

```
# =====
# 🚀 Install Dependencies
# =====
!pip install transformers datasets rouge-score

import torch
from transformers import T5Tokenizer, T5ForConditionalGeneration, Trainer, TrainingArguments
from datasets import load_dataset

# Check if GPU is available
device = "cuda" if torch.cuda.is_available() else "cpu"
print(f"Using device: {device}")

# Preparing metadata (setup.py) ... done
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) (0.28.1)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (1.26.4)
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) (0.21.0)
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: 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]<=2024.12.0) (2024.12.0)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-packages (from datasets) (3.11.13)
Requirement already satisfied: absl-py in /usr/local/lib/python3.11/dist-packages (from rouge-score) (1.4.0)
Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-packages (from rouge-score) (3.9.1)
Requirement already satisfied: six>=1.14.0 in /usr/local/lib/python3.11/dist-packages (from rouge-score) (1.17.0)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (2.4.6)
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) (25.1.0)
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<1.0,>=0.24.0) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.4.0)
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) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2025.1.1)
Requirement already satisfied: click in /usr/local/lib/python3.11/dist-packages (from nltk->rouge-score) (8.1.8)
Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (from nltk->rouge-score) (1.4.2)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2025.1)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2025.1)
Downloading datasets-3.3.2-py3-none-any.whl (485 kB)
  485.4/485.4 kB 4.8 MB/s eta 0:00:00
Downloading dill-0.3.8-py3-none-any.whl (116 kB)
  116.3/116.3 kB 3.7 MB/s eta 0:00:00
Downloading multiprocess-0.70.16-py311-none-any.whl (143 kB)
  143.5/143.5 kB 4.4 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 5.5 MB/s eta 0:00:00
Building wheels for collected packages: rouge-score
  Building wheel for rouge-score (setup.py) ... done
  Created wheel for rouge-score: filename=rouge_score-0.1.2-py3-none-any.whl size=24935 sha256=7ee75b9a94d41d786a9a7fc0ca2d09122e
  Stored in directory: /root/.cache/pip/wheels/1e/19/43/8a442dc83660ca25e163e1bd1f89919284ab0d0c1475475148
Successfully built rouge-score
Installing collected packages: xxhash, dill, rouge-score, multiprocess, datasets
Successfully installed datasets-3.3.2 dill-0.3.8 multiprocess-0.70.16 rouge-score-0.1.2 xxhash-3.5.0
Using device: cuda

# =====
# 🚀 Load Pretrained T5 Model (Baseline)
# =====
model_name = "t5-small" # You can also try "t5-base" or "t5-large"
tokenizer = T5Tokenizer.from_pretrained(model_name)
model_before = T5ForConditionalGeneration.from_pretrained(model_name).to(device) # Keep a copy of the base model
```

```

/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as :
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
warnings.warn(
tokenizer_config.json: 100%                2.32k/2.32k [00:00<00:00, 55.7kB/s]

spiece.model: 100%                        792k/792k [00:00<00:00, 4.57MB/s]

tokenizer.json: 100%                      1.39M/1.39M [00:00<00:00, 6.42MB/s]
You are using the default legacy behaviour of the <class 'transformers.models.t5.tokenization_t5.T5Tokenizer'>. This is expected, ar
config.json: 100%                        1.21k/1.21k [00:00<00:00, 40.0kB/s]

model.safetensors: 100%                  242M/242M [00:01<00:00, 189MB/s]

generation_config.json: 100%             147/147 [00:00<00:00, 3.32kB/s]

```

Start coding or [generate](#) with AI.

```

# =====
# 🚀 Load and Prepare XSum Dataset
# =====
dataset = load_dataset("xsum")
train_dataset = dataset["train"].select(range(2000)) # Small subset for faster training
val_dataset = dataset["validation"].select(range(500))

# Preprocessing function
def preprocess_function(examples):
    inputs = ["summarize: " + doc for doc in examples["document"]]
    model_inputs = tokenizer(inputs, padding="max_length", truncation=True, max_length=512)

    labels = tokenizer(examples["summary"], padding="max_length", truncation=True, max_length=150)
    model_inputs["labels"] = labels["input_ids"]
    return model_inputs

