In The Name of Allah
Pattern Recognition (Autumn 2018)
Instructor: Mahdi Yazdian
TA: Farzane Nikzad

Project: Hand-Written Digits Classification
Due Date: 97.10.29

The final project is defined as a 10-classe handwritten image classification problem. Here, USPS dataset will be used for classification purpose. You can do the project individually or as a 2-member group. The group(s) with the best results achieve higher score.

- **Data Description**

The dataset has been uploaded on the course home page. It contains 10 different digit classes. Each class have 2000 samples.

- **Goal**

According to the related researches, you are expected to evaluate several suitable pattern recognition approaches and finally propose your best solution to solve this problem. You are free to apply any other intermediate steps such as dimensionality reduction, feature selection or etc.

  - You are expected to at least consider following methods:
    - Bayesian classifier (Using Gaussian distribution)
    - KNN classifier (Determine a suitable K using cross-validation)
    - PCA
    - LDA
    - [optional] Any other methods in the literature, such as GMM, Naïve Bayes, etc.

- **Guidelines**
  - Try to find simple and effective features according to the state-of-the-are researches. As a simple method you can resize the images into a fix size and victories the image as a feature vector. You are allowed to use any code on the internet. This step is very important in a classification problem especially in image classification.
  - Evaluate the results.
  - Prepare the final reports and the PowerPoint representation. Note that to report the results of the all implemented methods.

- **Experimental Results and Discussion.**

In order to have a same framework for comparison, the data is partitioned into test set and train set. Please report the performance of your methods on both train and test sets. To evaluate your methods, please calculate the following metrics:

  - Create a confusion matrix: It describes the performance of a classifier so that you can see what types of errors your classifier is making.
  - Calculate the precision, recall, and F1 score for each class. Then calculate average F1 score, which shows the single-number metric you are looking for.

- Calculate per class accuracy as well as average accuracy.

Please provide any arbitrary table, figure or plot to analysis your results.

- **Delivery**

You are expected to deliver:

1- The final codes. Note that your code is needed to be self-comment. Provide all possible used toolbox and implement your codes as a functional form. Note that to have a main function called "PRProject_Main.m" to run the code.

2- A comprehensive report. The final report should be a two-column 4-page report in IEEE paper format. Therefore, same as a paper, your report should contain usual sections:
   i. Abstract
   ii. Introduction: including problem definition, a brief literature review, motivation, the general view of the proposed method
   iii. Proposed method: including intuition (why should it be the better solution?), description of your algorithms.
   iv. Evaluation results: including evaluation framework, the possible figures, tables or plots. Compare the methods (including the base method) and discus about the results. Enumerate the advances or disadvantages of the best proposed method.
   v. Conclusion
   vi. References.

3- An oral presentation (with power point) on Wed 97.04.20

Please upload the materials (Codes, Report and PowerPoint) before 23:55 on 97.10.29.

Good Luck