

Assignment 1: Hugging Face Model Exploration

Objective

The objective of this assignment is to explore an open-source language model (LLM) from the Hugging Face Hub, install it locally, and perform a simple NLP task to understand how transformers can be applied in real-world scenarios.

Steps

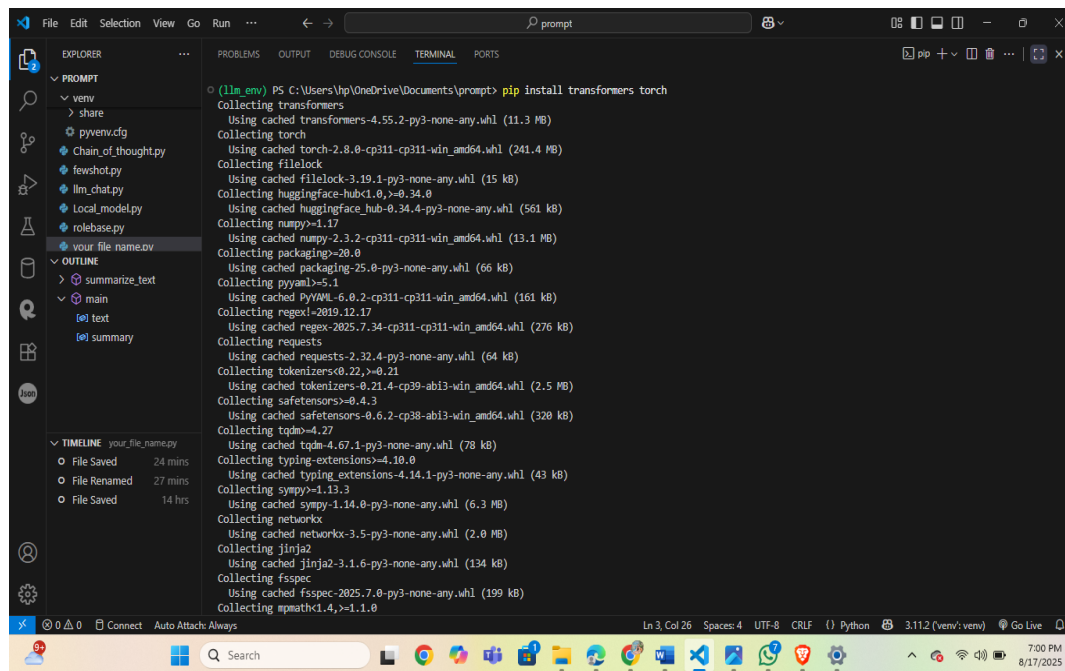
1. Environment Setup

```
# Create a new virtual environment
python -m venv llm_env
```

```
# Activate the environment (Windows)
llm_env\Scripts\activate
```

```
# Activate the environment (Linux/Mac)
source llm_env/bin/activate
```

```
# Install Hugging Face Transformers and Torch
pip install transformers torch
```



The screenshot shows a Windows terminal window with the command prompt. The user has entered the command `pip install transformers torch`. The terminal output shows the progress of the installation, including the collection of various packages and their versions. The packages listed include transformers, torch, filelock, huggingface-hub, numpy, packaging, pyyaml, regex, requests, tokenizers, safetensors, tqdm, typing-extensions, sympy, networkx, Jinja2, and fspec. The installation is complete, and the terminal shows the command prompt again.

2. Model Selection

For this task, the chosen model is facebook/bart-large-cnn, a transformer-based model widely used for text summarization tasks.