```

```

# Tokenize datasets
tokenized_train = train_dataset.map(preprocess_function, batched=True)
tokenized_val = val_dataset.map(preprocess_function, batched=True)

# Ensure dataset only has required features
tokenized_train.set_format(type="torch", columns=["input_ids", "attention_mask", "labels"])
tokenized_val.set_format(type="torch", columns=["input_ids", "attention_mask", "labels"])

```

```

README.md: 100%                6.24k/6.24k [00:00<00:00, 664kB/s]

xsum.py: 100%                   5.76k/5.76k [00:00<00:00, 572kB/s]
The repository for xsum contains custom code which must be executed to correctly load the dataset. You can inspect the repository cc
You can avoid this prompt in future by passing the argument `trust_remote_code=True`.

Do you wish to run the custom code? [y/N] y
(...)SUM-EMNLP18-Summary-Data-Original.tar.gz: 100%                255M/255M [00:01<00:00, 239MB/s]

Downloading data: 100%          2.72M/2.72M [00:00<00:00, 43.5MB/s]

Generating train split: 100%     204045/204045 [00:38<00:00, 6003.29 examples/s]

Generating validation split: 100% 11332/11332 [00:20<00:00, 412.50 examples/s]

Generating test split: 100%      11334/11334 [00:19<00:00, 611.97 examples/s]

Map: 100%                       2000/2000 [00:04<00:00, 395.00 examples/s]

Map: 100%                       500/500 [00:01<00:00, 401.14 examples/s]

```

```

# =====
# 🚀 Fine-Tuning T5 on XSum
# =====
model_after = T5ForConditionalGeneration.from_pretrained(model_name).to(device) # Separate model instance

training_args = TrainingArguments(
    output_dir="./t5_summarization",
    evaluation_strategy="epoch",
    learning_rate=5e-5,
    per_device_train_batch_size=4,
    per_device_eval_batch_size=4,
    num_train_epochs=5,
    weight_decay=0.01,
    save_strategy="epoch",

```

```

logging_dir="./logs",
logging_steps=10,
fp16=True if torch.cuda.is_available() else False
)

trainer = Trainer(
    model=model_after, # Fine-tuning a new model, keeping model_before unchanged
    args=training_args,
    train_dataset=tokenized_train,
    eval_dataset=tokenized_val,
    tokenizer=tokenizer,
)

trainer.train()

# Save fine-tuned model properly
trainer.model.save_pretrained("./fine_tuned_t5_xsum")
tokenizer.save_pretrained("./fine_tuned_t5_xsum")

print("\n===== Fine-Tuning Completed =====\n")

/usr/local/lib/python3.11/dist-packages/transformers/training_args.py:1575: FutureWarning: `evaluation_strategy` is deprecated and v
warnings.warn(
<ipython-input-4-8d5fb438ba31>:20: FutureWarning: `tokenizer` is deprecated and will be removed in version 5.0.0 for `Trainer.__init
trainer = Trainer(
wandb: WARNING The `run_name` is currently set to the same value as `TrainingArguments.output_dir`. If this was not intended, please
wandb: Using wandb-core as the SDK backend. Please refer to https://wandb.me/wandb-core for more information.
wandb: Logging into wandb.ai. (Learn how to deploy a W&B server locally: https://wandb.me/wandb-server)
wandb: You can find your API key in your browser here: https://wandb.ai/authorize
wandb: Paste an API key from your profile and hit enter: .....
wandb: WARNING If you're specifying your api key in code, ensure this code is not shared publicly.
wandb: WARNING Consider setting the WANDB_API_KEY environment variable, or running `wandb login` from the command line.
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: kanamarlapudi21bec18 (kanamarlapudi21bec18-iiitkottayam) to https://api.wandb.ai. Use `wandb login -
Tracking run with wandb version 0.19.7
Run data is saved locally in /content/wandb/run-20250304_174718-tc8pnxmh
Syncing run ./t5\_summarization to Weights & Biases \(docs\)
View project at https://wandb.ai/kanamarlapudi21bec18-iiitkottayam/huggingface
View run at https://wandb.ai/kanamarlapudi21bec18-iiitkottayam/huggingface/runs/tc8pnxmh
Passing a tuple of `past_key_values` is deprecated and will be removed in Transformers v4.48.0. You should pass an instance of `Enc
[2500/2500 06:04, Epoch 5/5]

Epoch  Training Loss  Validation Loss
-----
1         0.700500         0.552603
2         0.586000         0.542247
3         0.579800         0.539270
4         0.535800         0.537782
5         0.580000         0.537592

===== Fine-Tuning Completed =====

# =====
# 🚀 Load Fine-Tuned Model Separately
# =====
model_after = T5ForConditionalGeneration.from_pretrained("./fine_tuned_t5_xsum").to(device)

# =====
# 🚀 Functions to Generate Summaries
# =====

# Function for the base (pretrained) model
def generate_summary_before(article, max_input_length=512, max_output_tokens=150):
    inputs = tokenizer("summarize: " + article, return_tensors="pt", truncation=True, max_length=max_input_length).to(device)

    with torch.no_grad(): # Disable gradient calculation for inference
        output = model_before.generate(
            **inputs,
            max_new_tokens=max_output_tokens,
            no_repeat_ngram_size=2,
            early_stopping=True,
            pad_token_id=tokenizer.pad_token_id # 🚀 Fix padding issue
        )

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

