```
!git clone https://github.com/sahil280114/codealpaca.git
!cd codealpaca
```

```
Cloning into 'codealpaca'...
remote: Enumerating objects: 79, done.
remote: Counting objects: 100% (23/23), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 79 (delta 19), reused 15 (delta 15), pack-reused 56 (from 1)
Receiving objects: 100% (79/79), 7.67 MiB | 6.05 MiB/s, done.
Resolving deltas: 100% (34/34), done.
```

```
!nvidia-smi
```

```
Tue Mar 18 15:05:04 2025
+-----------------------------------------------------------------------------------------+
| NVIDIA-SMI 550.54.15              Driver Version: 550.54.15      CUDA Version: 12.4     |
|-----------------------------------------+------------------------+----------------------+
| GPU  Name                 Persistence-M | Bus-Id          Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |           Memory-Usage | GPU-Util  Compute M. |
|                                         |                        |               MIG M. |
|=========================================+========================+======================|
|   0  Tesla T4                       Off |   00000000:00:04.0 Off |                    0 |
| N/A   59C    P8             11W /   70W |       0MiB /  15360MiB |      0%      Default |
|                                         |                        |                  N/A |
+-----------------------------------------+------------------------+----------------------+

+-----------------------------------------------------------------------------------------+
| Processes:                                                                              |
|  GPU   GI   CI        PID   Type   Process name                              GPU Memory |
|        ID   ID                                                               Usage      |
|=========================================================================================|
|  No running processes found                                                             |
+-----------------------------------------------------------------------------------------+
```

```
%cd codealpaca
```

```
/content/codealpaca
```

```
ls -R codealpaca
```

```
ls: cannot access 'codealpaca': No such file or directory
```

```
!ls
!git clone https://github.com/sahil280114/codealpaca.git
```

```
convert_to_hf.py  ds_config.json           nolora.py   requirements.txt  utils.py
data              generate_instruction.py  prompt.txt  seed_tasks.jsonl
DATA_LICENSE      LICENSE                  README.md   train.py
Cloning into 'codealpaca'...
remote: Enumerating objects: 79, done.
remote: Counting objects: 100% (23/23), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 79 (delta 19), reused 15 (delta 15), pack-reused 56 (from 1)
Receiving objects: 100% (79/79), 7.67 MiB | 5.46 MiB/s, done.
Resolving deltas: 100% (34/34), done.
```

```
!ls
!ls -R codealpaca
```

```
codealpaca        DATA_LICENSE             LICENSE     README.md         train.py
convert_to_hf.py  ds_config.json           nolora.py   requirements.txt  utils.py
data              generate_instruction.py  prompt.txt  seed_tasks.jsonl
codealpaca:
convert_to_hf.py  ds_config.json           nolora.py   requirements.txt  utils.py
data              generate_instruction.py  prompt.txt  seed_tasks.jsonl
DATA_LICENSE      LICENSE                  README.md   train.py

codealpaca/data:
code_alpaca_20k.json  code_alpaca_2k.json  new_codealpaca.json  rosetta_alpaca.json
```

```python
import json
import pandas as pd

# Choose the dataset you want (modify if needed)
dataset_path = "codealpaca/data/code_alpaca_2k.json"  # Change to another file if needed
```

```
# Load dataset
with open(dataset_path, "r") as f:
    data = json.load(f)

# Convert to DataFrame
df = pd.DataFrame(data)

# Print first few rows
print(df.head())

# Print dataset shape
print("\nDataset Shape:", df.shape)

# Print column names
print("\nColumn Names:", df.columns)
```

```
                                         instruction  \
0  What are the distinct values from the given list?   
1  How would you order a sequence of letters alph...   
2  Write a JavaScript code to loop over all eleme...   
3  Write a Python function to calculate the facto...   
4  What would be the output of the following Java...   

                                          input  \
0              dataList = [3, 9, 3, 5, 7, 9, 5]   
1                                    A, B, C, D   
2            numbersArray = [45, 6, 23, 12, 35]   
3                                                 
4  let area = 6 * 5;\nlet radius = area / 3.14;   

                                              output  
0  The distinct values from the given list are 3,...  
1  The sequence of letters ordered alphabetically...  
2  for(let i = 0; i < numbersArray.length; i++) {...  
3  def factorial(number):\n    fact = 1\n    for ...  
4  The output of the JavaScript snippet is the ra...  

Dataset Shape: (2017, 3)

Column Names: Index(['instruction', 'input', 'output'], dtype='object')
```

```
dataset = Dataset.from_pandas(df)
```

```
df.head()
```

	instruction	input	output
0	What are the distinct values from the given list?	dataList = [3, 9, 3, 5, 7, 9, 5]	The distinct values from the given list are 3,...
