```
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (2024.5.15)
       Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (2.31.0)
       Requirement already satisfied: safetensors>=0.4.1 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (0.4.3)
       Requirement already satisfied: tokenizers<0.20,>=0.19 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (0.19.1)
       Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (4.66.4)
       Requirement already satisfied: accelerate>=0.21.0 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (0.31.0)
       Requirement already satisfied: torch in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (2.3.1)
       Requirement already satisfied: psutil in /usr/local/lib/python3.10/dist-packages (from accelerate>=0.21.0->transformers[torch]) (5.9.5)
       Requirement already satisfied: fsspec>-2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub<1.0,>=0.23.2->transform Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub<1.0,>=0.23.2->transfor
       Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (1.12.1)
       Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (3.3)
       Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (3.1.4)
       Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[tor
       Requirement already satisfied: nvidia-cuda-runtime-cu12=-12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[t
       Requirement already satisfied: nvidia-cuda-cupti-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[tor
       Requirement already satisfied: nvidia-cudnn-cu12==8.9.2.26 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch])
       Requirement already satisfied: nvidia-cublas-cu12==12.1.3.1 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch])
       Requirement already satisfied: nvidia-cufft-cu12==11.0.2.54 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch])
       Requirement already satisfied: nvidia-curand-cu12==10.3.2.106 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch
       Requirement already satisfied: nvidia-cusolver-cu12==11.4.5.107 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[tor
       Requirement already satisfied: nvidia-cusparse-cu12==12.1.0.106 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[tor
       Requirement already satisfied: nvidia-nccl-cu12==2.20.5 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (2. Requirement already satisfied: nvidia-nvtx-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (
       Requirement already satisfied: triton==2.3.1 in /usr/local/lib/python3.10/dist-packages (from torch->transformers[torch]) (2.3.1)
       Requirement already satisfied: nvidia-nvjitlink-cu12 in /usr/local/lib/python3.10/dist-packages (from nvidia-cusolver-cu12==11.4.5.107-)
       Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->transformers[torch])
       Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->transformers[torch]) (3.7)
       Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->transformers[torch]) (2.0.7 Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->transformers[torch]) (2024.
       Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch->transformers[torch]) (2.1
       Requirement already satisfied: mpmath<1.4.0,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy->torch->transformers[torch])
       Installing collected packages: transformers
          Attempting uninstall: transformers
             Found existing installation: transformers 4.41.2
             Uninstalling transformers-4.41.2:
               Successfully uninstalled transformers-4.41.2
       Successfully installed transformers-4.42.3
import pandas as pd
import numpy as np
import torch
from transformers import AutoTokenizer, AutoModelForSequenceClassification, Trainer, TrainingArguments
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
from sklearn.preprocessing import LabelEncoder
# Load and preprocess the data
data = pd.read_csv('/content/drive/MyDrive/Naive/training_data (2).csv')
# Check the column names
print(data.columns)
# Combine symptom columns into a single text column
data['text'] = data[['Symptom_1', 'Symptom_2', 'Symptom_3']].apply(lambda x: ' '.join(x.dropna().astype(str)), axis=1)
# Encode the labels to integers
label encoder = LabelEncoder()
data['label'] = label_encoder.fit_transform(data['Disease'])
# Extract features and labels
X = data['text'].tolist()
y = data['label'].tolist()
# Split the data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# Load BioBERT tokenizer and model
tokenizer = AutoTokenizer.from_pretrained("dmis-lab/biobert-base-cased-v1.1")
model = AutoModelForSequenceClassification.from_pretrained("dmis-lab/biobert-base-cased-v1.1", num_labels=len(set(y)))
# Tokenize the data
train encodings = tokenizer(X train, truncation=True, padding=True, max length=512)
test_encodings = tokenizer(X_test, truncation=True, padding=True, max_length=512)
# Convert to torch Dataset
class Dataset(torch.utils.data.Dataset):
     def __init__(self, encodings, labels):
           self.encodings = encodings
```

```
self.labels = labels
     def __getitem__(self, idx):
           item = {key: torch.tensor(val[idx]) for key, val in self.encodings.items()}
          item['labels'] = torch.tensor(self.labels[idx])
          return item
     def __len__(self):
           return len(self.labels)
train_dataset = Dataset(train_encodings, y_train)
test_dataset = Dataset(test_encodings, y_test)
# Define training arguments
training_args = TrainingArguments(
   output_dir='./results',
                                      # output directory
    num_train_epochs=3, # output directory
num_train_epochs=3, # number of training epochs
per_device_train_batch_size=8, # batch size for training
per_device_eval_batch_size=8, # batch size for evaluation
warmup_steps=500, # number of warmup steps for learning rate scheduler
weight_decay=0.01, # strength of weight decay
logging_dir='./logs', # directory for storing logs
logging_steps=10.
     logging_steps=10,
# Define the Trainer
trainer = Trainer(
     model=model,
                                                   # the instantiated @ Transformers model to be trained
    model=model,
args=training_args,
train_dataset=train_dataset,
eval_dataset=test_dataset
# training arguments, defined above
# training dataset
evaluation dataset
# Train the model
trainer.train()
# Evaluate the model
preds_output = trainer.predict(test_dataset)
preds = np.argmax(preds_output.predictions, axis=1)
accuracy = accuracy_score(y_test, preds)
precision = precision_score(y_test, preds, average='weighted')
recall = recall_score(y_test, preds, average='weighted')
f1 = f1_score(y_test, preds, average='weighted')
print(f'Accuracy: {accuracy}')
print(f'Precision: {precision}')
print(f'Recall: {recall}')
print(f'F1 Score: {f1}')
```