```

```

# Function for the fine-tuned model
def generate_summary_after(article, max_input_length=512, max_output_tokens=150):
    inputs = tokenizer("summarize: " + article, return_tensors="pt", truncation=True, max_length=max_input_length).to(device)

    with torch.no_grad():
        output = model_after.generate(
            **inputs,
            max_new_tokens=max_output_tokens,
            no_repeat_ngram_size=2,
            early_stopping=True,
            pad_token_id=tokenizer.pad_token_id # 🚀 Fix padding issue
        )

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

# =====
# 🚀 Compare Summaries Before and After Fine-Tuning
# =====
test_samples = dataset["test"].select(range(5)) # Pick some test examples

for i, example in enumerate(test_samples):
    article = example["document"]
    reference_summary = example["summary"]

    summary_before = generate_summary_before(article)
    summary_after = generate_summary_after(article)

    print(f"\n===== Example {i+1} =====")
    print("\n--- Original Article ---\n", article[:500], "...")
    print("\n--- Reference Summary ---\n", reference_summary)
    print("\n--- T5 Generated Summary (Before Fine-Tuning) ---\n", summary_before)
    print("\n--- T5 Generated Summary (After Fine-Tuning) ---\n", summary_after)

🔄 /usr/local/lib/python3.11/dist-packages/transformers/generation/configuration_utils.py:676: UserWarning: `num_beams` is set to 1.
  warnings.warn(

===== Example 1 =====

--- Original Article ---
Prison Link Cymru had 1,099 referrals in 2015-16 and said some ex-offenders were living rough for up to a year before finding su
Workers at the charity claim investment in housing would be cheaper than jailing homeless repeat offenders.
The Welsh Government said more people than ever were getting help to address housing problems.
Changes to the Housing Act in Wales, introduced in 2015, removed the right for prison leavers to be given priority for accommodat
Prison Link C ...

--- Reference Summary ---
There is a "chronic" need for more housing for prison leavers in Wales, according to a charity.

--- T5 Generated Summary (Before Fine-Tuning) ---
prison link Cymru says some ex-offenders were living rough for up to a year. the charity says it is cheaper than jailing homeless

--- T5 Generated Summary (After Fine-Tuning) ---
The Welsh Housing Authority has said it is "desperate" for people to find accommodation after their release.

===== Example 2 =====

--- Original Article ---
Officers searched properties in the Waterfront Park and Colonsay View areas of the city on Wednesday.
Detectives said three firearms, ammunition and a five-figure sum of money were recovered.
A 26-year-old man who was arrested and charged appeared at Edinburgh Sheriff Court on Thursday. ...

--- Reference Summary ---
A man has appeared in court after firearms, ammunition and cash were seized by police in Edinburgh.

--- T5 Generated Summary (Before Fine-Tuning) ---
police say three firearms, ammunition and a five-figure sum of money were recovered. he appeared at Edinburgh Sheriff Court on f

--- T5 Generated Summary (After Fine-Tuning) ---
A man has been arrested and charged with firearms offences in Edinburgh.

===== Example 3 =====

--- Original Article ---
Jordan Hill, Brittany Covington and Tesfaye Cooper, all 18, and Tanishia Covington, 24, appeared in a Chicago court on Friday.
The four have been charged with hate crimes and aggravated kidnapping and battery, among other things.
An online fundraiser for their victim has collected $51,000 (£42,500) so far.
Denying the four suspects bail, Judge Maria Kuriakos Ciesil asked: "Where was your sense of decency?"
Prosecutors told the court the beating started in a van and continued at a house, where t ...

--- Reference Summary ---
Four people accused of kidnapping and torturing a mentally disabled man in a "racially motivated" attack streamed on Facebook ha

--- T5 Generated Summary (Before Fine-Tuning) ---