1	How would you order a sequence of letters alph...	A, B, C, D	The sequence of letters ordered alphabetically...
2	Write a JavaScript code to loop over all eleme...	numbersArray = [45, 6, 23, 12, 35]	for(let i = 0; i < numbersArray.length; i++) {...
3	Write a Python function to calculate the facto...		def factorial(number):\n fact = 1\n for ...
4	What would be the output of the following Java...	let area = 6 * 5;\nlet radius = area / 3.14;	The output of the JavaScript snippet is the ra...

```
import json
import pandas as pd

# Choose the dataset you want (modify if needed)
dataset_path1 = "codealpaca/data/code_alpaca_20k.json"  # Change to another file if needed

# Load dataset
with open(dataset_path1, "r") as f:
    data = json.load(f)

# Convert to DataFrame
df1 = pd.DataFrame(data)

# Print first few rows
print(df1.head())

# Print dataset shape
print("\nDataset Shape:", df1.shape)

# Print column names
print("\nColumn Names:", df1.columns)

df1.head()
```

```
                                         instruction  \
0  Create an array of length 5 which contains all...
1  Formulate an equation to calculate the height ...
2  Write a replace method for a string class whic...
3  Create an array of length 15 containing number...
4  Write a function to find the number of distinc...

                                               input  \
0
1
2  string = "Hello World!"\nreplace_with = "Greet...
3
4  matrix = [[1, 0, 0],\n          [1, 0, 1],\n  ...

                                              output
0                             arr = [2, 4, 6, 8, 10]
1  Height of triangle = opposite side length * si...
2  def replace(self, replace_with):\n    new_stri...
3  arr = [3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33...
4  def find_num_distinct_states(matrix):\n    sta...

Dataset Shape: (20022, 3)

Column Names: Index(['instruction', 'input', 'output'], dtype='object')
```

	instruction	input	output
0	Create an array of length 5 which contains all...		arr = [2, 4, 6, 8, 10]
1	Formulate an equation to calculate the height ...		Height of triangle = opposite side length * si...
2	Write a replace method for a string class whic...	string = "Hello World!"\nreplace_with = "Greet...	def replace(self, replace_with):\n new_stri...
3	Create an array of length 15 containing number...		arr = [3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33...
4	Write a function to find the number of distinc...	matrix = [[1, 0, 0],\n [1, 0, 1],\n ...	def find_num_distinct_states(matrix):\n sta...

```
!pip install --upgrade gspread
```

```
Requirement already satisfied: gspread in /usr/local/lib/python3.11/dist-packages (6.1.4)
Collecting gspread
  Downloading gspread-6.2.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: google-auth>=1.12.0 in /usr/local/lib/python3.11/dist-packages (from gspread) (2.38.0)
Requirement already satisfied: google-auth-oauthlib>=0.4.1 in /usr/local/lib/python3.11/dist-packages (from gspread) (1.2.1)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from google-auth>=1.12.0->gsp
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.11/dist-packages (from google-auth>=1.12.0->gspr
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.11/dist-packages (from google-auth>=1.12.0->gspread) (4
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from google-auth-oauthlib>=
Requirement already satisfied: pyasn1<0.7.0,>=0.4.6 in /usr/local/lib/python3.11/dist-packages (from pyasn1-modules>=0.2.1->goo
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.11/dist-packages (from requests-oauthlib>=0.7.0->googl
Requirement already satisfied: requests>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from requests-oauthlib>=0.7.0->googl
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests>=2.0.0->reque
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests>=2.0.0->requests-oauthlib
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests>=2.0.0->requests-oa
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests>=2.0.0->requests-oa
Downloading gspread-6.2.0-py3-none-any.whl (59 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 59.9/59.9 kB 4.6 MB/s eta 0:00:00
Installing collected packages: gspread
  Attempting uninstall: gspread
    Found existing installation: gspread 6.1.4
    Uninstalling gspread-6.1.4:
      Successfully uninstalled gspread-6.1.4
Successfully installed gspread-6.2.0
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
import pandas as pd

# Load dataset
dataset_path = "codealpaca/data/code_alpaca_2k.json"  # Change file if needed
df = pd.read_json(dataset_path)

# Save CSV to Google Drive
csv_path = "/content/drive/My Drive/CodeAlpaca_Dataset.csv"
df.to_csv(csv_path, index=False)

print(f"✅ Dataset saved to Google Drive: {csv_path}")
```

```
✅ Dataset saved to Google Drive: /content/drive/My Drive/CodeAlpaca_Dataset.csv
```

```
!pip install transformers accelerate bitsandbytes peft trl datasets
