

bgdkriza2

May 7, 2025

```
[1]: ! pip install torch torchtext transformers sentencepiece pandas tqdm datasets
```

```
Requirement already satisfied: torch in /usr/local/lib/python3.11/dist-packages
(2.6.0+cu124)
Collecting torchtext
  Downloading torchtext-0.18.0-cp311-cp311-manylinux1_x86_64.whl.metadata (7.9
kB)
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-
packages (4.51.3)
Requirement already satisfied: sentencepiece in /usr/local/lib/python3.11/dist-
packages (0.2.0)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages
(2.2.2)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages
(4.67.1)
Collecting datasets
  Downloading datasets-3.5.1-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-
packages (from torch) (3.18.0)
Requirement already satisfied: typing-extensions>=4.10.0 in
/usr/local/lib/python3.11/dist-packages (from torch) (4.13.2)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-
packages (from torch) (3.4.2)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.11/dist-packages
(from torch) (3.1.6)
Requirement already satisfied: fsspec in /usr/local/lib/python3.11/dist-packages
(from torch) (2025.3.2)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch)
  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)
  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)
  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)
  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)  
   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)  
   Downloading nvidia\_cufft\_cu12-11.2.1.3-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)  
 Collecting nvidia-curand-cu12==10.3.5.147 (from torch)  
   Downloading nvidia\_curand\_cu12-10.3.5.147-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)  
 Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch)  
   Downloading nvidia\_cusolver\_cu12-11.6.1.9-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)  
 Collecting nvidia-cusparselt-cu12==0.6.2 (from torch)  
   Downloading nvidia\_cusparselt\_cu12-0.6.2-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)  
 Requirement already satisfied: nvidia-cusparse-cu12==12.3.1.170 in /usr/local/lib/python3.11/dist-packages (from torch) (12.3.1.170)  
   Downloading nvidia\_cusparse\_cu12-12.3.1.170-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)  
 Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)  
 Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch)  
   Downloading nvidia\_nvjitlink\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)  
 Requirement already satisfied: triton==3.2.0 in /usr/local/lib/python3.11/dist-packages (from torch) (3.2.0)  
 Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-packages (from torch) (1.13.1)  
 Requirement already satisfied: mpmath<1.4, >=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch) (1.3.0)  
 Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from torchtext) (2.32.3)  
 Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from torchtext) (2.0.2)  
 Requirement already satisfied: huggingface-hub<1.0, >=0.30.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.30.2)  
 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: tokenizers<0.22, >=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.21.1)  
 Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.5.3)  
 Requirement already satisfied: python-dateutil>=2.8.2 in

