UNIT 1 - HANDS-ON DOCUMENTATION

NAME - SWATHI M
SRN - PES2UG23CS635
SECTION - J

Email id : pes2ug23cs635@pesu.pes.edu
Github link : https://github.com/TechieTweet/GenAI-HandsOn-Unit1

Assignment 1: Phase 3

Observation table:

Observation Table

Task	Model	Success / Failure	Observation	Why did this happen?
Generation	BERT	Failure	Output repeated dots instead of meaningful text.	BERT is encoder-only and trained for masked token prediction, not next-word generation.
	RoBERTa	Failure	Returned only the prompt without continuation.	RoBERTa is encoder-only and cannot generate tokens autoregressively.
	BART	Partial Success	Generated continuation but text was nonsensical.	BART can generate text but base checkpoint is not fine-tuned for open-ended generation.
Fill-Mask	BERT	Failure	Error when using `<mask>` token.	BERT tokenizer expects `[MASK]` token; mismatch caused failure.
	RoBERTa	Success	Predicted correct words like generate/create.	RoBERTa is MLM-trained and uses `<mask>` token.
	BART	Partial Success	Produced reasonable masked-word predictions.	BART uses denoising pretraining and can reconstruct missing text but is not specialized for MLM.
QA	BERT	Failure	Did not produce meaningful answer.	Base BERT is not fine-tuned for QA; QA head randomly initialized.
	RoBERTa	Partial Success	Extracted partial correct answer.	Strong encoder representations but not QA fine-tuned.
	BART	Failure	Produced incoherent answer fragment.	bart-base not fine-tuned for extractive QA.

Assignment 2 : Project 6 - TL;DR for News Articles

```
summarizer = pipeline("summarization", model="facebook/bart-large-cnn")
```

```
config.json: 1.58k/? [00:00<00:00, 104kB/s]
model.safetensors: 100% 1.63G/1.63G [00:30<00:00, 33.8MB/s]
generation_config.json: 100% 363/363 [00:00<00:00, 28.5kB/s]
vocab.json: 899k/? [00:00<00:00, 16.4MB/s]
merges.txt: 456k/? [00:00<00:00, 21.3MB/s]
tokenizer.json: 1.36M/? [00:00<00:00, 40.4MB/s]
Device set to use cpu
```

```
# # generate summary

summary = summarizer(text, max_length=50, min_length=10, do_sample=False)
print(summary[0]['summary_text'])
```

```
Generative Artificial Intelligence is a form of artificial intelligence. It can produce content such as text, images, audio, and code. Generative AI also introduces risks suc
```

Key Concepts Understood

- Difference between Encoder-only and Encoder–Decoder Transformer architectures.

- How BERT and RoBERTa perform Masked Language Modeling for text understanding.
- How BART performs sequence-to-sequence generation tasks.
- Use of Hugging Face pipelines for text generation, fill-mask, question answering, and summarization.
- Importance of fine-tuning for task-specific performance.

Solution Implemented

Assignment (Model Benchmark)
Implemented 3experiments comparing BERT, RoBERTa, and BART on:
1. Text Generation
2. Fill-Mask Prediction
3. Question Answering

Observed model behavior and explained results based on architecture differences.

Project (TL;DR for News Articles)
Implemented a text summarization system using Hugging Face summarization pipeline with bart-large-cnn model.
The system takes long text as input and produces a concise summary.

Conclusion
The hands-on exercises demonstrated how different transformer architectures behave on various NLP tasks and how pretrained Hugging Face models can be used to quickly build Generative AI applications.