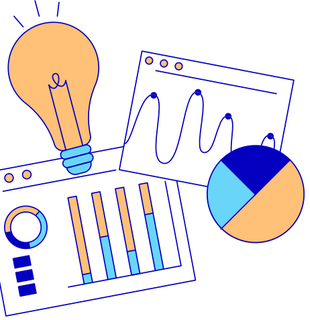


Mistral Local LLM Summarizer

by using Hugging Face Transformers

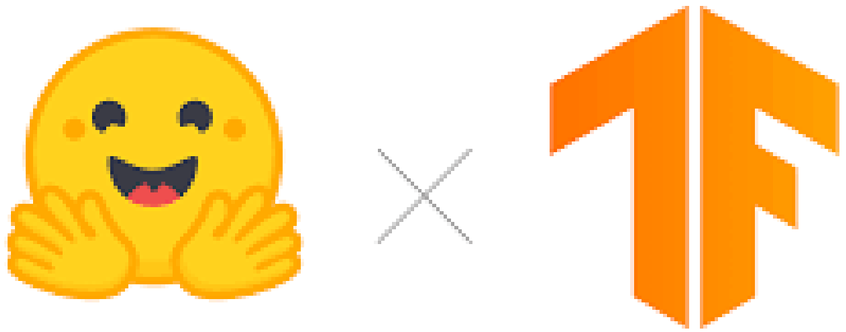
Subham Dey

What is Hugging Face Transformers?

Hugging Face Transformers is an open-source library for natural language processing (NLP) and computer vision, offering pre-trained models like BERT, GPT, and T5. Key features include:

- Intuitive APIs for tasks such as text generation, summarization, translation, and classification.
- The `pipeline()` function for easy access without extensive machine learning knowledge.
- Token processing and contextual embeddings generation.
- Compatibility with both CPU and GPU, and integration with PyTorch and TensorFlow.

Models are readily available on the Hugging Face Hub for easy sharing.



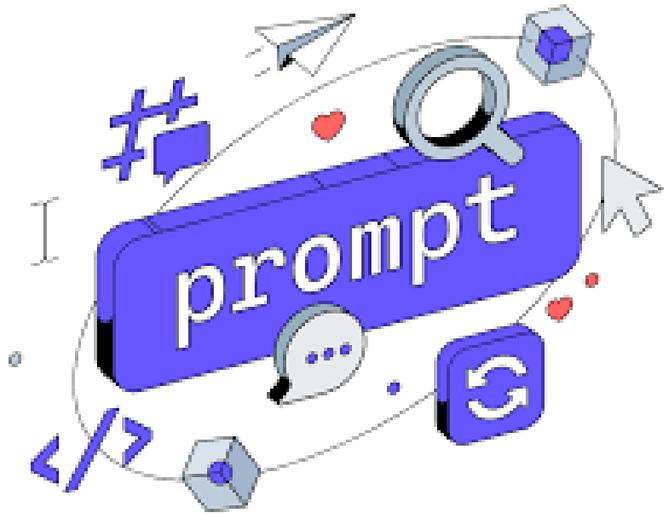
Subham Dey

What is NLP ?

Natural Language Processing (NLP) is a field of Artificial Intelligence (AI) that focuses on the interaction between computers and human languages. It enables machines to read, understand, generate, and respond to human language in a meaningful way.

What is Prompt Engineering ?

Prompt engineering is the practice of designing and refining input instructions (prompts) to guide language models like ChatGPT or T5 to produce accurate, relevant, and useful responses .



Subham Dey

To create Mistral Local LLM Summarizer you need to follow the following steps :

Step-by-Step Process to Build the Project

- **Set Up the Environment**
- **Choose and Load a Model**
- **Create the Summarizer Script**
- **Test with Real-world Texts**
- **Prompt Engineering Experiments**
- **Organize the Project Structure**

Subham Dey

Mistral Local LLM Summarizer :

This project demonstrates how to build and run a local language model (LLM) summarizer using Hugging Face's transformers library. The model used is , a free, open-source transformer optimized for text-to-text generation tasks.
No paid APIs or GPUs are required — this project is entirely CPU-based and runs locally in a Jupyter environment using Python.