```

the four suspects have been charged with hate crimes and aggravated kidnapping. the victim, 18, suffers from schizophrenia and a

--- T5 Generated Summary (After Fine-Tuning) ---

A man has been charged with hate crimes and battery for beating a white man in solitary confinement.

===== Example 4 =====

!pip install evaluate

```
Collecting evaluate
  Downloading evaluate-0.4.3-py3-none-any.whl.metadata (9.2 kB)
Requirement already satisfied: datasets>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from evaluate) (3.3.2)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from evaluate) (1.26.4)
Requirement already satisfied: dill in /usr/local/lib/python3.11/dist-packages (from evaluate) (0.3.8)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (from evaluate) (2.2.2)
Requirement already satisfied: requests>=2.19.0 in /usr/local/lib/python3.11/dist-packages (from evaluate) (2.32.3)
Requirement already satisfied: tqdm>=4.62.1 in /usr/local/lib/python3.11/dist-packages (from evaluate) (4.67.1)
Requirement already satisfied: xxhash in /usr/local/lib/python3.11/dist-packages (from evaluate) (3.5.0)
Requirement already satisfied: multiprocessing in /usr/local/lib/python3.11/dist-packages (from evaluate) (0.70.16)
Requirement already satisfied: fsspec>=2021.05.0 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]>=2021.05.0->evaluate) (2025.1)
Requirement already satisfied: huggingface-hub>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (0.28.1)
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (from evaluate) (24.2)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (3.17.0)
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (18.1.0)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (3.11.13)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (6.0.2)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.7.0->evaluate) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->evaluate) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->evaluate) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->evaluate) (2.3.1)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->evaluate) (2025.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate) (2025.1)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate) (2025.1)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (2.4.4)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (25.1.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (6.1.0)
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (0.2.0)
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets>=2.0.0->evaluate) (1.18.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas->evaluate) (1.17.0)
Downloading evaluate-0.4.3-py3-none-any.whl (84 kB)
84.0/84.0 kB 3.8 MB/s eta 0:00:00
Installing collected packages: evaluate
Successfully installed evaluate-0.4.3
```

```
import evaluate # New library for evaluation metrics

# Load ROUGE metric
rouge = evaluate.load("rouge")

def compute_rouge_scores(reference_summaries, generated_summaries):
    scores = rouge.compute(predictions=generated_summaries, references=reference_summaries)
    # Access scores directly, removing the .mid attribute
    return {
        "ROUGE-1": scores["rouge1"],
        "ROUGE-2": scores["rouge2"],
        "ROUGE-L": scores["rougeL"]
    }

# =====
# Summaries & Computed ROUGE Scores
# =====

test_samples = dataset["test"].select(range(400)) # Use 400 test samples for evaluation

# Store summaries
reference_summaries = []
generated_summaries_before = []
generated_summaries_after = []

for example in test_samples:
    article = example["document"]
    reference_summary = example["summary"]

    # Generate summaries
    summary_before = generate_summary_before(article)
    summary_after = generate_summary_after(article)

    # Store results
    reference_summaries.append(reference_summary)
    generated_summaries_before.append(summary_before)
```

```

generated_summaries_after.append(summary_after)

# Compute ROUGE scores
rouge_scores_before = compute_rouge_scores(reference_summaries, generated_summaries_before)
rouge_scores_after = compute_rouge_scores(reference_summaries, generated_summaries_after)

# =====
# 🚀 Print Results
# =====
print("\n===== ROUGE Score Comparison =====\n")
print("🚀 ROUGE Scores (Before Fine-Tuning):")
for key, value in rouge_scores_before.items():
    print(f"    {key}: {value:.4f}")

print("\n✅ ROUGE Scores (After Fine-Tuning):")
for key, value in rouge_scores_after.items():
    print(f"    {key}: {value:.4f}")

# =====
# 🚀 Optional: Show Percentage Improvement
# =====
print("\n📊 Percentage Improvement After Fine-Tuning:")
for key in rouge_scores_before.keys():
    improvement = ((rouge_scores_after[key] - rouge_scores_before[key]) / rouge_scores_before[key]) * 100
    print(f"    {key}: {improvement:.2f}% improvement")