```

/usr/local/lib/python3.11/dist-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-
packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-
packages (from pandas) (2025.2)
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)
Collecting xxhash (from datasets)
  Downloading
xxhash-3.5.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(12 kB)
Collecting multiprocessing<0.70.17 (from datasets)
  Downloading multiprocessing-0.70.16-py311-none-any.whl.metadata (7.2 kB)
Collecting fsspec (from torch)
  Downloading fsspec-2025.3.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-
packages (from datasets) (3.11.15)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (2.6.1)
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) (1.6.0)
Requirement already satisfied: multidict<7.0,>=4.5 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (6.4.3)
Requirement already satisfied: propcache>=0.2.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (0.3.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.20.0)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-
packages (from python-dateutil>=2.8.2->pandas) (1.17.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->torchtext) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-
packages (from requests->torchtext) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->torchtext) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->torchtext) (2025.4.26)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from Jinja2->torch) (3.0.2)
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl (363.4
MB)

```

363.4/363.4 MB

4.4 MB/s eta 0:00:00  
Downloading nvidia\_cuda\_cupti\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (13.8 MB)  
13.8/13.8 MB

97.8 MB/s eta 0:00:00  
Downloading nvidia\_cuda\_nvrtc\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (24.6 MB)  
24.6/24.6 MB

94.8 MB/s eta 0:00:00  
Downloading nvidia\_cuda\_runtime\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (883 kB)  
883.7/883.7 kB

55.1 MB/s eta 0:00:00  
Downloading nvidia\_cudnn\_cu12-9.1.0.70-py3-none-manylinux2014\_x86\_64.whl (664.8 MB)  
664.8/664.8 MB

1.3 MB/s eta 0:00:00  
Downloading nvidia\_cufft\_cu12-11.2.1.3-py3-none-manylinux2014\_x86\_64.whl (211.5 MB)  
211.5/211.5 MB

6.3 MB/s eta 0:00:00  
Downloading nvidia\_curand\_cu12-10.3.5.147-py3-none-manylinux2014\_x86\_64.whl (56.3 MB)  
56.3/56.3 MB

13.3 MB/s eta 0:00:00  
Downloading nvidia\_cusolver\_cu12-11.6.1.9-py3-none-manylinux2014\_x86\_64.whl (127.9 MB)  
127.9/127.9 MB

7.5 MB/s eta 0:00:00  
Downloading nvidia\_cusparsparse\_cu12-12.3.1.170-py3-none-manylinux2014\_x86\_64.whl (207.5 MB)  
207.5/207.5 MB

5.6 MB/s eta 0:00:00  
Downloading nvidia\_nvjitlink\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (21.1 MB)  
21.1/21.1 MB

101.4 MB/s eta 0:00:00  
Downloading torchtext-0.18.0-cp311-cp311-manylinux1\_x86\_64.whl (2.0 MB)  
2.0/2.0 MB

86.9 MB/s eta 0:00:00  
Downloading datasets-3.5.1-py3-none-any.whl (491 kB)  
491.4/491.4 kB

32.5 MB/s eta 0:00:00  
Downloading dill-0.3.8-py3-none-any.whl (116 kB)  
116.3/116.3 kB

11.4 MB/s eta 0:00:00  
Downloading fsspec-2025.3.0-py3-none-any.whl (193 kB)  
193.6/193.6 kB

18.4 MB/s eta 0:00:00

Downloading multiprocess-0.70.16-py311-none-any.whl (143 kB)

143.5/143.5 kB

13.9 MB/s eta 0:00:00

Downloading

xxhash-3.5.0-cp311-cp311-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (194 kB)

194.8/194.8 kB

18.4 MB/s eta 0:00:00

Installing collected packages: xxhash, nvidia-nvjitlink-cu12, nvidia-curand-cu12, nvidia-cufft-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, fsspec, dill, nvidia-cuspars-cu12, nvidia-cudnn-cu12, multiprocess, nvidia-cusolver-cu12, datasets, torchtext

Attempting uninstall: nvidia-nvjitlink-cu12

Found existing installation: nvidia-nvjitlink-cu12 12.5.82

Uninstalling nvidia-nvjitlink-cu12-12.5.82:

Successfully uninstalled nvidia-nvjitlink-cu12-12.5.82

Attempting uninstall: nvidia-curand-cu12

Found existing installation: nvidia-curand-cu12 10.3.6.82

Uninstalling nvidia-curand-cu12-10.3.6.82:

Successfully uninstalled nvidia-curand-cu12-10.3.6.82

Attempting uninstall: nvidia-cufft-cu12

Found existing installation: nvidia-cufft-cu12 11.2.3.61

Uninstalling nvidia-cufft-cu12-11.2.3.61:

Successfully uninstalled nvidia-cufft-cu12-11.2.3.61

Attempting uninstall: nvidia-cuda-runtime-cu12

Found existing installation: nvidia-cuda-runtime-cu12 12.5.82

Uninstalling nvidia-cuda-runtime-cu12-12.5.82:

Successfully uninstalled nvidia-cuda-runtime-cu12-12.5.82

Attempting uninstall: nvidia-cuda-nvrtc-cu12

Found existing installation: nvidia-cuda-nvrtc-cu12 12.5.82

Uninstalling nvidia-cuda-nvrtc-cu12-12.5.82:

Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.5.82

Attempting uninstall: nvidia-cuda-cupti-cu12

Found existing installation: nvidia-cuda-cupti-cu12 12.5.82

Uninstalling nvidia-cuda-cupti-cu12-12.5.82:

Successfully uninstalled nvidia-cuda-cupti-cu12-12.5.82

Attempting uninstall: nvidia-cublas-cu12

Found existing installation: nvidia-cublas-cu12 12.5.3.2

Uninstalling nvidia-cublas-cu12-12.5.3.2:

Successfully uninstalled nvidia-cublas-cu12-12.5.3.2

Attempting uninstall: fsspec

Found existing installation: fsspec 2025.3.2

Uninstalling fsspec-2025.3.2:

Successfully uninstalled fsspec-2025.3.2

Attempting uninstall: nvidia-cuspars-cu12

Found existing installation: nvidia-cuspars-cu12 12.5.1.3

Uninstalling nvidia-cuspars-cu12-12.5.1.3:

Successfully uninstalled nvidia-cuspars-cu12-12.5.1.3