3. Model Loading in Python

from transformers import pipeline

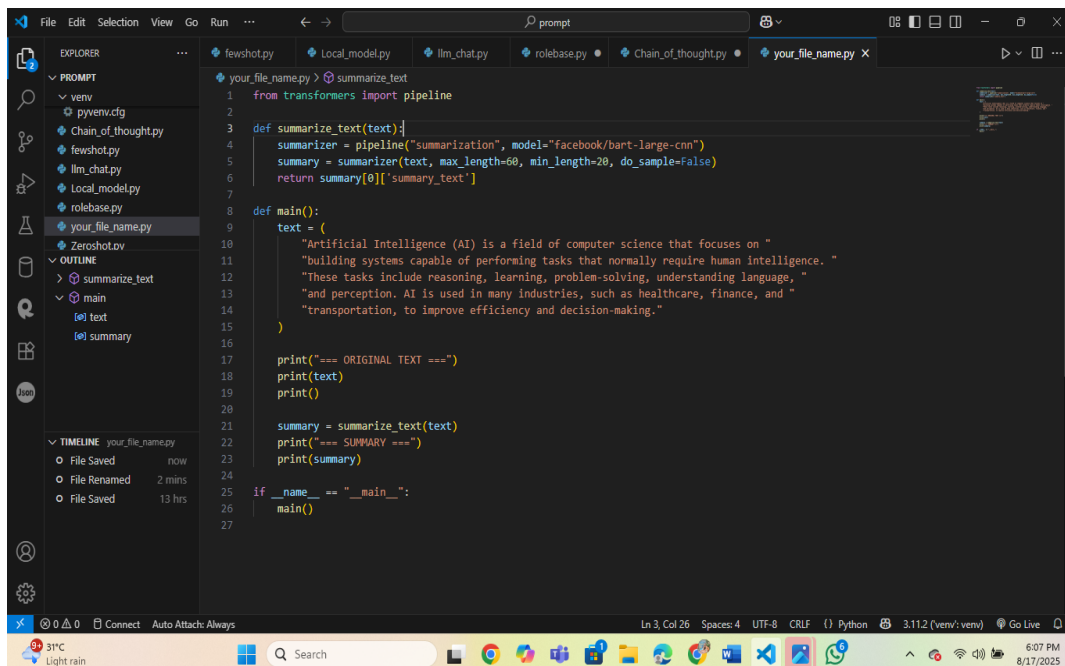
```
def summarize_text(text):
    summarizer = pipeline("summarization", model="facebook/bart-large-cnn")
    summary = summarizer(text, max_length=60, min_length=20, do_sample=False)
    return summary[0]['summary_text']

def main():
    text = (
        "Artificial Intelligence (AI) is a field of computer science that focuses on "
        "building systems capable of performing tasks that normally require human intelligence. "
        "These tasks include reasoning, learning, problem-solving, understanding language, "
        "and perception. AI is used in many industries, such as healthcare, finance, and "
        "transportation, to improve efficiency and decision-making."
    )

    print("=== ORIGINAL TEXT ===")
    print(text)
    print()

    summary = summarize_text(text)
    print("=== SUMMARY ===")
    print(summary)

if __name__ == "__main__":
    main()
```



```
your_file_name.py > summarize_text
1 from transformers import pipeline
2
3 def summarize_text(text):
4     summarizer = pipeline("summarization", model="facebook/bart-large-cnn")
5     summary = summarizer(text, max_length=60, min_length=20, do_sample=False)
6     return summary[0]['summary_text']
7
8 def main():
9     text = (
10         "Artificial Intelligence (AI) is a field of computer science that focuses on "
11         "building systems capable of performing tasks that normally require human intelligence. "
12         "These tasks include reasoning, learning, problem-solving, understanding language, "
13         "and perception. AI is used in many industries, such as healthcare, finance, and "
14         "transportation, to improve efficiency and decision-making."
15     )
16
17     print("=== ORIGINAL TEXT ===")
18     print(text)
19     print()
20
21     summary = summarize_text(text)
22     print("=== SUMMARY ===")
23     print(summary)
24
25 if __name__ == "__main__":
26     main()
27
```

