Email from Sadri

Dear students

Please find a brief explanation of your team project as follows:

Title: Learning, Designing, Implementation and Comparison of Machine Learning Algorithms for Classification and Clustering of
Handwritten Digits.

Steps of the project:

1) Please download these three papers (or similar papers) and read them carefully. (a copy of these papers also will be posted on the course web page)

A) Paper#1: Javad Sadri, Mohammad Reza Yeganehzad, Javad Saghi, "A Novel Comprehensive Database for Offline Persian Handwriting Recognition", Published in the Journal of "Pattern Recognition (Elsevier)", Volume 60, Pages 378-393, December 2016. Currently available on-line at: http://www.sciencedirect.com/science/article/pii/S0031320316300097, December 2016. Link for Downloading the Database:
https://users.encs.concordia.ca/~j_sadri/PersianDatabase.htm

B) Paper#2: Javad Sadri, C. Y. Suen, T. D. Bui, "Application of Support Vector Machines for Recognition of Handwritten Arabic/Persian Digits", Published in the Proceedings of the Second Conference on Machine Vision and Image Processing & Applications (MVIP 2003), Vol. 1, pp. 300-307, Tehran, Iran, Feb. 2003.

C) Paper#3: Javad Sadri, C. Y. Suen, T. D. Bui, "A New Clustering Method for Improving Plasticity and Stability in Handwritten Character Recognition Systems", Published in the Proceedings of the International Conference on Pattern Recognition (ICPR 2006), pp. 1130-1133, Hong Kong , August 2006.

2) Download the sample database (digit part) from
the following links: https://users.encs.concordia.ca/~j_sadri/PersianDatabase.htm
More data and more samples of handwritten digits based on paper#1 will be given to you in next few weeks.

3) Develop Different Methods for feature extraction for digits digits according to the methods explained in Paper#1 and #2 or other similar methods in the literature.

4) Develop Different methods for classification of digits. At least 4 different classification methods for digits should be implemented, compared and tested.

5) Develop Different methods for classifier evaluations and comparison of your classifiers, show some of your correct classifications,
and misclassification cases. Show and interpreting the results of your classification visually and explain them.

6) Develop Different methods or approaches for clustering of digits's shapes. (at least 3 different methods of clustering of digits

in each group should be implemented and tested). For example different shapes and styles of writing digit 5 should be clustered and

presented visually. (for more info. look at Paper #3, and other similar papers). You should use three different algorithm for clustering or use three different types of features or 3 different types

of distance functions/or similarity measure for your clustering, and you should compare results of your clustering visually similar to Paper#3.

7) More information and details about this projects will be given in the class. These projects can be implemented in team of 1,2, 3, 4, or 5 students.

the due date for the submission of a full report and source codes will be on April 9, 2017. Projects should be presented during the same week starting

(between April 10 to April 17, 2017).

Important note: If you wish, You can choose, database of handwritten digits of any script in the world for this project.

I will be happy to meet you in my office hours, or in our class to talk about more details of this project or to answer your all questions.

Wishing you success

J. Sadr