📰 Article Summarization with Transformers

This project demonstrates the use of a transformer-based model to perform text summarization on a dataset of news articles using the Hugging Face `transformers` library.

📂 Files Included

- `article-summarization-documented.ipynb` - The complete Jupyter Notebook with code and markdown documentation.
- `summarized_articles.csv` - CSV file containing original articles and their corresponding summaries.
- `sample_output.png` - (Optional) A sample screenshot of results for quick preview.

📚 Dataset

The dataset used contains articles to be summarized. Each article is processed to a max length of 1024 characters to suit the summarizer's input constraints.

⚙️ Model

The pipeline initially uses Facebook's `bart-large-cnn` model for summarization. Due to memory constraints and performance optimization, a smaller model such as `sshleifer/distilbart-cnn-12-6` is recommended.

🧠 Steps

1. Load the model and tokenizer using Hugging Face Transformers.
2. Process the input dataset in batches to avoid memory overload.
3. Generate summaries and collect them with the original text.
4. Save results to a CSV file.

🚀 How to Run

Make sure you have the following Python packages:

- transformers
- torch
- pandas
- tqdm

Then run the notebook step by step.

✅ Sample Output

A sample of the generated summary:

| Original Text (shortened) | Summary |

|---------|---------|

| Article about economic... | The article discusses economic growth... |

Notes

- GPU recommended for faster inference.
- If running into CUDA memory issues, reduce batch size or use a smaller model.