Emotion Prediction Project Documentation

All the files required for running code and technical details are present on google drive link here:
https://drive.google.com/drive/folders/1GYHa-f-oIy9LNdXsC9d60ArBPeMn4QHp?usp=sharing

1. main.py

This script performs emotion prediction on text data using a pretrained RoBERTa-based PyTorch model.

How to Run:

Call the function:

predict(csv_path, output_csv_path="predicted_emotions.csv")

- **Input:**
  csv_path – Path to a CSV file containing test data.
  Note: The function assumes the test data has a column named **"text"**.
- **Output:**
  A file named **predicted_emotions.csv** will be generated in the same directory.
  This output file will contain:
    - A column with the original **text**.
    - A **predicted_labels** column with predictions in the format:
      [anger, fear, joy, sadness, surprise].

2. emoModel

This is a **pretrained main.py model**, fine-tuned from the RoBERTa Transformer architecture.

- The training code is **not included** in main.py to avoid delays during execution.
- Ensure that the emoModel file is located in the **same directory** as main.py for the code to run properly

3. train.ipynb

This Jupyter Notebook contains the **complete pipeline** for:

- Data preprocessing
- Model training and evaluation
- Hyperparameter tuning and experimentation

The code is well-documented with inline comments. The notebook explores different preprocessing techniques and tuning strategies to improve performance