Week 6: System Integration & Final Testing

Task 1: Integrate System Components (Tai)

Goal: Combine all components into one complete accent recognition pipeline.

- A new Google Colab notebook was created and saved as: **scripts/full_pipeline_demo.ipynb**
- The notebook loads .wav audio files and applies preprocessing using librosa
- Feature extraction includes:

1. MFCCs (13)
2. Delta MFCCs (13)
3. Spectral Contrast (7)
4. Pitch (1)
5. Formants F1–F3 (3)
6. Total features used: 37

- A new **final_model.pkl** was retrained using these 37 features and saved to **models/**
- The system prints predictions clearly, e.g.: filename.wav ➜ Predicted Accent: Vietnamese
- A graph was added at the end to visualize the extracted feature vector.

Files Updated:

- scripts/full_pipeline_demo.ipynb
- models/final_model.pkl

Task 2: System Testing with New Audio Samples (Kyle, James, Wissam)

Goal: Test the final model's accuracy using new audio recordings (not seen during training).

What You Need To Do:

- Collect or record 3 new **.wav files** (one per accent: American, British, Vietnamese) and download them to your computer. (Make sure these are **not the same audio files used during training).**
- Download these files from GitHub to your computer: **models/final_model.pkl**
- Run these files through **scripts/full_pipeline_demo.ipynb**
    1. First upload: final_model.pkl
    2. Then upload your 3 test **.wav files**
- Observe if the predictions are accurate

Save the 3 test audio files in:

- data/new_test_samples/

Task 3: Write the Testing Report (Kyle, James, Wissam)

Goal: Document model performance on real-world samples.

What To Include in the Report:

- The actual predictions for each .wav file
- Whether the prediction was correct or incorrect
- How fast the prediction ran (inference time)
- Observations: Did the model struggle with any accent?

Save the report as:

- results/system_testing_report.txt