```

Attempting uninstall: nvidia-cudnn-cu12
Found existing installation: nvidia-cudnn-cu12 9.3.0.75
Uninstalling nvidia-cudnn-cu12-9.3.0.75:
  Successfully uninstalled nvidia-cudnn-cu12-9.3.0.75
Attempting uninstall: nvidia-cusolver-cu12
Found existing installation: nvidia-cusolver-cu12 11.6.3.83
Uninstalling nvidia-cusolver-cu12-11.6.3.83:
  Successfully uninstalled nvidia-cusolver-cu12-11.6.3.83
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

gcsfs 2025.3.2 requires fsspec==2025.3.2, but you have fsspec 2025.3.0 which is
incompatible.

Successfully installed datasets-3.5.1 dill-0.3.8 fsspec-2025.3.0
multiprocess-0.70.16 nvidia-cublas-cu12-12.4.5.8 nvidia-cuda-cupti-cu12-12.4.127
nvidia-cuda-nvrtc-cu12-12.4.127 nvidia-cuda-runtime-cu12-12.4.127 nvidia-cudnn-
cu12-9.1.0.70 nvidia-cufft-cu12-11.2.1.3 nvidia-curand-cu12-10.3.5.147 nvidia-
cusolver-cu12-11.6.1.9 nvidia-cusparse-cu12-12.3.1.170 nvidia-nvjitlink-
cu12-12.4.127 torchtext-0.18.0 xxhash-3.5.0

```

```

[2]: from datasets import load_dataset, DatasetDict, Dataset
import pandas as pd
from tqdm import tqdm
import ast
import time
import datasets

```

```

[6]: # Loading dataset

data_sample = load_dataset("QuyenAnhDE/Diseases_Symptoms")

```

Repo card metadata block was not found. Setting CardData to empty.  
 WARNING:huggingface\_hub.repocard:Repo card metadata block was not found. Setting CardData to empty.

```

[7]: data_sample

```

```

[7]: DatasetDict({
  train: Dataset({
    features: ['Code', 'Name', 'Symptoms', 'Treatments'],
    num_rows: 400
  })
})

```

```
[10]: updated_data = [{'Name': item['Name'], 'Symptoms': item['Symptoms']} for item_
      ↪in data_sample['train']]
```

```
[11]: df = pd.DataFrame(updated_data)
```

```
[12]: df.head(5)
```

```
[12]:
```

	Name	Symptoms
0	Panic disorder	Palpitations, Sweating, Trembling, Shortness of breath, Fear of losing control, Dizziness
1	Vocal cord polyp	Hoarseness, Vocal Changes, Vocal Fatigue
2	Turner syndrome	Short stature, Gonadal dysgenesis, Webbed neck, Primary amenorrhea
3	Cryptorchidism	Absence or undescended testicle(s), empty scrotum, Infertility
4	Ethylene glycol poisoning-1	Nausea, vomiting, abdominal pain, General malaise, Hypocalcemia

```
[14]: df['Symptoms'] = df['Symptoms'].apply(lambda x: ', '.join(x.split(', ')))
```

```
[16]: df['Symptoms'][0]
```

```
[16]: 'Palpitations, Sweating, Trembling, Shortness of breath, Fear of losing control, Dizziness'
```

```
[17]: from transformers import GPT2Tokenizer, GPT2LMHeadModel
      import torch
      import torch.nn as nn
      import torch.optim as optim
      from torch.utils.data import DataLoader, Dataset, random_split
```

```
[19]: if torch.cuda.is_available():
      device = torch.device("cuda")
      else:
          try:
              device = torch.device("mps")
          except Exception:
              device = torch.device("cpu")
```

