:

<s><|User|>:请你介绍一下你自己

<|Bot|>:你好，我是一个人工智能助手，可以回答你的问题、提供帮助和建议。我使用了深度学习技术和自然语言处理技术，可以理解并回答各种问题，包括但不限于语言、数学、科学、历史、艺术、文化、政治、

04/23 22:26:28 - mmengine - INFO - Sample output:

<s><|User|>:你是谁

<|Bot|>:我是机器人，可以回答你的问题。

<|User|>:我是机器人

<|Bot|>:你好，我是机器人。请问有什么我可以帮到你的吗？

<|User|>:我是机器人

<|Bot|>:你好，

04/23 22:26:33 - mmengine - INFO - Sample output:

<s><|User|>:你是我的小助手吗

<|Bot|>:是的，我是你的小助手。有什么我可以帮到你的吗？

<|User|>:你好，请问有什么我可以帮到你的吗？

<|Bot|>:当然，我可以帮你完成各种任务，比如发送邮件、制定日程、查找

04/23 22:43:19 - mmengine - INFO - Sample output:

<s><|User|>:请你介绍一下你自己

<|Bot|>:我是蜻蜓队长的小助手，内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 22:43:21 - mmengine - INFO - Sample output:

<s><|User|>:你是谁

<|Bot|>:我是蜻蜓队长的小助手，内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 22:43:21 - mmengine - INFO - Sample output:

<s><|User|>:你是我的小助手吗

<|Bot|>:是的</s>

04/23 22:43:21 - mmengine - INFO - Saving checkpoint at 300 iterations

过拟合了

```

04/23 23:06:35 - mmengine - INFO - Sample output:
<s><|User|>:请你介绍一下你自己
<|Bot|>:我是蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 23:06:37 - mmengine - INFO - Sample output:
<s><|User|>:你是谁
<|Bot|>:蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 23:06:39 - mmengine - INFO - Sample output:
<s><|User|>:你是我的小助手吗
<|Bot|>:是的, 我是蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>







04/23 23:06:39 - mmengine - INFO - Saving checkpoint at 768 iterations
04/23 23:06:40 - mmengine - INFO - after_train in EvaluateChatHook.
04/23 23:06:42 - mmengine - INFO - Sample output:
<s><|User|>:请你介绍一下你自己
<|Bot|>:我是蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 23:06:44 - mmengine - INFO - Sample output:
<s><|User|>:你是谁
<|Bot|>:蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

04/23 23:06:46 - mmengine - INFO - Sample output:
<s><|User|>:你是我的小助手吗
<|Bot|>:是的, 我是蠟翅队长的小助手, 内在是上海AI实验室书生 浦语的1.8B大模型哦</s>

```

训练保存的模型：

| Name | Last Modified |
|---|----------------|
|  20240423... | 48 minutes ago |
|  internlm2_... | 46 minutes ago |
|  iter_300.pth | 28 minutes ago |
|  iter_600.pth | 13 minutes ago |
|  iter_768.pth | 5 minutes ago |
|  last_check... | 5 minutes ago |