Name:
ID:
Date:
ITU, Computer Engineering Dept.
BLG454E, Learning From Data HW3
Due. May 21, 2017, 23:00 . NO LATE SUBMISSION WILL BE ACCEPTED. Upload your report and matlab codes through Ninova. DO NOT SUBMIT THROUGH E-MAIL.

Instructors: Yusuf Yaslan (yyaslan@itu.edu.tr)
Assistant : Resul Tugay(tugayr@itu.edu)

Policy:
Please do your homeworks on your own. You are encouraged to discuss the questions with your class mates, but the code and the hw you submitted must be your own work. **Cheating is highly discouraged** for it could mean a zero or negative grade from the homework.
If a question is not clear, please let me know (via email or in class). Unless we indicate otherwise, do not use libraries for machine learning methods. When in doubt, email me.

In order to be able to take the final exam for BLG454E you have to have a **weighted average score of 40 (over 100) for midterm, homeworks and term project.** Otherwise you will get a VF from the course.

DO NOT USE ANY BUILT-IN FUNCTION OR LIBRARY, WRITE YOUR OWN CODES. OTHERWISE YOU WILL GET 0(ZERO) GRADE.

MATLAB ASSINGMENTS

You will use the following dataset given for this hw. Please check the website carefully.
https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits

Q1) Examine the dataset (You have train and test set separetely in this homework). Classify given dataset using Neural Network (MLP). Write comments below codes you implemented. (You should evaluate classification results using 10-fold cross validation)

Q2) Reduce dimension of the data to 2 dimension using PCA. After reduction, classify test set using Neural Network (MLP) and compare your result with the results of the first and third question. Plot the result, analyse it and put it into your report. [You need to write down the PCA code

yourself, do not use a library pca() function. Do not use cov() function, but you may use eig() function] Is the result of the this question different than the result of the first and third question ? What do you think Which one is better ? Plot the result, analyse it and put it into your report.

Q3) Reduce dimension of the data to 2 dimension using autoencoder method. After reduction, classify test set using Neural Network (MLP) and compare your result with the results of the first and second question. Is the result of the this question different than the result of the first and second question ? What do you think Which one is better ? Plot the result, analyse it and put it into your report.

Note: Dimensionality reduction will be first applied to training dataset. (Eigenvalues will be computed on training set. Then they will be applied to test set. Similarly, autoencoder will be trained on the same training dataset.)