```
pip install transformers[torch] datasets
```

```
Requirement already satisfied: datasets in /usr/local/lib/python3.12/dist-packages (4.0.0)
Requirement already satisfied: transformers[torch] in /usr/local/lib/python3.12/dist-packages (4.57.6)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (3.20.3)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transformers[
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (25.0
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (20
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers[t
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (0
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (4.67.1)
Requirement already satisfied: torch>=2.2 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (2.9.0+cu1
Requirement already satisfied: accelerate>=0.26.0 in /usr/local/lib/python3.12/dist-packages (from transformers[torch]) (1
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.12/dist-packages (from datasets) (18.1.0)
Requirement already satisfied: dill<0.3.9,>=0.3.0 in /usr/local/lib/python3.12/dist-packages (from datasets) (0.3.8)
Requirement already satisfied: pandas in /usr/local/lib/python3.12/dist-packages (from datasets) (2.2.2)
Requirement already satisfied: xxhash in /usr/local/lib/python3.12/dist-packages (from datasets) (3.6.0)
Requirement already satisfied: multiprocess<0.70.17 in /usr/local/lib/python3.12/dist-packages (from datasets) (0.70.16)
Requirement already satisfied: fsspec<=2025.3.0,>=2023.1.0 in /usr/local/lib/python3.12/dist-packages (from fsspec[http]<=
Requirement already satisfied: psutil in /usr/local/lib/python3.12/dist-packages (from accelerate>=0.26.0->transformers[to
Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python3.12/dist-packages (from fsspec[http]<=2
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->transfo
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers[torch]
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transformers[
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transformers[
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->transformers[torch]
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->transformers[tor
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->transformers[t
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->transformers[torch]) (3
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=2
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2
Requirement already satisfied: nvidia-cudnn-cu12==9.10.2.21 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->t
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->t
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->tr
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2-
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2-
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->
Requirement already satisfied: nvidia-nccl-cu12==2.27.5 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->trans
Requirement already satisfied: nvidia-nvshmem-cu12==3.3.20 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->tr
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->tran
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2-
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->t
Requirement already satisfied: triton==3.5.0 in /usr/local/lib/python3.12/dist-packages (from torch>=2.2->transformers[tor
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets) (
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.12/dist-packages (from pandas->datasets) (2025.3)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!=4.0
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a
```

```python
import torch
from transformers import GPT2Tokenizer, GPT2LMHeadModel, Trainer, TrainingArguments, TextDataset, DataCollatorForLanguageMod

model_name = "gpt2"
tokenizer = GPT2Tokenizer.from_pretrained(model_name)
model = GPT2LMHeadModel.from_pretrained(model_name)

with open("train.txt", "w") as f:
    for _ in range(100):
        f.write("Artificial intelligence is transforming the world.\n")
        f.write("GPT-2 is a powerful transformer model for text generation.\n")
        f.write("Fine-tuning allows models to learn specific styles and structures.\n")

def load_dataset(file_path, tokenizer, block_size=128):
    dataset = TextDataset(
        tokenizer=tokenizer,
        file_path=file_path,
        block_size=block_size,
    )
    return dataset

train_dataset = load_dataset("train.txt", tokenizer)
data_collator = DataCollatorForLanguageModeling(tokenizer=tokenizer, mlm=False)
```

```python
    training_args = TrainingArguments(
        output_dir="./gpt2-finetuned",
        overwrite_output_dir=True,
        num_train_epochs=3,
        per_device_train_batch_size=4,
        save_steps=100,
        logging_steps=10,
    )

    trainer = Trainer(
        model=model,
        args=training_args,
        data_collator=data_collator,
        train_dataset=train_dataset,
    )

    print("Starting Training...")
    trainer.train()

    model.save_pretrained("./gpt2-finetuned")
    tokenizer.save_pretrained("./gpt2-finetuned")

    def generate_text(prompt, max_length=50):
        device = "cuda" if torch.cuda.is_available() else "cpu"
        model.to(device)

        input_ids = tokenizer.encode(prompt, return_tensors='pt').to(device)

        output = model.generate(
            input_ids,
            max_length=max_length,
            num_return_sequences=1,
            no_repeat_ngram_size=2,
            do_sample=True,
            top_k=50,
            top_p=0.95,
            temperature=0.7
        )

        return tokenizer.decode(output[0], skip_special_tokens=True)

    print("\n--- Generated Text ---")
    prompt = "The future of AI"
    generated_output = generate_text(prompt)
    print(generated_output)
```

```
Map (num_proc=4): 100%  ████████████████████████  3000/3000 [00:00<00:00, 1339.01 examples/s]
Map (num_proc=4): 100%  ████████████████████████  3000/3000 [00:00<00:00, 1844.76 examples/s]
Starting Training...
/usr/local/lib/python3.12/dist-packages/notebook/notebookapp.py:191: SyntaxWarning: invalid escape sequence '\/'
  | |_| | '_ \/ _` / _` |  _/ -_)
wandb: (1) Create a W&B account
wandb: (2) Use an existing W&B account
wandb: (3) Don't visualize my results
wandb: Enter your choice: 2
wandb: You chose 'Use an existing W&B account'
wandb: Logging into https://api.wandb.ai. (Learn how to deploy a W&B server locally: https://wandb.me/wandb-server)
wandb: Find your API key here: https://wandb.ai/authorize?ref=models
wandb: Paste an API key from your profile and hit enter: ··········
wandb: No netrc file found, creating one.
wandb: Appending key for api.wandb.ai to your netrc file: /root/.netrc
wandb: Currently logged in as: dishagandhi809 (dishagandhi809-prodigy-information-technologies) to https://api.wandb.ai. Use `wandb login --relogin` to force relogin
Tracking run with wandb version 0.24.0
Run data is saved locally in /content/wandb/run-20260128_164321-n4fr9rmc
Syncing run soft-oath-3 to Weights & Biases (docs)
View project at https://wandb.ai/dishagandhi809-prodigy-information-technologies/huggingface
View run at https://wandb.ai/dishagandhi809-prodigy-information-technologies/huggingface/runs/n4fr9rmc
`loss_type=None` was set in the config but it is unrecognized. Using the default loss: `ForCausalLMLoss`.
████████████████████████  [198/198 00:39, Epoch 3/3]
```

Step	Training Loss
10	0.812000
20	0.112800
30	0.032500
40	0.030600
50	0.032800
60	0.024900
70	0.017200
80	0.015700
90	0.021000
100	0.016800
110	0.019900
120	0.010700
130	0.012800
140	0.013700
150	0.021300
160	0.010300
170	0.015300
180	0.011700
190	0.019200

```
The attention mask and the pad token id were not set. As a consequence, you may observe unexpected behavior. Please pass your input's `attention_mask` to obtain reliable results.
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's `attention_mask` to obtain reliable results.

--- Generated Text ---
The future of AI is transforming the world.GPT-2 is a powerful transformer model for text generation.Fine-tuning allows models to learn specific styles and structures.Artificial intelligence is reshaping the universe.Geometry allows model to explore
```