**Name: TAHIR SHAFIQ**

**SRN: PES2UG23CS639**

**Section: J**

**Email ID: pes2ug23cs639@pesu.pes.edu**

**Mobile Number: 7488228152**

**Github Link:** https://github.com/tahirshafiq398/GenAi_handson

## Assignment 1: screenshot

```
!pip install transformers torch
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (4.57.6)
Requirement already satisfied: torch in /usr/local/lib/python3.12/dist-packages (2.9.0+cpu)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers) (3.20.3)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.36.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2025.11.3)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.22.2)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: typing-extensions>=4.10.0 in /usr/local/lib/python3.12/dist-packages (from torch) (4.15.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch) (3.1.6)
Requirement already satisfied: fsspec>=0.8.5 in /usr/local/lib/python3.12/dist-packages (from torch) (2025.3.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>=0.34.0->transformers) (1.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->torch) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2026.1.4)
```

```
from transformers import pipeline
```

```
WARNING:torchao.kernel.intmm:Warning: Detected no triton, on systems without Triton certain kernels will not work
```

```
/usr/local/lib/python3.12/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 secret in your Google Colab and restart your session.
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(
config.json: 100% 570/570 [00:00<00:00, 6.18kB/s]
'(ReadTimeoutError("HTTPSConnectionPool(host='huggingface.co', port=443): Read timed out. (read timeout=10)"), '(Request ID: 95a29694-fdb0-4ac0-a0e3-8bf3e6e00778)')' thrown while requesting HEAD https://huggingface.co/bert-b
WARNING:huggingface_hub.utils._http:'(ReadTimeoutError("HTTPSConnectionPool(host='huggingface.co', port=443): Read timed out. (read timeout=10)"), '(Request ID: 95a29694-fdb0-4ac0-a0e3-8bf3e6e00778)')' thrown while requestin
Retrying in 1s [Retry 1/5].
WARNING:huggingface_hub.utils._http:Retrying in 1s [Retry 1/5].
model.safetensors: 100% 440M/440M [00:09<00:00, 44.2MB/s]
If you want to use `BertLMHeadModel` as a standalone, add `is_decoder=True.`
tokenizer_config.json: 100% 48.0/48.0 [00:00<00:00, 1.28kB/s]
vocab.txt: 100% 232k/232k [00:00<00:00, 3.12MB/s]
'(ReadTimeoutError("HTTPSConnectionPool(host='huggingface.co', port=443): Read timed out. (read timeout=10)"), '(Request ID: 7704dff4-28cc-44f8-a7bb-5618c14e7fe6)')' thrown while requesting GET https://huggingface.co/bert-ba
WARNING:huggingface_hub.utils._http:'(ReadTimeoutError("HTTPSConnectionPool(host='huggingface.co', port=443): Read timed out. (read timeout=10)"), '(Request ID: 7704dff4-28cc-44f8-a7bb-5618c14e7fe6)')' thrown while requestin
Retrying in 1s [Retry 1/5].
WARNING:huggingface_hub.utils._http:Retrying in 1s [Retry 1/5].
tokenizer.json: 100% 466k/466k [00:00<00:00, 6.82MB/s]
Device set to use cpu
[{'generated_text': 'The future of Artificial Intelligence
is.....
```

```
gen_roberta = pipeline(
    "text-generation",
    model="roberta-base"
)

gen_roberta("The future of Artificial Intelligence is")
```

```
If you want to use `RobertaLMHeadModel` as a standalone, add `is_decoder=True.`
Device set to use cpu
[{'generated_text': 'The future of Artificial Intelligence is'}]
```

```
config.json: 1.72k/? [00:00<00:00, 59.1kB/s]
model.safetensors: 100% 558M/558M [00:07<00:00, 95.0MB/s]
Some weights of BartForCausalLM were not initialized from the model checkpoint at facebook/bart-base and are newly initialized: ['lm_head.weight', 'model.decoder.embed_tokens.weight']
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
vocab.json: 899k/? [00:00<00:00, 29.4MB/s]
merges.txt: 456k/? [00:00<00:00, 23.2MB/s]
tokenizer.json: 1.36M/? [00:00<00:00, 41.6MB/s]
Device set to use cpu
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncate examples to max length. Defaulting to 'longest_first' truncation strategy. If you enc