Index(['Disease', 'Symptom\_1', 'Symptom\_2', 'Symptom\_3'], dtype='object')

Some weights of BertForSequenceClassification were not initialized from the model checkpy you should probably TRAIN this model on a down-stream task to be able to use it for preceived.

You	sho	ould probably	TRAIN	this							sk to Epoch		to
Ste	ер	Training Loss				[141	10/14	110	1.07	.50,	Еросі	13/3]	
	10	3.808600											
	20	3.713200											
,	30	3.695500											
	40	3.717400											
	50	3.671900											
	60	3.663800											
-	70	3.619800											
	80	3.646000											
9	90	3.575000											
10	00	3.518300											
1	10	3.467000											
13	20	3.433600											
1:	30	3.313800											
14	40	3.211600											
1	50	3.221900											
16	60	3.147200											
1	70	2.965100											
18	80	2.834800											
15	90	2.870000											
20	00	2.719600											
2	10	2.579800											
2	20	2.512300											
2	30	2.388800											
2	40	2.278400											
2	50	2.130500											
20	60	2.125800											
2	70	2.077000											
28	80	1.823200											
29	90	1.749600	Ę.										
30	00	1.588200											
3	10	1.574800											
32	20	1.395000											
33	30	1.251200											
34	40	1.274100											
3	50	1.065500											
36	60	1.001300											
3	70	0.935200											
38	80	0.829100											
35	90	0.789700											
40	00	0.876500											
4	10	0.651300											
43	20	0.674500											
	30	0.575700											
0.0													