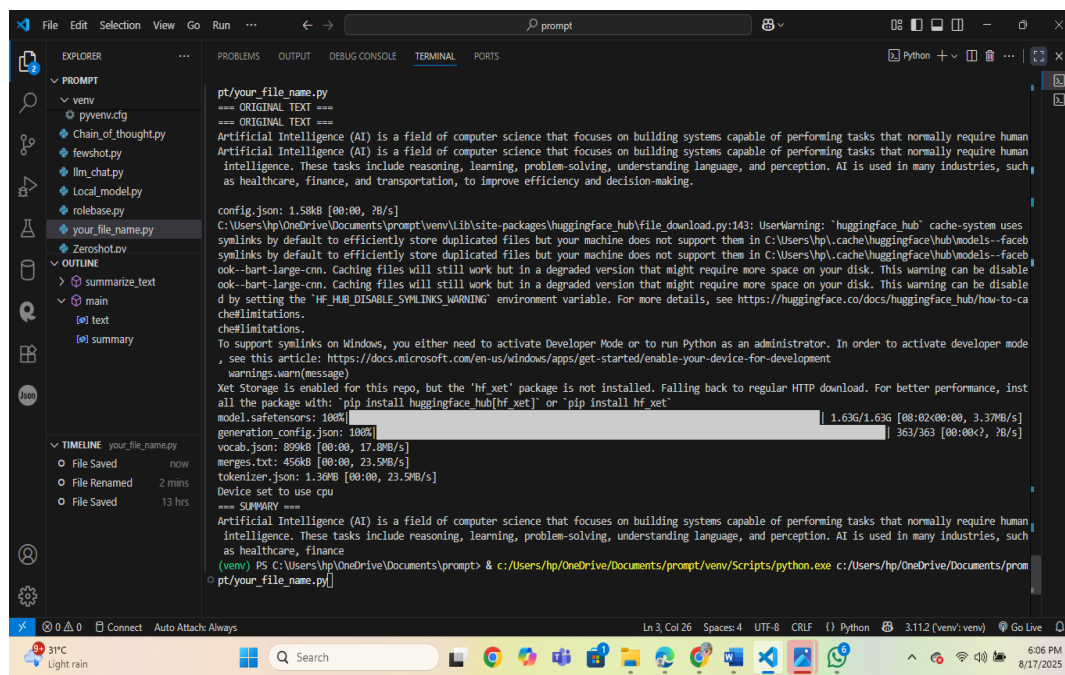
4. Sample Output

=== ORIGINAL TEXT ===

Artificial Intelligence (AI) is a field of computer science that focuses on building systems capable of performing tasks that normally require human intelligence. These tasks include reasoning, learning, problem-solving, understanding language, and perception. AI is used in many industries, such as healthcare, finance, and transportation, to improve efficiency and decision-making.

=== SUMMARY ===

AI is a branch of computer science focused on creating systems that perform tasks requiring human intelligence. It is applied across industries like healthcare, finance, and transportation to enhance efficiency and decision-making.



```
pt/your_file_name.py
=== ORIGINAL TEXT ===
=== ORIGINAL TEXT ===
Artificial Intelligence (AI) is a field of computer science that focuses on building systems capable of performing tasks that normally require human intelligence. These tasks include reasoning, learning, problem-solving, understanding language, and perception. AI is used in many industries, such as healthcare, finance, and transportation, to improve efficiency and decision-making.

config.json: 1.58kB [00:00, 2B/s]
C:\Users\hp\OneDrive\Documents\prompt\venv\Lib\site-packages\huggingface_hub\file_download.py:143: UserWarning: 'huggingface_hub' cache-system uses symlinks by default to efficiently store duplicated files but your machine does not support them in C:\Users\hp\OneDrive\Documents\prompt\venv\Lib\site-packages\huggingface_hub\models--facebook--bart-large-cnn. Caching files will still work but in a degraded version that might require more space on your disk. This warning can be disabled by setting the 'HF_HUB_DISABLE_SYMLINKS_WARNING' environment variable. For more details, see https://huggingface.co/docs/huggingface_hub/how-to-cache#limitations.
To support symlinks on Windows, you either need to activate Developer Mode or to run Python as an administrator. In order to activate developer mode, see this article: https://docs.microsoft.com/en-us/windows/apps/get-started/enable-your-device-for-development
warnings.warn(message)
Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed. Falling back to regular HTTP download. For better performance, install all the package with: 'pip install huggingface_hub[hf_xet]' or 'pip install hf_xet'
model.safetensors: 100%| 1.63G/1.63G [08:02<00:00, 3.37MB/s]
generation_config.json: 100%| 363/363 [00:00<?, 2B/s]
vocab.json: 899kB [00:00, 17.8MB/s]
merges.txt: 456kB [00:00, 23.5MB/s]
tokenizer.json: 1.30MB [00:00, 23.5MB/s]
Device set to use cpu
=== SUMMARY ===
Artificial Intelligence (AI) is a field of computer science that focuses on building systems capable of performing tasks that normally require human intelligence. These tasks include reasoning, learning, problem-solving, understanding language, and perception. AI is used in many industries, such as healthcare, finance
(venv) PS C:\Users\hp\OneDrive\Documents\prompt> & c:/Users/hp/OneDrive/Documents/prompt/venv/Scripts/python.exe c:/Users/hp/OneDrive/Documents/prompt/your_file_name.py
```

Reflection

This exercise demonstrated how to set up a Python environment, install machine learning libraries, and use an open-source model from Hugging Face. By working with BART, we performed text summarization, gaining insights into how transformer-based models condense natural language while preserving key meaning. This hands-on approach builds the foundation for working with larger LLMs in local or cloud environments.

Deliverables Completed:

- Environment setup commands
- Model loading code
- Sample output with explanation

- Report format with placeholders replaced by images