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (4.48.3)
Requirement already satisfied: accelerate in /usr/local/lib/python3.11/dist-packages (1.3.0)
Collecting bitsandbytes
  Downloading bitsandbytes-0.45.3-py3-none-manylinux_2_24_x86_64.whl.metadata (5.0 kB)
Requirement already satisfied: peft in /usr/local/lib/python3.11/dist-packages (0.14.0)
Collecting trl
  Downloading trl-0.15.2-py3-none-any.whl.metadata (11 kB)
Collecting datasets
  Downloading datasets-3.4.1-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from transformers) (3.17.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.24.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (0
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (24.2)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2024.11.6)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)
Requirement already satisfied: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.1 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.5.3)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from accelerate) (5.9.5)
Requirement already satisfied: torch>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from accelerate) (2.6.0+cu124)
Requirement already satisfied: rich in /usr/local/lib/python3.11/dist-packages (from trl) (13.9.4)
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/dist-packages (from datasets) (18.1.0)
Collecting dill<0.3.9,>=0.3.0 (from datasets)
  Downloading dill-0.3.8-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (from datasets) (2.2.2)
Collecting xxhash (from datasets)
  Downloading xxhash-3.5.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (12 kB)
Collecting multiprocess<0.70.17 (from datasets)
  Downloading multiprocess-0.70.16-py311-none-any.whl.metadata (7.2 kB)
Requirement already satisfied: fsspec<=2024.12.0,>=2023.1.0 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]<=20
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-packages (from datasets) (3.11.13)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (2
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.3.2)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (25.3.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.5.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (6.1.0
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (0.3.0)
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.18.3)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->transforme
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.1.6)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-manylinux2014_x86_64.whl.metadata (1.5 kB)
```

```
# import torch
# import bitsandbytes as bnb
# from transformers import AutoModelForCausalLM

# print(f"CUDA Available: {torch.cuda.is_available()}")
# print(f"BitsAndBytes Installed: {hasattr(bnb, 'nn')}")



# Import necessary libraries
import torch
import bitsandbytes as bnb
from transformers import AutoModelForCausalLM, AutoTokenizer, TrainingArguments, Trainer
from peft import prepare_model_for_kbit_training, LoraConfig, get_peft_model
from datasets import Dataset
import pandas as pd
import json

# Check GPU availability
print(f"CUDA Available: {torch.cuda.is_available()}")
print(f"BitsAndBytes Installed: {hasattr(bnb, 'nn')}")
```

```
CUDA Available: True
BitsAndBytes Installed: True
```

```
from huggingface_hub import login

login("hf_QpnMxxxxxxSObBQMkTbOe")
```

```
from transformers import AutoModelForCausalLM, AutoTokenizer
import torch

# Correct model name
model_name = "bigcode/starcoderbase-3b"

# Load the model in 4-bit quantization
model = AutoModelForCausalLM.from_pretrained(
    model_name,
    load_in_4bit=True,
    torch_dtype=torch.float16,
    device_map="auto"
)

# Load the tokenizer
tokenizer = AutoTokenizer.from_pretrained(model_name)

print("✅ Model Loaded Successfully")
```

```
config.json: 100%                                    1.05k/1.05k [00:00<00:00, 64.9kB/s]
The `load_in_4bit` and `load_in_8bit` arguments are deprecated and will be removed in the future versions. Please, pass a `Bits
pytorch_model.bin.index.json: 100%                   32.7k/32.7k [00:00<00:00, 3.41MB/s]
Downloading shards: 100%                             2/2 [02:07<00:00, 56.01s/it]
pytorch_model-00001-of-00002.bin: 100%               10.0G/10.0G [01:47<00:00, 144MB/s]
model.safetensors.index.json: 100%                   34.4k/34.4k [00:00<00:00, 2.62MB/s]
pytorch_model-00002-of-00002.bin: 100%               2.18G/2.18G [00:19<00:00, 195MB/s]
Loading checkpoint shards: 100%                      2/2 [00:52<00:00, 23.38s/it]
generation_config.json: 100%                         111/111 [00:00<00:00, 7.10kB/s]
tokenizer_config.json: 100%                          677/677 [00:00<00:00, 65.4kB/s]
vocab.json: 100%                                     777k/777k [00:00<00:00, 1.85MB/s]
merges.txt: 100%                                     442k/442k [00:00<00:00, 2.09MB/s]