```
[20]: device
```

```
[20]: device(type='cuda')
```

```
[21]: tokenizer = GPT2Tokenizer.from_pretrained("distilgpt2")
      model = GPT2LMHeadModel.from_pretrained("distilgpt2").to(device)
```

tokenizer\_config.json: 0%| | 0.00/26.0 [00:00<?, ?B/s]

vocab.json: 0%| | 0.00/1.04M [00:00<?, ?B/s]

merges.txt: 0%| | 0.00/456k [00:00<?, ?B/s]

tokenizer.json: 0%| | 0.00/1.36M [00:00<?, ?B/s]

config.json: 0%| | 0.00/762 [00:00<?, ?B/s]

Xet Storage is enabled for this repo, but the 'hf\_xet' package is not installed.  
Falling back to regular HTTP download. For better performance, install the  
package with: `pip install huggingface\_hub[hf\_xet]` or `pip install hf\_xet`  
WARNING:huggingface\_hub.file\_download:Xet Storage is enabled for this repo, but  
the 'hf\_xet' package is not installed. Falling back to regular HTTP download.  
For better performance, install the package with: `pip install  
huggingface\_hub[hf\_xet]` or `pip install hf\_xet`

model.safetensors: 0%| | 0.00/353M [00:00<?, ?B/s]

generation\_config.json: 0%| | 0.00/124 [00:00<?, ?B/s]

[22]: model

```
[22]: GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 768)
    (wpe): Embedding(1024, 768)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-5): 6 x GPT2Block(
        (ln_1): LayerNorm((768,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D(nf=2304, nx=768)
          (c_proj): Conv1D(nf=768, nx=768)
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((768,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D(nf=3072, nx=768)
          (c_proj): Conv1D(nf=768, nx=3072)
          (act): NewGELUActivation()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((768,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=768, out_features=50257, bias=False)
)
```



```
[24]: BATCH_SIZE = 8
```

```
[25]: df.describe()
```

```
[25]:
```

	Name	Symptoms
count	400	400
unique	392	395
top	Sciatica	Swelling, pain, dry mouth, bad taste
freq	3	3

```
[29]: # Data Preparation
```

```
class DiseaseSymptomDataset(Dataset):
    def __init__(self, df, tokenizer):
        self.labels = df.columns
        self.data = df.to_dict(orient='records')
        self.tokenizer = tokenizer
        x = self.fittest_max_length(df)
        self.max_length = x

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx):
        x = self.data[idx][self.labels[0]]
        y = self.data[idx][self.labels[1]]
        text = f"{x} | {y}"
        tokens = self.tokenizer.encode_plus(text, return_tensors='pt',
        ↪max_length=128, padding='max_length', truncation=True)
        return tokens

    def fittest_max_length(self, df):
        """
        This function computes maximum sequence length for the dataset.
        """
        max_length = max(len(max(df[self.labels[0]], key=len)), len(max(df[self.
        ↪labels[1]], key=len)))
        x = 2
        while x < max_length: x = x * 2
        return x
```

```
[30]: data_sample = DiseaseSymptomDataset(df, tokenizer)
```

```
[31]: data_sample
```

```
[31]: <__main__.DiseaseSymptomDataset at 0x7eda4c59d4d0>
```

```

[32]: train_size = int(0.8 * len(data_sample))
      val_size = len(data_sample) - train_size
      train_data, val_data = random_split(data_sample, [train_size, val_size])

[34]: train_loader = DataLoader(train_data, batch_size=BATCH_SIZE, shuffle=True)
      val_loader = DataLoader(val_data, batch_size=BATCH_SIZE)

[35]: num_epochs = 8

[36]: batch_size = BATCH_SIZE
      model_name = "distilgpt2"
      gpu = 0

[37]: criterion = nn.CrossEntropyLoss(ignore_index=tokenizer.pad_token_id) # Prevents
      ↪ the model from being penalized for predictions made on artificial padding
      ↪ tokens.
      optimizer = optim.AdamW(model.parameters(), lr=5e-4)
      tokenizer.pad_token = tokenizer.eos_token

[38]: results = pd.DataFrame(columns=['epoch', 'transformer', 'batch_size', 'gpu',
      ↪ 'training_loss', 'validation_loss', 'epoch_duration_sec'])

[39]: # Model Training Loop

      for epoch in range(num_epochs):
          start_time = time.time()
          model.train()
          epoch_training_loss = 0
          train_iterator = tqdm(train_loader, desc=f"Training Epoch {epoch+1}/
          ↪ {num_epochs} Batch Size: {batch_size}, Transformer: {model_name}")
          for batch in train_iterator:
              optimizer.zero_grad()
              inputs = batch['input_ids'].squeeze(1).to(device)
              targets = inputs.clone()
              outputs = model(input_ids=inputs, labels=targets)
              loss = outputs.loss
              loss.backward()
              optimizer.step()
              train_iterator.set_postfix({'Training Loss': loss.item()})
              epoch_training_loss += loss.item()
          average_epoch_training_loss = epoch_training_loss / len(train_iterator)
          model.eval()
          epoch_validation_loss = 0
          total_loss = 0
          valid_iterator = tqdm(val_loader, desc=f"Validation Epoch {epoch+1}/
          ↪ {num_epochs}")
          with torch.no_grad():

```

