

```
!git clone https://github.com/hiyogua/LLaMA-Factory
```

```
Cloning into 'LLaMA-Factory'...
remote: Enumerating objects: 5981, done.
remote: Total 5981 (delta 0), reused 0 (delta 0), pack-reused 5981
Receiving objects: 100% (5981/5981), 186.06 MiB | 20.10 MiB/s, done.
Resolving deltas: 100% (4351/4351), done.
Updating files: 100% (132/132), done.
```

```
%cd LLaMA-Factory
```

```
/content/LLaMA-Factory
```

```
!pip install -r requirements.txt
```

```
Requirement already satisfied: torch>=1.13.1 in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 1)) (2.1.0+cu121)
Collecting transformers>=4.36.2 (from -r requirements.txt (line 2))
  Downloading transformers-4.36.2-py3-none-any.whl (8.2 MB)
----- 8.2/8.2 MB 24.3 MB/s eta 0:00:00
Collecting datasets>=2.14.3 (from -r requirements.txt (line 3))
  Downloading datasets-2.16.1-py3-none-any.whl (507 kB)
----- 507.1/507.1 kB 35.3 MB/s eta 0:00:00
Collecting accelerate>=0.21.0 (from -r requirements.txt (line 4))
  Downloading accelerate-0.25.0-py3-none-any.whl (265 kB)
----- 265.7/265.7 kB 38.5 MB/s eta 0:00:00
Collecting peft>=0.7.0 (from -r requirements.txt (line 5))
  Downloading peft-0.7.1-py3-none-any.whl (168 kB)
----- 168.3/168.3 kB 26.5 MB/s eta 0:00:00
Collecting trl>=0.7.6 (from -r requirements.txt (line 6))
  Downloading trl-0.7.8-py3-none-any.whl (140 kB)
----- 140.1/140.1 kB 22.8 MB/s eta 0:00:00
Collecting gradio<4.0.0,>=3.38.0 (from -r requirements.txt (line 7))
  Downloading gradio-3.50.2-py3-none-any.whl (20.3 MB)
----- 20.3/20.3 MB 71.5 MB/s eta 0:00:00
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 8)) (1.11.4)
Collecting einops (from -r requirements.txt (line 9))
  Downloading einops-0.7.0-py3-none-any.whl (44 kB)
----- 44.6/44.6 kB 7.6 MB/s eta 0:00:00
Collecting sentencepiece (from -r requirements.txt (line 10))
  Downloading sentencepiece-0.1.99-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.3 MB)
----- 1.3/1.3 MB 86.0 MB/s eta 0:00:00
Requirement already satisfied: protobuf in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 11)) (3.20.3)
Collecting tiktoken (from -r requirements.txt (line 12))
  Downloading tiktoken-0.5.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.0 MB)
----- 2.0/2.0 MB 94.5 MB/s eta 0:00:00
Requirement already satisfied: jieba in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 13)) (0.42.1)
Collecting rouge-chinese (from -r requirements.txt (line 14))
  Downloading rouge_chinese-1.0.3-py3-none-any.whl (21 kB)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 15)) (3.8.1)
Collecting uvicorn (from -r requirements.txt (line 16))
  Downloading uvicorn-0.25.0-py3-none-any.whl (60 kB)
----- 60.3/60.3 kB 9.9 MB/s eta 0:00:00
Requirement already satisfied: pydantic in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 17)) (1.10.13)
Collecting fastapi (from -r requirements.txt (line 18))
  Downloading fastapi-0.108.0-py3-none-any.whl (92 kB)
----- 92.0/92.0 kB 15.5 MB/s eta 0:00:00
Collecting sse-starlette (from -r requirements.txt (line 19))
  Downloading sse_starlette-1.8.2-py3-none-any.whl (8.9 kB)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (from -r requirements.txt (line 20)) (3.7.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (3.13.1)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (3.1.2)
Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (1.12)
Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (3.2.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (3.1.2)
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (2023.6.0)
Requirement already satisfied: triton==2.1.0 in /usr/local/lib/python3.10/dist-packages (from torch>=1.13.1->-r requirements.txt (line 1)) (2.1)
Requirement already satisfied: huggingface-hub<1.0,>=0.19.3 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2))
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2))
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2))
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2))
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2))
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers>=4.36.2->-r requirements.txt (line 2)) (2.28.1)
```

```
!pip install einops transformers_stream_generator
```

```
Requirement already satisfied: einops in /usr/local/lib/python3.10/dist-packages (0.7.0)
Collecting transformers_stream_generator
  Downloading transformers_stream_generator-0.0.4.tar.gz (12 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: transformers>=4.26.1 in /usr/local/lib/python3.10/dist-packages (from transformers_stream_generator) (4.36.2)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (3.13.1)
Requirement already satisfied: huggingface-hub<1.0,>=0.19.3 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (0.19.3)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (1.24.4)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (23.1)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (6.0.1)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (2023.12.31)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (2.31.0)
```

```
Requirement already satisfied: tokenizers<0.19,>=0.14 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_ge
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_genera
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers>=4.26.1->transformers_stream_generator) (4.
