DATA MINING ENDSEM

CLASSIFICATION

Requirements:

- Python version 3.8 or 3.9
- **Python Libraries needed:**
    - Pandas
    - Numpy
    - imblearn
    - scikit-learn

Submission Structure:

1) training-s145.csv : Classifier Training data set provided to us.
2) answer-s145.csv : output file containing class labels for test data.
3) classifier_endsem.py : Python script that trains the model, takes input test file and outputs 'answer-s145.csv' containing predicted labels for test data.
4) requirements_classifier.txt: Contains all the python library names required to successfully run given python script.
5) classifier-s145.sh: Shell script file to execute python script with given test file.
6) Makefile: Installs required python libraries if not already installed and takes test file path as argument and calls classifier-s145.sh and passes test file path as argument.

Steps To Execute Assignment:

1. Move into the folder where 'Makefile' is present and open Terminal/Command Prompt etc.
2. Type 'make setup' and press enter.
3. Type 'make run path='test file path/testfilename.csv' and press enter.
4. 'answer-s145.csv' will be generated and stored in the folder.

Note:

1. Do not Delete 'training-s145.csv' as it is required by classification model.

HARSH AGARWAL
21111030