# Data Mining Project Guidelines
## Steps to run the project

This file provides instructions on how to configure, compile, and execute the Python code for **Accent Recognition using Mel Frequency Cepstral Coefficients and Supervised Learning Algorithms A Comparative Study.**

**Requirements:**

- Python 3.7+
- TensorFlow
- Keras
- Pandas
- NumPy
- Scikit-learn
- Seaborn
- Matplotlib

**Installation:**

1. Install all required libraries using pip:
   pip install tensorflow keras pandas numpy scikit-learn seaborn matplotlib
2. Download the dataset:
   https://archive.ics.uci.edu/dataset/518/speaker+accent+recognition.
   Ensure you have access to the CSV file named accent-mfcc-data-1.csv.

**Configuration:**
1. Edit the code to specify the path to the accent-mfcc-data-1.csv file:

   df = pd.read_csv('/path/to/accent-mfcc-data-1.csv')

Adjust hyperparameters for the chosen models:
- Decision Tree Classifier
- K-nearest Neighbors Classifier
- Support Vector Machine
- Neural Network
- Random Forests

**Compilation:**
1. Open a terminal or command prompt or Open the Jupyter Notebooks (**Accent_Recognition.ipynb**) in your preferred environment.
2. Alternatively, you can Navigate to the directory containing the Python script and run (**Accent_Recognition.py**) directly.
3. If using Run Jupyter Notebook:

3.1 Run Jupyter Notebook from the terminal: jupyter notebook

3.2 Open the **Accent_Recognition.ipynb** file and run each cell sequentially. Or run the direct command to compile and execute the code

**Execution:**
1. The code will perform the following tasks:
    - Import necessary libraries.
    - Load the dataset.
    - Perform data preprocessing and exploration.
    - Train and evaluate different models for accent classification.
    - Plot visualizations of results
2. The output will include:
    - Model accuracy scores
    - Confusion matrices
    - ROC curves
    - Other relevant metrics and visualizations