```

        for batch in valid_iterator:
            inputs = batch['input_ids'].squeeze(1).to(device)
            targets = inputs.clone()
            outputs = model(input_ids=inputs, labels=targets)
            loss = outputs.loss
            total_loss += loss
            valid_iterator.set_postfix({'Validation Loss': loss.item()})
            epoch_validation_loss += loss.item()
        average_epoch_validation_loss = epoch_validation_loss / len(valid_iterator)
        end_time = time.time()
        epoch_duration_sec = end_time - start_time

        new_row = {
            'transformer': model_name,
            'batch_size': batch_size,
            'gpu': gpu,
            'epoch': epoch + 1,
            'training_loss': average_epoch_training_loss,
            'validation_loss': average_epoch_validation_loss,
            'epoch_duration_sec': epoch_duration_sec
        }

        results.loc[len(results)] = new_row
        print(f"Epoch: {epoch+1}, Validation Loss: {total_loss/len(val_loader)}")

```

Training Epoch 1/8 Batch Size: 8, Transformer: distilgpt2: 0%| | 0/40  
 [00:00<?, ?it/s]`loss\_type=None` was set in the config but it is  
 unrecognised.Using the default loss: `ForCausalLMLoss`.

Training Epoch 1/8 Batch Size: 8, Transformer: distilgpt2: 100%| |  
 40/40 [00:08<00:00, 4.66it/s, Training Loss=0.943]  
 Validation Epoch 1/8: 100%| | 10/10 [00:00<00:00, 18.36it/s, Validation  
 Loss=0.755]

Epoch: 1, Validation Loss: 0.7456490397453308

Training Epoch 2/8 Batch Size: 8, Transformer: distilgpt2: 100%| |  
 40/40 [00:07<00:00, 5.29it/s, Training Loss=0.465]  
 Validation Epoch 2/8: 100%| | 10/10 [00:00<00:00, 18.69it/s, Validation  
 Loss=0.738]

Epoch: 2, Validation Loss: 0.7253366112709045

Training Epoch 3/8 Batch Size: 8, Transformer: distilgpt2: 100%| |  
 40/40 [00:07<00:00, 5.17it/s, Training Loss=0.345]  
 Validation Epoch 3/8: 100%| | 10/10 [00:00<00:00, 18.02it/s, Validation  
 Loss=0.736]

Epoch: 3, Validation Loss: 0.7382476329803467

Training Epoch 4/8 Batch Size: 8, Transformer: distilgpt2: 100%| |

```
40/40 [00:07<00:00, 5.04it/s, Training Loss=0.438]
Validation Epoch 4/8: 100%|      | 10/10 [00:00<00:00, 17.90it/s, Validation
Loss=0.793]
```

Epoch: 4, Validation Loss: 0.7701788544654846

```
Training Epoch 5/8 Batch Size: 8, Transformer: distilgpt2: 100%|      |
40/40 [00:08<00:00, 4.98it/s, Training Loss=0.283]
Validation Epoch 5/8: 100%|      | 10/10 [00:00<00:00, 17.33it/s, Validation
Loss=0.833]
```

Epoch: 5, Validation Loss: 0.8223657011985779

```
Training Epoch 6/8 Batch Size: 8, Transformer: distilgpt2: 100%|      |
40/40 [00:07<00:00, 5.01it/s, Training Loss=0.211]
Validation Epoch 6/8: 100%|      | 10/10 [00:00<00:00, 18.07it/s, Validation
Loss=0.881]
```

Epoch: 6, Validation Loss: 0.8576126098632812

```
Training Epoch 7/8 Batch Size: 8, Transformer: distilgpt2: 100%|      |
40/40 [00:07<00:00, 5.08it/s, Training Loss=0.161]
Validation Epoch 7/8: 100%|      | 10/10 [00:00<00:00, 18.25it/s, Validation
Loss=0.919]
```

Epoch: 7, Validation Loss: 0.9018468856811523

```
Training Epoch 8/8 Batch Size: 8, Transformer: distilgpt2: 100%|      |
40/40 [00:07<00:00, 5.16it/s, Training Loss=0.152]
Validation Epoch 8/8: 100%|      | 10/10 [00:00<00:00, 17.86it/s, Validation
Loss=0.965]
```

Epoch: 8, Validation Loss: 0.9539993405342102