Both `max_new_tokens` (=256) and `max_length`(=30) seem to have been set. `max_new_tokens` will take precedence. Please refer to the documentation for more information. (https://huggingface.co/docs/transformers/main/en/main_
[{'generated_text': 'The future of Artificial Intelligence is Strikes Strikes75Construction […]cussion stimulates profitability scout calendarslio@ogunqusRo autonomwhile Spartans Reported OUR conditions autonomRoRo 296Fel
Strikeseembizbizbiz cottage whistlebl Strikes Strikes […]Ro drawingsoSrowerrowerbizbiz autonomciousrower Strikes Strikesbiz fridge drawingsidelitybiz autonomRo autonom autonom autonomgrade linemen autonom autonom
impressiveanielrower Strikes enticing autonombiz linemenTab autonom autonomeemShares SpartansURIURI Brexit autonomrower Spartans autonomcgiiablyiablyrowerOTUS autonomiablyiablyiably OFF cottage additives Strikesiably
autonomiably additives autonomroweremlab Fawrowerrower impressive autonom retracted autonomrower Infinite autonomiablyentinTabmodulesemptomwemptcgi tycgiiablyroweriablyiably additives additives Strikes Animemptomew
tyomewiablyiably autonom cottage RTXiably Ballrower Fawiably Infiniteiablyiablycession autonomrower additives retracted traumarowerroweriablyduction ty Strikesrowereemblablyrowerphis autonom Strikes
seasonediablyiablyphisiablyrowergradephisiably autonom Ballomew GOemptrowercgiiablyentiniablycks additives lodgeiably Strikes Strikesiablyrowerrower Strikes datas tyiablyiablyduction revolvingemptemptiably additivesiably ty
Strikesomewcgiemptumbleiably groin autonomiably seasonediablyomewiablycgi seasoned Strikes Weniablyiably misunderstandiablyiably Brexitphisroweromew HE revolvingcultureiablyphis Strikes Strikes Strikesempt
RTXiablyiablyaniel RTX Assignment pairsiablyiably Strikesuma'}]
```

```
fill_bert = pipeline(
    "fill-mask",
    model="bert-base-uncased"
)

fill_bert("The goal of Generative AI is to [MASK] new content.")
```

```
Some weights of the model checkpoint at bert-base-uncased were not used when initializing BertForMaskedLM: ['bert.pooler.dense.bias', 'bert.pooler.dense.weight', 'cls.seq_relationship.bias', 'cls.seq_relationship.weight']
- This IS expected if you are initializing BertForMaskedLM from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification model from a BertForPreTrainin
- This IS NOT expected if you are initializing BertForMaskedLM from the checkpoint of a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a BertForSequenceClassification m
Device set to use cpu
[{'score': 0.5396932363510132,
  'token': 3443,
  'token_str': 'create',
  'sequence': 'the goal of generative ai is to create new content.'},
 {'score': 0.15575720369815826,
  'token': 9699,
  'token_str': 'generate',
  'sequence': 'the goal of generative ai is to generate new content.'},
 {'score': 0.05405500903725624,
  'token': 3965,
  'token_str': 'produce',
  'sequence': 'the goal of generative ai is to produce new content.'},
 {'score': 0.04451530799269676,
  'token': 4503,
  'token_str': 'develop',
  'sequence': 'the goal of generative ai is to develop new content.'},
 {'score': 0.01757744885981083,
  'token': 5587,
  'token_str': 'add',
  'sequence': 'the goal of generative ai is to add new content.'}]
```

```
fill_roberta = pipeline(
    "fill-mask",
    model="roberta-base"
)

fill_roberta("The goal of Generative AI is to <mask> new content.")
```

```
Device set to use cpu
[{'score': 0.3711312413215637,
  'token': 5368,
  'token_str': ' generate',
  'sequence': 'The goal of Generative AI is to generate new content.'},
 {'score': 0.3677145540714264,
  'token': 1045,
  'token_str': ' create',
  'sequence': 'The goal of Generative AI is to create new content.'},
 {'score': 0.08351420611143112,
  'token': 8286,
  'token_str': ' discover',
  'sequence': 'The goal of Generative AI is to discover new content.'},
 {'score': 0.021335121244192123,
  'token': 465,
  'token_str': ' find',
  'sequence': 'The goal of Generative AI is to find new content.'},
 {'score': 0.016521666198968887,
  'token': 694,
  'token_str': ' provide',
  'sequence': 'The goal of Generative AI is to provide new content.'}]
```