```



===== ROUGE Score Comparison =====

🚀 ROUGE Scores (Before Fine-Tuning):
 ROUGE-1: 0.1680
 ROUGE-2: 0.0226
 ROUGE-L: 0.1147

✅ ROUGE Scores (After Fine-Tuning):
 ROUGE-1: 0.2573
 ROUGE-2: 0.0618
 ROUGE-L: 0.1996

📊 Percentage Improvement After Fine-Tuning:
 ROUGE-1: 53.11% improvement
 ROUGE-2: 173.35% improvement
 ROUGE-L: 73.98% improvement

```
custom_text = input("Enter your text: ")
```

```
baseline_summary = generate_summary_before(custom_text)
fine_tuned_summary = generate_summary_after(custom_text)
```

```

print("\n===== Model Comparison =====")
print("\n--- Original Input ---\n", custom_text)
print("\n--- T5 Generated Summary (Before Fine-Tuning) ---\n", baseline_summary)
print("\n--- T5 Generated Summary (After Fine-Tuning) ---\n", fine_tuned_summary)

```



Enter your text: There seems to be no end in sight to the agonising wait of family members of eight persons trapped under the debris

===== Model Comparison =====

```

--- Original Input ---
There seems to be no end in sight to the agonising wait of family members of eight persons trapped under the debris about 14 km in:

--- T5 Generated Summary (Before Fine-Tuning) ---
the conveyor belt was synchronised with the functioning of the TBM. it would shift excavated material, soil mixed with small stones:

--- T5 Generated Summary (After Fine-Tuning) ---
A conveyor belt has been restored after 10 days of hard work to trace the dead people trapped in the tunnel.

```

```

from transformers import T5ForConditionalGeneration, T5Tokenizer
import torch
import os

```

```

# Mount Google Drive (only required in Google Colab)
from google.colab import drive
drive.mount('/content/drive')

```

```

# Define the save path inside Google Drive
drive_path = "/content/drive/MyDrive/t5_finetuned_model"
os.makedirs(drive_path, exist_ok=True)

```

```
# Load your fine-tuned model (if already in memory, skip this step)
model = T5ForConditionalGeneration.from_pretrained("fine_tuned_t5_xsum") # Replace with your fine-tuned model
tokenizer = T5Tokenizer.from_pretrained("fine_tuned_t5_xsum")

# Save model and tokenizer to Drive
model.save_pretrained(drive_path)
tokenizer.save_pretrained(drive_path)

print(f"Model saved to {drive_path}")

🔗 Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
Model saved to /content/drive/MyDrive/t5_finetuned_model
```