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub<1.0,>=0.19.3->transformers>=4.26.1
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub<1.0,>=0.19.3->transforme
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->transformers>=4.26.1->transforme
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->transformers>=4.26.1->transformers_stream_ge
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->transformers>=4.26.1->transformers_str
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->transformers>=4.26.1->transformers_str
Building wheels for collected packages: transformers_stream_generator
  Building wheel for transformers_stream_generator (setup.py) ... done
  Created wheel for transformers_stream_generator: filename=transformers_stream_generator-0.0.4-py3-none-any.whl size=12316 sha256=a5711960d168a48e
  Stored in directory: /root/.cache/pip/wheels/47/1d/3c/92d88493ed40c0d9be60a391eb76c9a56e9f9b7542cb789401
Successfully built transformers_stream_generator
Installing collected packages: transformers_stream_generator
Successfully installed transformers_stream_generator-0.0.4
```

```
!sed -i 's/share=False/share=True/' src/train_web.py
!sed -i 's/share=False/share=True/' src/web_demo.py
```

```
!python src/train_web.py
```

```
fp16=True,
fp16_backend=auto,
fp16_full_eval=False,
fp16_opt_level=01,
fsdp=[],
fsdp_config={'min_num_params': 0, 'xla': False, 'xla_fsdp_grad_ckpt': False},
fsdp_min_num_params=0,
fsdp_transformer_layer_cls_to_wrap=None,
full_determinism=False,
generation_config=None,
generation_max_length=None,
generation_num_beams=None,
gradient_accumulation_steps=4,
gradient_checkpointing=False,
gradient_checkpointing_kwargs=None,
greater_is_better=None,
group_by_length=False,
half_precision_backend=auto,
hub_always_push=False,
hub_model_id=None,
hub_private_repo=False,
hub_strategy=every_save,
hub_token=<HUB_TOKEN>,
ignore_data_skip=False,
include_inputs_for_metrics=False,
include_num_input_tokens_seen=False,
include_tokens_per_second=False,
jit_mode_eval=False,
label_names=None,
label_smoothing_factor=0.0,
learning_rate=0.0005,
length_column_name=length,
load_best_model_at_end=False,
local_rank=0,
log_level=passive,
log_level_replica=warning,
log_on_each_node=True,
logging_dir=saves/ChineseLLaMA2-1.3B-Chat/lora/train_2024-01-09-09-58-02/runs/Jan09_10-13-25_e9ff40fc51d9,
logging_first_step=False,
logging_nan_inf_filter=True,
logging_steps=5,
logging_strategy=steps,
lr_scheduler_kwargs={},
lr_scheduler_type=cosine,
max_grad_norm=1.0,
max_steps=-1,
metric_for_best_model=None,
mp_parameters=,
neptune_noise_alpha=None,
no_cuda=False,
num_train_epochs=8.0,
optim=adamw_torch,
optim_args=None,
output_dir=saves/ChineseLLaMA2-1.3B-Chat/lora/train_2024-01-09-09-58-02,
overwrite_output_dir=False,
past_index=-1,
per_device_eval_batch_size=8,
per_device_train_batch_size=6
```

