Data Science and Machine Learning Lab 2
Assignment 1 (Given Dec. 6, 2019; