

```

!pip install datasets
!pip install transformers

Requirement already satisfied: pyarrow>=12.0.0 in /usr/local/lib/python3.10/dist-packages (from datasets) (14.0.2)
Requirement already satisfied: pyarrow-hotfix in /usr/local/lib/python3.10/dist-packages (from datasets) (0.6)
Collecting dill<0.3.9,>=0.3.0 (from datasets)
  Downloading dill-0.3.8-py3-none-any.whl (116 kB)
    116.3/116.3 kB 6.9 MB/s eta 0:00:00
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from datasets) (2.0.3)
Requirement already satisfied: requests>=2.19.0 in /usr/local/lib/python3.10/dist-packages (from datasets) (2.31.0)
Requirement already satisfied: tqdm>=4.62.1 in /usr/local/lib/python3.10/dist-packages (from datasets) (4.66.2)
Collecting xxhash (from datasets)
  Downloading xxhash-3.4.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (194 kB)
    194.1/194.1 kB 8.2 MB/s eta 0:00:00
Collecting multiprocessing (from datasets)
  Downloading multiprocessing-0.70.16-py310-none-any.whl (134 kB)
    134.8/134.8 kB 8.4 MB/s eta 0:00:00
Requirement already satisfied: fsspec[http]<=2024.3.1,>=2023.1.0 in /usr/local/lib/python3.10/dist-packages (from datasets) (2024.3.1)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.10/dist-packages (from datasets) (3.9.5)
Collecting huggingface-hub>=0.21.2 (from datasets)
  Downloading huggingface_hub-0.22.2-py3-none-any.whl (388 kB)
    388.9/388.9 kB 12.0 MB/s eta 0:00:00
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from datasets) (24.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from datasets) (6.0.1)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (23.2.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (6.0.5)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (1.9.7)
Requirement already satisfied: async-timeout<5.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp->datasets) (4.0.3)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub) (4.11.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests>=2.19.0->datasets) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests>=2.19.0->datasets) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests>=2.19.0->datasets) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests>=2.19.0->datasets) (2024.2.2)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas->datasets) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->datasets) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->datasets) (2024.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: xxhash, dill, multiprocessing, huggingface-hub, datasets
  Attempting uninstall: huggingface-hub
    Found existing installation: huggingface-hub 0.20.3
    Uninstalling huggingface-hub-0.20.3:
      Successfully uninstalled huggingface-hub-0.20.3
Successfully installed datasets-2.19.0 dill-0.3.8 huggingface-hub-0.22.2 multiprocessing-0.70.16 xxhash-3.4.1
Requirement already satisfied: transformers in /usr/local/lib/python3.10/dist-packages (4.40.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers) (3.13.4)
Requirement already satisfied: huggingface-hub<1.0,>=0.19.3 in /usr/local/lib/python3.10/dist-packages (from transformers) (0.22.2)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers) (1.25.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers) (24.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers) (6.0.1)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers) (2023.12.1)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers) (2.31.0)
Requirement already satisfied: tokenizers<0.20,>=0.19 in /usr/local/lib/python3.10/dist-packages (from transformers) (0.19.1)
Requirement already satisfied: safetensors>=0.4.1 in /usr/local/lib/python3.10/dist-packages (from transformers) (0.4.3)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers) (4.66.2)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub<1.0,>=0.19.3->transformers) (2024.3.1)
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->transformers) (4.11.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->transformers) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->transformers) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->transformers) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->transformers) (2024.2.2)

from google.colab import drive

# Mount Google Drive
drive.mount('/content/drive')

Mounted at /content/drive

!pip install rouge

Collecting rouge
  Downloading rouge-1.0.1-py3-none-any.whl (13 kB)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from rouge) (1.16.0)
Installing collected packages: rouge
Successfully installed rouge-1.0.1

from datasets import load_dataset
from transformers import BartForConditionalGeneration, BartTokenizer
from rouge import Rouge
import random

# Load test dataset
test_dataset = load_dataset("code_x_glue_ct_code_to_text", "python", split="test")

```

```

# Shuffle the dataset
test_dataset = test_dataset.shuffle(seed=42)

# Select the first 1000 data points
test_dataset = test_dataset.select(range(100))

# Load fine-tuned model and tokenizer
model = BartForConditionalGeneration.from_pretrained("/content/drive/MyDrive/finetuned_model")
tokenizer = BartTokenizer.from_pretrained("facebook/bart-base")

# Initialize Rouge
rouge = Rouge()

# Evaluate on the test dataset
references = []
predictions = []

for example in test_dataset:
    code = example["code"]
    docstring = example["docstring"]

    # Tokenize input for this particular example
    inputs = tokenizer(f"code: {code}", return_tensors="pt", max_length=512, truncation=True)

    # Generate summary
    summary_ids = model.generate(inputs.input_ids.to(model.device), num_beams=4, max_length=128, early_stopping=True)
    summary = tokenizer.decode(summary_ids[0], skip_special_tokens=True)

    # Append reference and prediction for ROUGE calculation
    references.append(docstring)
    predictions.append(summary)

# Calculate ROUGE scores
rouge_scores = rouge.get_scores(predictions, references, avg=True)

# Format and print ROUGE scores
formatted_scores = "\n".join([f"{metric}: {score}" for metric, score in rouge_scores.items()])
print("ROUGE scores:")
print(formatted_scores)

ROUGE scores:
rouge-1: {'r': 0.658703392198503, 'p': 0.9824118011618013, 'f': 0.72996800280963}
rouge-2: {'r': 0.6242197470439058, 'p': 0.9766361273135468, 'f': 0.6907215548955585}
rouge-l: {'r': 0.658703392198503, 'p': 0.9824118011618013, 'f': 0.72996800280963}

from datasets import load_dataset
from transformers import BartForConditionalGeneration, BartTokenizer
from rouge import Rouge
import random

# Load test dataset
test_dataset = load_dataset("code_search_net", "python", split="test")

# Shuffle the dataset
test_dataset = test_dataset.shuffle(seed=42)

# Select the first 1000 data points
test_dataset = test_dataset.select(range(100))

# Load fine-tuned model and tokenizer
model = BartForConditionalGeneration.from_pretrained("/content/drive/MyDrive/finetuned_model")
tokenizer = BartTokenizer.from_pretrained("facebook/bart-base")

# Initialize Rouge
rouge = Rouge()

# Evaluate on the test dataset
references = []
predictions = []

for example in test_dataset:
    code = example["func_code_string"]
    docstring = example["func_documentation_string"]

    # Tokenize input for this particular example
    inputs = tokenizer(f"code: {code}", return_tensors="pt", max_length=512, truncation=True)

    # Generate summary
    summary_ids = model.generate(inputs.input_ids.to(model.device), num_beams=4, max_length=128, early_stopping=True)
    summary = tokenizer.decode(summary_ids[0], skip_special_tokens=True)

    # Append reference and prediction for ROUGE calculation

```

```
references.append(docstring)
predictions.append(summary)

# Calculate ROUGE scores
rouge_scores = rouge.get_scores(predictions, references, avg=True)

# Format and print ROUGE scores
formatted_scores = "\n".join([f"{metric}: {score}" for metric, score in rouge_scores.items()])
print("ROUGE scores:")
print(formatted_scores)

ROUGE scores:
rouge-1: {'r': 0.6238417906335681, 'p': 0.9763567364534065, 'f': 0.697041148411199}
rouge-2: {'r': 0.5893947903515329, 'p': 0.968195302289356, 'f': 0.6583109449495957}
rouge-l: {'r': 0.6238417906335681, 'p': 0.9763567364534065, 'f': 0.697041148411199}
```