Subham Dey

Features

Text Summarization

- Summarizes both short and long documents

Prompt Engineering

- Flexible prompt-based instruction input

Local Execution

- Runs fully on CPU

Real-world Test Cases

- Includes examples from AI, healthcare, and education domains

Subham Dey

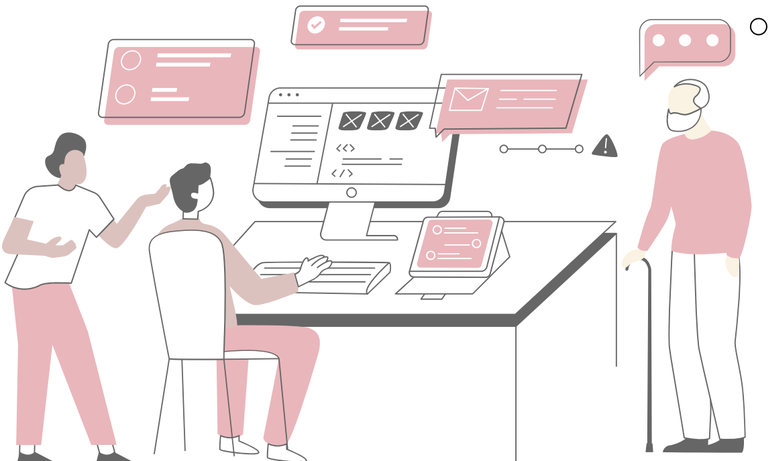
ESSENTIAL TOOLS AND TECHNOLOGIES

- PYTHON : THE FUNDAMENTAL PROGRAMMING LANGUAGE
- HUGGING FACE TRANSFORMERS : PIPELINE AND MODELS FOR LARGE LANGUAGE MODELS (LLMS)
- JUPYTER NOTEBOOK : DEVELOPMENT AND TESTING ENVIRONMENT
- FLAN-T5-SMALL : THE MODEL UTILIZED FOR SUMMARIZATION

Project Structure

Mistral Test/

- `mistral_test.py` # Test environment loading
- `mistral_api_test.ipynb` # Main summarization notebook
- `.env` # API keys (if necessary)
- `README.md` # Project documentation (this file)



SNAPSHOTS

Subham Dey

```
from transformers import pipeline

# Load the model pipeline (Flan-T5 is public and free)
generator = pipeline("text2text-generation", model="google/flan-t5-small")

# Run the prompt
output = generator("Tell me a fun fact about AI!", max_new_tokens=50)

print("Answer:", output[0]['generated_text'])
```

```
Device set to use cpu
Answer: An AI is a device used to measure distance between two objects.
```

```
generator("Summarize this: Artificial Intelligence is transforming every industry...", max_new_tokens=50)
```

```
[{'generated_text': 'Artificial intelligence is transforming every industry.'}]
```

```
long_text = """
Artificial Intelligence (AI) is transforming every industry — from automating repetitive tasks
to making sense of massive amounts of data. In healthcare, AI is used for faster diagnoses,
in finance for fraud detection, and in education for personalized learning experiences.
The pace of innovation continues to grow.
"""

result = generator("Summarize this: " + long_text, max_new_tokens=100)
print("➤ Summary:", result[0]['generated_text'])
```

```
➤ Summary: Artificial Intelligence (AI) is transforming every industry — from automating repetitive tasks to making sense of massive amounts of dat
```

SNAPSHOTS

Subham Dey

```
while True:
    user_input = input("You: ")
    if user_input.lower() in ["exit", "quit"]:
        break
    response = generator(user_input, max_new_tokens=60)
    print("Bot:", response[0]['generated_text'])
```

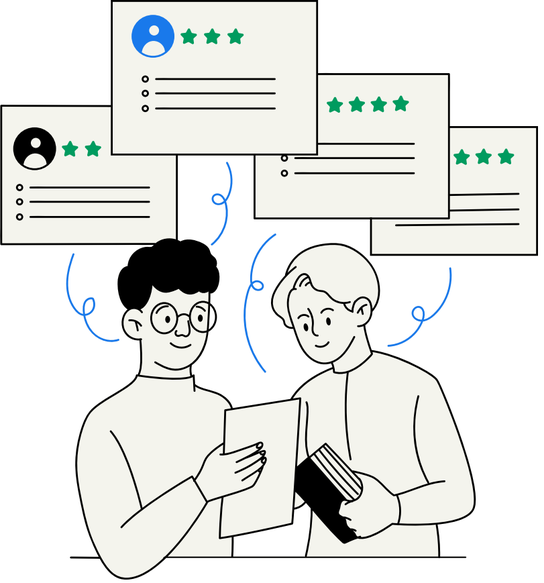
```
You:  Say some interesting facts about Rome
Bot: Rome is a city in the southern hemisphere.
```

Short Prompt Output :-
Prompt: Tell me a fun fact about AI!
Answer: An AI is a device used to measure distance between two objects.

Long Document Summarization :-
Prompt: Summarize this: "Artificial Intelligence is transforming every industry..."
Summary: AI is transforming every industry — from automating repetitive tasks to making sense of massive amounts of data.

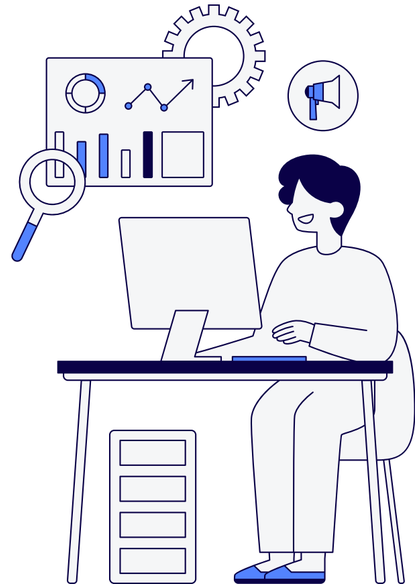
python code :-

```
from transformers import pipeline

# Load the model pipeline
generator = pipeline("text2text-generation",
model="google/flan-t5-small")

# Prompt the model
output = generator("Summarize this: Artificial Intelligence is
transforming every industry...", max_new_tokens=50)

print("Summary:", output[0]['generated_text'])
```



The End

Thank Your For Reading



Subham Dey