Name: Rithvik Rajesh Matta

SRN: PES2UG23CS485

Section: 6H

Email ID: Rithvik.matta@gmail.com

Mobile Number: 9606126580

Github Link:

- Assignment 1-> https://github.com/rtk5/sem6/blob/main/GEN-AI/unit%201/Unit1_Benchmark.ipynb
- assignment 2-> https://github.com/rtk5/TL-DR-for-News-Articles

Assignment 1: screenshot

Experiment 1)

```
models = {
    "BERT": "bert-base-uncased",
    "RoBERTa": "roberta-base",
    "BART": "facebook/bart-base"
}

prompt = "The future of Artificial Intelligence is"

for name, model in models.items():
    print(f"\n{name} OUTPUT:")
    try:
        generator = pipeline("text-generation", model=model)
        output = generator(prompt, max_length=30, num_return_sequences=1)
        print(output[0]["generated_text"])
    except Exception as e:
        print("ERROR:", e)
```

```
BERT OUTPUT:
If you want to use `BertLMHeadModel` as a standalone, add `is_decoder=True.`
Device set to use cpu
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncat
Both `max_new_tokens` (=256) and `max_length`(=30) seem to have been set. `max_new_tokens` will take precedence. Please refer to the docu
If you want to use `RobertaLMHeadModel` as a standalone, add `is_decoder=True.`
The future of Artificial Intelligence is.....

RoBERTa OUTPUT:
Device set to use cpu
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncat
Both `max_new_tokens` (=256) and `max_length`(=30) seem to have been set. `max_new_tokens` will take precedence. Please refer to the docu
The future of Artificial Intelligence is

BART OUTPUT:
Some weights of BartForCausalLM were not initialized from the model checkpoint at facebook/bart-base and are newly initialized: ['lm_head
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
Device set to use cpu
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncat
Both `max_new_tokens` (=256) and `max_length`(=30) seem to have been set. `max_new_tokens` will take precedence. Please refer to the docu
The future of Artificial Intelligence isnation mansassin ain785 176 Angle multiplayer multiplayerhesion Roth multiplayer multiplayer cont
```

Experiment 2)

```python
fill_mask_models = {
    "BERT": ("bert-base-uncased", "[MASK]"),
    "RoBERTa": ("roberta-base", "<mask>"),
    "BART": ("facebook/bart-base", "<mask>")
}

for name, (model, mask_token) in fill_mask_models.items():
    print(f"\n{name} OUTPUT:")
    pipe = pipeline("fill-mask", model=model)
    sentence = f"The goal of Generative AI is to {mask_token} new content."
    results = pipe(sentence)
    for r in results[:3]:
        print(r["token_str"], " | score:", round(r["score"], 4))
```

```
Some weights of the model checkpoint at bert-base-uncased were not used when initializing BertForMaskedLM: ['bert.pooler.den
- This IS expected if you are initializing BertForMaskedLM from the checkpoint of a model trained on another task or with an
- This IS NOT expected if you are initializing BertForMaskedLM from the checkpoint of a model that you expect to be exactly

BERT OUTPUT:
Device set to use cpu
create  | score: 0.5397
generate  | score: 0.1558
produce  | score: 0.0541

RoBERTa OUTPUT:
Device set to use cpu
 generate  | score: 0.3711
 create  | score: 0.3677
 discover  | score: 0.0835

BART OUTPUT:
Device set to use cpu
 create  | score: 0.0746
 help  | score: 0.0657
 provide  | score: 0.0609
```

Experiment 3)

```python
qa_models = {
    "BERT": "bert-base-uncased",
    "RoBERTa": "roberta-base",
    "BART": "facebook/bart-base"
}

context = "Generative AI poses significant risks such as hallucinations, bias, and deepfakes."
question = "What are the risks?"

for name, model in qa_models.items():
    print(f"\n{name} OUTPUT:")
    qa = pipeline("question-answering", model=model)
    result = qa(question=question, context=context)
    print("Answer:", result["answer"])
    print("Score:", round(result["score"], 4))
```

```
BERT OUTPUT:
Some weights of BertForQuestionAnswering were not initialized from the model checkpoint at bert-base-uncased and are newly initialized: ['qa_outp
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
Device set to use cpu
Some weights of RobertaForQuestionAnswering were not initialized from the model checkpoint at roberta-base and are newly initialized: ['qa_output
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
Answer: poses significant risks such as hallucinations
Score: 0.0078

RoBERTa OUTPUT:
Device set to use cpu
Answer: deepfakes.
Score: 0.0131

BART OUTPUT:
Some weights of BartForQuestionAnswering were not initialized from the model checkpoint at facebook/bart-base and are newly initialized: ['qa_out
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
Device set to use cpu
Answer: such
Score: 0.0284
```

Table)

Task	Model	Classification (Success/Failure)	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BERT	*Failure*	The model repeated punctuation and failed to generate a meaningful continuation beyond the prompt.	BERT is an encoder-only model trained for understanding and masked token
	RoBERTa	*Failure*	The model echoed the prompt without generating any additional content.	RoBERTa is also encoder-only and lacks a decoder to generate new tokens s
	BART	*Partial Success*	The model generated a long sequence of text, but the output was incoherent and largely gibberish.	Although BART has a decoder, it is trained for sequence-to-sequence tasks,
Fill-Mask	BERT	*Success*	The model confidently predicted appropriate words such as "create", "generate", and "produce".	BERT was explicitly trained using Masked Language Modeling (MLM), making
	RoBERTa	*Success*	The model correctly predicted words like "generate" and "create" with high probability scores.	RoBERTa uses an optimized MLM training strategy, improving its ability to pre
	BART	*Partial Success*	The model predicted reasonable words such as "create", but with much lower confidence scores.	BART is trained using denoising objectives rather than pure MLM, so masked
QA	BERT	*Partial Success*	The model extracted the correct answer span but with a very low confidence score.	While BERT's encoder architecture supports extractive QA, it was not fine-tu
	RoBERTa	*Partial Success*	The model returned a more complete answer span but still showed very low confidence.	RoBERTa provides strong contextual representations but lacks a trained ques
	BART	*Failure*	The model returned an incorrect and unrelated answer ("Generative").	BART is not designed for extractive question answering and requires task-sp

Assignment 2: Screenshots

TL - DR for News Articles

```
url = "https://www.bbc.com/news/articles/c39428dv18yo"
print(summarize_article(url))
```

```
... AI model from Google's DeepMind reads recipe for life in DNA . Called AlphaGenome, it could help scientists discover why subtle differences in our DNA put us
```

```
long_text = """
India announced a new policy initiative today aimed at boosting semiconductor manufacturing, marking a significant step in the country's push toward technolog

Under the new initiative, the Government of India plans to expand financial incentives, streamline regulatory approvals, and improve infrastructure support for

A key focus of the initiative is talent and research. By supporting specialized training programs, strengthening university-industry partnerships, and funding

Industry experts view this policy as a timely intervention, especially amid global supply chain uncertainties. If executed effectively, the initiative could p
"""
print(summarize_article(long_text))
```

```
Your max_length is set to 75, but your input_length is only 56. Since this is a summarization task, where outputs shorter than the input are typically wanted, y
 India announced a new policy initiative aimed at boosting semiconductor manufacturing . The move comes as part of a broader strategy to strengthen domestic cap
```