tokenizer.json: 100%                                 2.06M/2.06M [00:00<00:00, 3.15MB/s]
special_tokens_map.json: 100%                        532/532 [00:00<00:00, 52.9kB/s]
✅ Model Loaded Successfully
```

```
#  hf_QpnxxxxxxQMkTbOe  token
```

```
from peft import prepare_model_for_kbit_training, LoraConfig, get_peft_model
from transformers import TrainingArguments, Trainer
from datasets import Dataset
```

```
# Set the EOS token as the padding token
tokenizer.pad_token = tokenizer.eos_token
```

```
# Prepare the model for QLoRA fine-tuning
model = prepare_model_for_kbit_training(model)
```

```
# Print the model architecture
print(model)
```

```
GPTBigCodeForCausalLM(
  (transformer): GPTBigCodeModel(
    (wte): Embedding(49152, 2816)
    (wpe): Embedding(8192, 2816)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPTBigCodeBlock(
        (ln_1): LayerNorm((2816,), eps=1e-05, elementwise_affine=True)
        (attn): GPTBigCodeSdpaAttention(
          (c_attn): Linear4bit(in_features=2816, out_features=3072, bias=True)
          (c_proj): Linear4bit(in_features=2816, out_features=2816, bias=True)
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((2816,), eps=1e-05, elementwise_affine=True)
        (mlp): GPTBigCodeMLP(
          (c_fc): Linear4bit(in_features=2816, out_features=11264, bias=True)
          (c_proj): Linear4bit(in_features=11264, out_features=2816, bias=True)
          (act): PytorchGELUTanh()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((2816,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=2816, out_features=49152, bias=False)
)
```

```
# Define LoRA configuration for StarCoder models
lora_config = LoraConfig(
    r=8,  # Rank of the low-rank adaptation
    lora_alpha=32,  # Scaling factor
    target_modules=["c_attn"],  # Target the combined query/key/value projection layer
    lora_dropout=0.1,  # Dropout for LoRA
    bias="none",  # No bias
    task_type="CAUSAL_LM"  # Task type
)

# Apply LoRA to the model
model = get_peft_model(model, lora_config)
```

```
def tokenize_function(examples):
    # Concatenate instruction and input for each example in the batch
    texts = [inst + " " + inp for inst, inp in zip(examples["instruction"], examples["input"])]

    # Tokenize the batch of texts
    return tokenizer(
        texts,
        padding="max_length",  # Pad to max_length
        truncation=True,       # Truncate to max_length
        max_length=512,       # Set max_length
        return_tensors="pt"   # Return PyTorch tensors
    )

tokenized_dataset = dataset.map(tokenize_function, batched=True)
```

```
Map: 100%                                    2017/2017 [00:01<00:00, 1444.99 examples/s]
```

```
# Tokenize the dataset
def tokenize_function(examples):
    # Concatenate instruction and input for each example in the batch
    texts = [inst + " " + inp for inst, inp in zip(examples["instruction"], examples["input"])]

    # Tokenize the batch of texts
    tokenized = tokenizer(
        texts,
        padding="max_length",  # Pad to max_length
        truncation=True,       # Truncate to max_length
        max_length=512,       # Set max_length
        return_tensors="pt"   # Return PyTorch tensors
    )

    # Add labels for causal language modeling
    tokenized["labels"] = tokenized["input_ids"].clone()
    return tokenized

tokenized_dataset = dataset.map(tokenize_function, batched=True)

# Set up training arguments
training_args = TrainingArguments(
    output_dir="./results",
    run_name="starcoder-finetune",  # Custom run name for W&B
    per_device_train_batch_size=2,  # Small batch size to save memory
    per_device_eval_batch_size=2,
    gradient_accumulation_steps=4,  # Accumulate gradients to simulate larger batch size
    num_train_epochs=3,
    logging_dir="./logs",
    logging_steps=10,
    save_steps=500,
    evaluation_strategy="steps",
    eval_steps=500,
    save_total_limit=2,
    fp16=True,  # Mixed precision training
    gradient_checkpointing=True,  # Enable gradient checkpointing
    push_to_hub=False,
    report_to="none"  # Disable W&B if you don't want to use it
)

# Initialize the Trainer
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=tokenized_dataset,
    eval_dataset=tokenized_dataset  # Use the same dataset for evaluation (or split it)
)

# Start training
trainer.train()

# Save the fine-tuned model
model.save_pretrained("/content/fine-tuned-model")
tokenizer.save_pretrained("/content/fine-tuned-model")

print("✅ Fine-tuning complete and model saved!")
```

```
Map: 100%                                   2017/2017 [00:01<00:00, 1534.03 examples/s]
/usr/local/lib/python3.11/dist-packages/transformers/training_args.py:1575: FutureWarning: `evaluation_strategy` is deprecated
  warnings.warn(
[756/756 1:31:07, Epoch 2/3]
```

Step	Training Loss	Validation Loss
500	0.077500	0.077198

```
✅ Fine-tuning complete and model saved!