0.195500

0.277600 0.309100

880 890

900

BioBERT\_Naive\_Bayes\_Code\_final.ipynb - Colab

910	0.073200
920	0.291300
930	0.216300
940	0.071300
950	0.085900
960	0.114000
970	0.106100
980	0.197100
990	0.063800
1000	0.132300
1010	0.097800
1020	0.079100
1030	0.106800
1040	0.103500
1050	0.159400
1060	0.109800
1070	0.182400
1080	0.074300
1090	0.086600
1100	0.146100
1110	0.196800
1120	0.132600
1130	0.095700
1140	0.102100
1150	0.280100
1160	0.070800
1170	
1180	
1190	
1200	
1210	
1220	
1230	
1240	
1250	
1260	
1270	
1280	
1290	
1300	
1310 1320	
1320	
1340	
1350	
1360	
1370	
maca.	www.P.F.J.S.FV

```
import matplotlib.pyplot as plt
# Extract training loss from log_history
train_losses = [item['loss'] for item in trainer.state.log_history]
# Plot eval F1 score
plt.figure()
plt.plot(trainer_state['eval_f1s'], label='Eval F1 Score')
plt.xlabel('Epoch')
plt.ylabel('F1 Score')
plt.title('Evaluation F1 Score')
plt.legend()
plt.show()
Show hidden output
nlt figure/figsize=(10 5))
# Plot the training loss
plt.figure(figsize=(10, 5))
# Extract loss values from log_history
loss_values = [entry['loss'] for entry in trainer.state.log_history if 'loss' in entry]
plt.plot(loss_values, label='Training Loss')
plt.xlabel('Steps')
plt.ylabel('Loss')
plt.title('Training Loss')
plt.legend()
plt.show()
Show hidden output
par. ..... , =poe... ,
n_classes = len(list(data.Disease.unique()))
n_classes
<del>→</del> 41
plt.subplot(2, 2, 3)
def get_model_metrics():
    model = GaussianNB()
    model.fit(X_train, y_train)
    # Make predictions on the test set
    y_pred = model.predict(X_test)
    # Calculate precision, recall, and F1 score
    precision = precision_score(y_test, y_pred, average='weighted')
    recall = recall_score(y_test, y_pred, average='weighted')
    f1 = f1_score(y_test, y_pred, average='weighted')
    accuracy = accuracy_score(y_test, y_pred)
```

```
from google.colab import drive
drive.mount('/content/drive')
Fr Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
pip install accelerate -U
Requirement already satisfied: accelerate in /usr/local/lib/python3.10/dist-packages (0.31.0)
     Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from accelerate) (1.25.2)
     Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from accelerate) (24.1)
     Requirement already satisfied: psutil in /usr/local/lib/python3.10/dist-packages (from accelerate) (5.9.5)
     Requirement already satisfied: pyyaml in /usr/local/lib/python3.10/dist-packages (from accelerate) (6.0.1)
     Requirement already satisfied: torch>=1.10.0 in /usr/local/lib/python3.10/dist-packages (from accelerate) (2.3.1)
     Requirement already satisfied: huggingface-hub in /usr/local/lib/python3.10/dist-packages (from accelerate) (0.23.4)
     Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.10/dist-packages (from accelerate) (0.4.3)
     Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (3.15.4)
     Requirement already satisfied: typing-extensions>=4.8.0 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (4.1
     Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (1.12.1)
     Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (3.3)
     Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (3.1.4)
     Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (2023.6.0)
     Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelera
     Requirement already satisfied: nvidia-cuda-runtime-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accele
     Requirement already satisfied: nvidia-cuda-cupti-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelera
     Requirement already satisfied: nvidia-cudnn-cu12==8.9.2.26 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (
     Requirement already satisfied: nvidia-cublas-cu12==12.1.3.1 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate)
     Requirement already satisfied: nvidia-cufft-cu12==11.0.2.54 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate)
     Requirement already satisfied: nvidia-curand-cu12==10.3.2.106 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate
     Requirement already satisfied: nvidia-cusolver-cu12==11.4.5.107 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->acceleration
     Requirement already satisfied: nvidia-cusparse-cu12==12.1.0.106 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelera
     Requirement already satisfied: nvidia-nccl-cu12==2.20.5 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (2.2
     Requirement already satisfied: nvidia-nvtx-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (1
     Requirement already satisfied: triton==2.3.1 in /usr/local/lib/python3.10/dist-packages (from torch>=1.10.0->accelerate) (2.3.1)
     Requirement already satisfied: nvidia-nvjitlink-cu12 in /usr/local/lib/python3.10/dist-packages (from nvidia-cusolver-cu12==11.4.5.107-)
     Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from huggingface-hub->accelerate) (2.31.0)
     Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub->accelerate) (4.66.4)
     Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch>=1.10.0->accelerate) (2.1.
     Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->huggingface-hub->acce
     Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->huggingface-hub->accelerate) (3.7
     Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->huggingface-hub->accelerate
     Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->huggingface-hub->accelerate
     Requirement already satisfied: mpmath<1.4.0,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy->torch>=1.10.0->accelerate) (
!pip install torch -U
Requirement already satisfied: torch in /usr/local/lib/python3.10/dist-packages (2.3.1)
     Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch) (3.15.4)
     Requirement already satisfied: typing-extensions>=4.8.0 in /usr/local/lib/python3.10/dist-packages (from torch) (4.12.2)
     Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch) (1.12.1)
     Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch) (3.3)
     Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch) (3.1.4)
     Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from torch) (2023.6.0)
     Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.105)
     Requirement already satisfied: nvidia-cuda-runtime-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.105)
     Requirement already satisfied: nvidia-cuda-cupti-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.105)
     Requirement already satisfied: nvidia-cudnn-cu12==8.9.2.26 in /usr/local/lib/python3.10/dist-packages (from torch) (8.9.2.26)
     Requirement already satisfied: nvidia-cublas-cu12==12.1.3.1 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.3.1)
     Requirement already satisfied: nvidia-cufft-cu12==11.0.2.54 in /usr/local/lib/python3.10/dist-packages (from torch) (11.0.2.54)
     Requirement already satisfied: nvidia-curand-cu12==10.3.2.106 in /usr/local/lib/python3.10/dist-packages (from torch) (10.3.2.106)
     Requirement already satisfied: nvidia-cusolver-cu12==11.4.5.107 in /usr/local/lib/python3.10/dist-packages (from torch) (11.4.5.107)
     Requirement already satisfied: nvidia-cusparse-cu12==12.1.0.106 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.0.106)
     Requirement already satisfied: nvidia-nccl-cu12==2.20.5 in /usr/local/lib/python3.10/dist-packages (from torch) (2.20.5)
     Requirement already satisfied: nvidia-nvtx-cu12==12.1.105 in /usr/local/lib/python3.10/dist-packages (from torch) (12.1.105)
     Requirement already satisfied: triton==2.3.1 in /usr/local/lib/python3.10/dist-packages (from torch) (2.3.1)
     Requirement already satisfied: nvidia-nvjitlink-cu12 in /usr/local/lib/python3.10/dist-packages (from nvidia-cusolver-cu12==11.4.5.107->
     Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch) (2.1.5)
     Requirement already satisfied: mpmath<1.4.0,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy->torch) (1.3.0)
!pip install transformers[torch] -U
Requirement already satisfied: transformers[torch] in /usr/local/lib/python3.10/dist-packages (4.41.2)
     Collecting transformers[torch]
       Using cached transformers-4.42.3-py3-none-any.whl (9.3 MB)
     Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (3.15.4)
     Requirement already satisfied: huggingface-hub<1.0,>=0.23.2 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (0.23.
     Requirement already satisfied: numpy<2.0,>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (1.25.2)
     Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (24.1)
     Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers[torch]) (6.0.1)
```