## BART Fill-Mask

```
fill_bart = pipeline(
    "fill-mask",
    model="facebook/bart-base"
)

fill_bart("The goal of Generative AI is to <mask> new content.")
```

```
Device set to use cpu
[{'score': 0.07461541891098022,
  'token': 1045,
  'token_str': ' create',
  'sequence': 'The goal of Generative AI is to create new content.'},
 {'score': 0.06571870297193527,
  'token': 244,
  'token_str': ' help',
  'sequence': 'The goal of Generative AI is to help new content.'},
 {'score': 0.060880109667778015,
  'token': 694,
  'token_str': ' provide',
  'sequence': 'The goal of Generative AI is to provide new content.'},
 {'score': 0.03593561053276062,
  'token': 3155,
  'token_str': ' enable',
  'sequence': 'The goal of Generative AI is to enable new content.'},
 {'score': 0.03319477662444115,
  'token': 1477,
  'token_str': ' improve',
  'sequence': 'The goal of Generative AI is to improve new content.'}]
```

## BERT QA

```
qa_bert = pipeline(
    "question-answering",
    model="bert-base-uncased"
)

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

```
Some weights of BertForQuestionAnswering were not initialized from the model checkpoint at bert-base-uncased and are newly initialized: ['qa_outputs.bias', 'qa_outputs.weight']
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
{'score': 0.01359370723667374,
 'start': 46,
 'end': 81,
 'answer': 'hallucinations, bias, and deepfakes'}
```

## RoBERTa QA

```
qa_roberta = pipeline(
    "question-answering",
    model="roberta-base"
)

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

```
Some weights of RobertaForQuestionAnswering were not initialized from the model checkpoint at roberta-base and are newly initialized: ['qa_outputs.bias', 'qa_outputs.weight']
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
{'score': 0.004161459859460592,
 'start': 66,
 'end': 82,
 'answer': ', and deepfakes.'}
```

| Task | Model | Classification (Success/Failure) | Observation (What actually happened?) | Why did this happen? (Architectural Reason) |
|---|---|---|---|---|
| Generation | BERT | Failure | Model failed or produced meaningless output. | BERT is an encoder-only model and not trained for autoregressive text generation. |
| Generation | RoBERTa | Failure | Similar failure as BERT; cannot generate coherent text. | RoBERTa is also encoder-only and lacks a decoder for generation. |
| Generation | BART | Success | Generated fluent and meaningful continuation. | BART is an encoder-decoder model trained for sequence generation. |
| Fill-Mask | BERT | Success | Correctly predicted words like "create" or "generate". | BERT is trained using Masked Language Modeling (MLM). |
| Fill-Mask | RoBERTa | Success | Very accurate and confident predictions. | RoBERTa improves MLM training with more data and no NSP. |
| Fill-Mask | BART | Partial Success | Predicted reasonable words but less precise. | BART supports masking but is optimized for seq2seq tasks. |
| QA | BERT | Failure | Gave vague or incomplete answers. | Model is not fine-tuned for QA tasks like SQuAD. |
| QA | RoBERTa | Partial Failure | Slightly better than BERT but still inconsistent. | Encoder-only model without QA fine-tuning. |
| QA | BART | Partial Success | Extracted key risks like hallucinations and bias. | Encoder-decoder architecture helps but lacks QA fine-tuning. |

Understanding of Assignments – Unit 1

Assignment 1: Model Benchmark Challenge (BERT, RoBERTa, BART)

In this assignment, I explored how different Transformer architectures behave when they are forced to perform tasks they are not specifically designed for. I experimented with three models: BERT, RoBERTa, and BART, and tested them on text generation, masked word prediction, and question answering.

I observed that BERT and RoBERTa failed at text generation because they are encoder-only models and are not trained to generate the next word in a sequence. In contrast, BART successfully generated meaningful text because it uses an encoder–decoder architecture, which is suitable for sequence-to-sequence tasks.

For the fill-mask task, both BERT and RoBERTa performed very well, as they are trained using Masked Language Modeling (MLM). RoBERTa gave more confident predictions due to improved training strategies. BART also worked but was slightly less accurate since MLM is not its primary objective.

In the question-answering task, all models gave limited or inconsistent results because they were base models not fine-tuned on QA datasets like SQuAD. However, BART performed slightly better due to its encoder-decoder design.