```

```
!ls /content/fine-tuned-model
```

```
adapter_config.json        merges.txt   special_tokens_map.json  tokenizer.json
adapter_model.safetensors  README.md    tokenizer_config.json    vocab.json
```

```
from google.colab import files

# Zip the model directory
!zip -r fine-tuned-model.zip /content/fine-tuned-model

# Download the zip file
files.download("fine-tuned-model.zip")
```

```
  adding: content/fine-tuned-model/ (stored 0%)
  adding: content/fine-tuned-model/vocab.json (deflated 57%)
  adding: content/fine-tuned-model/merges.txt (deflated 51%)
  adding: content/fine-tuned-model/tokenizer.json (deflated 81%)
  adding: content/fine-tuned-model/README.md (deflated 66%)
  adding: content/fine-tuned-model/special_tokens_map.json (deflated 73%)
  adding: content/fine-tuned-model/adapter_config.json (deflated 54%)
  adding: content/fine-tuned-model/adapter_model.safetensors (deflated 8%)
  adding: content/fine-tuned-model/tokenizer_config.json (deflated 88%)
```

```
# from google.colab import drive
# drive.mount('/content/drive')

# Copy the model to Google Drive
!cp -r /content/fine-tuned-model /content/drive/My\ Drive/
```

```
from transformers import AutoModelForCausalLM, AutoTokenizer

# Load the fine-tuned model
model = AutoModelForCausalLM.from_pretrained("/content/fine-tuned-model")

# Load the tokenizer
tokenizer = AutoTokenizer.from_pretrained("/content/fine-tuned-model")
```

```
Loading checkpoint shards: 100%                                   2/2 [00:02<00:00,  1.07s/it]
```

```
# Sample input
input_text = "Write a Python function to calculate the factorial of a number."

# Tokenize the input
inputs = tokenizer(input_text, return_tensors="pt")
```

```
# Generate output
outputs = model.generate(**inputs, max_length=100)

# Decode and print the output
print(tokenizer.decode(outputs[0], skip_special_tokens=True))
```

```
Setting `pad_token_id` to `eos_token_id`:0 for open-end generation.
Write a Python function to calculate the factorial of a number.
```

```
!pip install huggingface_hub
```

```
Requirement already satisfied: huggingface_hub in /usr/local/lib/python3.11/dist-packages (0.28.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (3.17.0)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (2024.10.0)
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (24.2)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (6.0.2)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (4
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface_
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface_hub) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface_hub) (
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface_hub) (
```

```
from huggingface_hub import login

# Log in with the new token
login(token="hf_ouyQCVxxxxxxKFWwGaqcW")
```

```
from huggingface_hub import HfApi, login, create_repo

# # Log in to Hugging Face with a write token
# login(token="your-new-write-token")

 # Initialize the API
api = HfApi()

# Replace with your Hugging Face username and model name
repo_id = "key-life/starcoder-finetuned-codealpaca"

# Create the repository if it doesn't exist
try:
    create_repo(repo_id, repo_type="model", exist_ok=True)
    print(f"✅ Repository '{repo_id}' created or already exists.")
except Exception as e:
    print(f"❌ Failed to create repository: {e}")
    raise

# Upload the model
try:
    api.upload_folder(
        folder_path="/content/fine-tuned-model",  # Path to your fine-tuned model
        repo_id = repo_id,  # Replace with your Hugging Face repo ID
        repo_type="model"
    )
    print("✅ Model uploaded to Hugging Face Hub!")
except Exception as e:
    print(f"❌ Failed to upload model: {e}")
    raise
```

```
No files have been modified since last commit. Skipping to prevent empty commit.
WARNING:huggingface_hub.hf_api:No files have been modified since last commit. Skipping to prevent empty commit.
```

Start coding or generate with AI.

```
from huggingface_hub import HfApi, login

# Log in to Hugging Face
login(token="hf_ouyQCxxxxxxxxFWwGaqcW")

# Initialize the API
api = HfApi()

# Upload the model
api.upload_folder(
    folder_path="/content/fine-tuned-model",
    repo_id="key-life/starcoder-finetuned-codealpaca",
    repo_type="model"
)
print("✅ Model uploaded to Hugging Face Hub!")
```

```
No files have been modified since last commit. Skipping to prevent empty commit.
WARNING:huggingface_hub.hf_api:No files have been modified since last commit. Skipping to prevent empty commit.
```