提醒：基本功能+報告品質普通=基本分數。欲得高分者應思考充實作業成果之各項可能作法。

Homework-1 (Chapter 3. Classification)

This is a **Multiclass Classification** homework.

Part 1: The MNIST dataset

1. Use SGDClassifier for the MNIST dataset and measure the accuracy (the ratio of correct predictions) using cross-validation (cv=3).
2. Use Data Augmentation (artificially growing the training set) to see if the accuracy can be improved.

   Note: For example, you may write a function that can shift an MNIST image in any direction (left, right, up, or down) by one pixel. (You should not allow data obtained by augmentation of the training part leak into the test/validation set.)

   Ref: You can use the shift() function from the scipy.ndimage.interpolation module. For example, shift(image, [2, 1], cval=0) shifts the image two pixels down and one pixel to the right.
3. Use the confusion matrix to gain insights and find ways to improve the performance, such as normalization and hyperparameter tuning.

Part 2: The Fashion MNIST dataset

1. Use the Fashion MNIST dataset as the target for classification and repeat steps in Part 1.

   Note: You may need to write a function to load the Fashion MNIST dataset (including a training set and a test set) and flatten the features of each image into a 1D array of size 784.
2. Compare SGDClassifier's performance on the two datasets.

Part 3: Writing a report

1. Write a report within 10 pages discussing your findings. (ID and names of the group members should be listed on the cover page. Please do not include any code in the report)
2. Please upload your report on Moodle.