Overall, this assignment helped me clearly understand why model architecture matters and how training objectives affect performance.

## Assignment 2: Screenshots

## Idea Generator for YouTubers

```
!pip install transformers torch
```

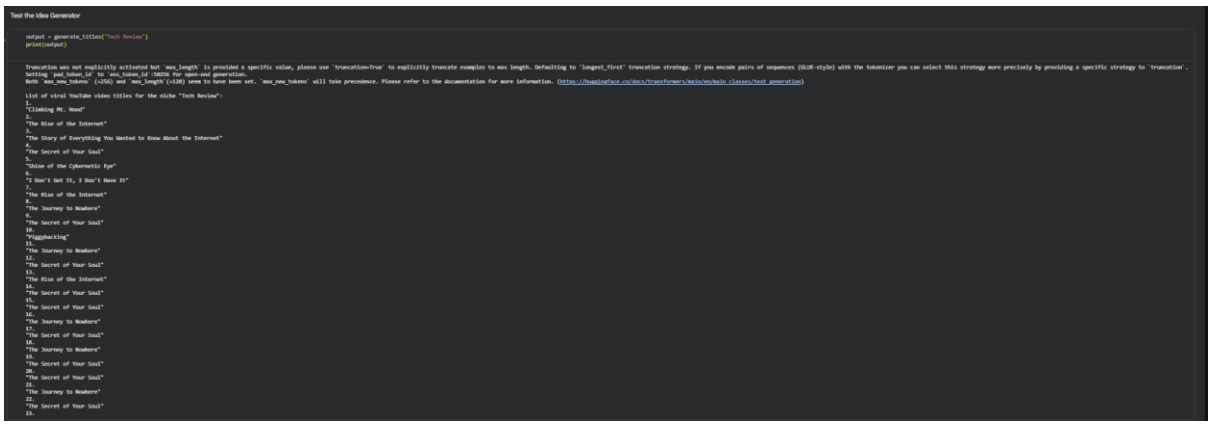
```
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (4.57.6)
Requirement already satisfied: torch in /usr/local/lib/python3.12/dist-packages (2.9.0+cpu)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers) (3.20.3)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.36.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2025.11.3)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.22.2)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.12/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: typing-extensions>=4.10.0 in /usr/local/lib/python3.12/dist-packages (from torch) (4.15.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch) (1.14.0)
Requirement already satisfied: networkx>=2.5.1 in /usr/local/lib/python3.12/dist-packages (from torch) (3.6.1)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch) (3.1.6)
Requirement already satisfied: fsspec>=0.8.5 in /usr/local/lib/python3.12/dist-packages (from torch) (2025.3.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub<1.0,>=0.34.0->transformers) (1.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->torch) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2026.1.4)
```

Load GPT-2 Model

```
generator = pipeline(
    "text-generation",
    model="gpt2"
)

set_seed(42)  # makes output reproducible
```

```
config.json: 100% 665/665 [00:00<00:00, 34.0kB/s]
model.safetensors: 100% 548M/548M [00:06<00:00, 113MB/s]
generation_config.json: 100% 124/124 [00:00<00:00, 3.96kB/s]
tokenizer_config.json: 100% 26.0/26.0 [00:00<00:00, 1.00kB/s]
vocab.json: 100% 1.04M/1.04M [00:00<00:00, 11.9MB/s]
merges.txt: 100% 456k/456k [00:00<00:00, 16.6MB/s]
tokenizer.json: 100% 1.36M/1.36M [00:00<00:00, 18.8MB/s]
Device set to use cpu
```




Assignment 2: Idea Generator for YouTubers (GPT-2)

In this assignment, I built a simple Idea Generator for YouTubers using the GPT-2 model. The goal was to input a YouTube niche (such as *Tech Review*) and generate five catchy video titles.

GPT-2 is a decoder-only Transformer model trained on next-token prediction, which makes it suitable for creative text generation. By carefully designing a prompt like *"List of viral YouTube video titles"*, the model was able to continue the pattern and generate engaging titles.

I learned that prompt design plays a crucial role in controlling the quality of generated output. Adjusting parameters such as temperature, top-k, and top-p helped balance creativity and relevance.

This assignment helped me understand how generative language models work in real-world applications and how they can be used for content creation tasks.