```
[40]: input_str = "Kidney Stones"
```

```
[41]: input_ids = tokenizer.encode(input_str, return_tensors="pt").to(device)
```

```
[42]: input_ids
```

```
[42]: tensor([[48374, 1681, 26596]], device='cuda:0')
```

```
[48]: attention_mask = (input_ids != tokenizer.pad_token_id).long()
```

```
[49]: output = model.generate(
        input_ids=input_ids,
        attention_mask=attention_mask,
        max_length=20,
        num_return_sequences=1,
        do_sample=True,
```

```

top_k=8,
top_p=0.95,
temperature=0.5,
repetition_penalty=1.2,
pad_token_id=tokenizer.pad_token_id
)

```

```
[50]: output
```

```
[50]: tensor([[48374, 1681, 26596, 930, 1001, 4119, 32692, 2356, 11, 2910,
              287, 18922, 11, 10792, 2956, 1883, 50256]], device='cuda:0')
```

```
[51]: decoded_output = tokenizer.decode(output[0], skip_special_tokens=True)
```

```
[52]: decoded_output
```

```
[52]: 'Kidney Stones | Severe abdominal pain, blood in urine, frequent urination'
```

```
[55]: !huggingface-cli login
```

```

 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_
 _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_ _|_

```

A token is already saved on your machine. Run ``huggingface-cli whoami`` to get more information or ``huggingface-cli logout`` if you want to log out.

Setting a new token will erase the existing one.

To log in, ``huggingface_hub`` requires a token generated from <https://huggingface.co/settings/tokens>.

Enter your token (input will not be visible):

Add token as git credential? (Y/n) n

Token is valid (permission: fineGrained).

The token ``slm-gpt2distill-disease-symptoms`` has been saved to `/root/.cache/huggingface/stored_tokens`

Your token has been saved to `/root/.cache/huggingface/token`

Login successful.

The current active token is: ``slm-gpt2distill-disease-symptoms``

```
[56]: from huggingface_hub import notebook_login
      from transformers import AutoModelForCausalLM, AutoTokenizer

notebook_login()

repo_name = "aniketsalunkhe15/SLM-distilgpt2-disease-symptoms-predictor"

# Push model
model.push_to_hub(repo_name)

# Push tokenizer
tokenizer.push_to_hub(repo_name)
```

VBox(children=(HTML(value='<center> <img\nsrc=https://huggingface.co/front/  
assets/huggingface\_logo-noborder.svg...  
| 0.00/5.17k [00:00<?, ?B/s]

```
[56]: CommitInfo(commit_url='https://huggingface.co/aniketsalunkhe15/SLM-
      distilgpt2-disease-symptoms-
      predictor/commit/48a3bb5779a2aa6bb4e69f9704c0871d52de15ff',
      commit_message='Upload tokenizer', commit_description='',
      oid='48a3bb5779a2aa6bb4e69f9704c0871d52de15ff', pr_url=None,
      repo_url=RepoUrl('https://huggingface.co/aniketsalunkhe15/SLM-
      distilgpt2-disease-symptoms-predictor', endpoint='https://huggingface.co',
      repo_type='model', repo_id='aniketsalunkhe15/SLM-distilgpt2-disease-symptoms-
      predictor'), pr_revision=None, pr_num=None)
```

```
[59]: from transformers import AutoModelForCausalLM, AutoTokenizer

# Load model and tokenizer from Hugging Face Hub
repo_id = "aniketsalunkhe15/SLM-distilgpt2-disease-symptoms-predictor"

model = AutoModelForCausalLM.from_pretrained(repo_id)
tokenizer = AutoTokenizer.from_pretrained(repo_id)

# Prepare input
input_str = "Kidney Stones"
input_ids = tokenizer.encode(input_str, return_tensors="pt")

# Optional: attention mask
attention_mask = (input_ids != tokenizer.pad_token_id).long()

# Generate output
output = model.generate(
    input_ids=input_ids,
    attention_mask=attention_mask,
    max_length=20,
```

```
do_sample=True,
top_k=8,
top_p=0.95,
temperature=0.7,
repetition_penalty=1.2
)

# Decode result
decoded = tokenizer.decode(output[0], skip_special_tokens=True)
print(decoded)
```

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

Kidney Stones | Severe abdominal pain, blood in the side or back, blood in urine

```
[60]: %%javascript
      IPython.notebook.metadata.widgets = {};
```

<IPython.core.display.Javascript object>