Implementation and Experimentation on the Perceptron Learning Algorithm (PLA).

Problem Statement:
Given a dataset where items belong to either of the two classes [Positive (+1) and Negative (-1)], train and validate a machine learning model based on the *Perceptron Learning Algorithm*. The datasets may or may not be linearly separable and that may affect the training process.

[DO NOT USE ANY LIBRARY FUNCTION EXCEPT FOR K-FOLD CROSS VALIDATION]

The implementation should contain: **[5 points]**

1. Implementation of the basic PLA.
2. Implementation of K-fold cross validation
3. Implementation of the performance measures: Accuracy, Precision, Recall and F1

The experiments should contain: **[5+4+4=13 points]**

1. For the dataset [Dataset-1],
    a. Report values of the performance metrics for different values of K.
    b. For each K report the mean and variances of the performance metrics computed over all the folds.
    c. Make an 80:20 train-test split. For each iteration intraining, count the number of misclassified instances. Plot iteration vs #misclassified instances.
2. For the dataset [Dataset-2], make an 80:20 train-test split.
    a. For each iteration intraining, count the number of misclassified instances. Plot iteration vs #misclassified instances.
    b. Report your observation on training with an explanation in one or two sentences.
    c. Report values of the performance metrics for the test data.
3. For the dataset [Dataset-3], make an 80:20 train-test split.
    a. For each iteration intraining, count the number of misclassified instances. Plot iteration vs #misclassified instances.
    b. Report your observation on training with an explanation in one or two sentences.
    c. Report values of the performance metrics to the test data.
    d. Connect Experiment 1, Experiment 2 and Experiment 3 by commenting on the nature of the datasets.

*** The PLA may not converge for some dataset and will converge for other dataset. For the datasets where PLA does not converge use MAX iteration number to be 2000.

Datasets:
Dataset-1: https://tinyurl.com/m44j6zp6
Dataset-2: https://tinyurl.com/yn2e9ub2
Dataset-3: https://tinyurl.com/23c22rff

Submission Items:
A .zip file containing the python source code and a PDF report file. The final name should follow the template: <Assign-No>_<Your Roll No>.zip. For example, if your roll no is 15CE30021, the filename for Assignment 1 will be: `Assign-1_15ce30021.zip`

1. The python code (.py) containing the implementations with comments at function level. The first two lines should contain your name and roll no.

2. A report [PDF] containing **[2 points]**
    a. Result tables:
        i. Table 1: Containing values of the performance metrics and their variances for different values of K.
        ii. Table 2: Report training and test accuracy for experiment 2 and 3.
    b. Conclusion:
        i. What can you say about the datasets used in three experiments?
        ii. How do you connect the nature of the datasets with the experimental results?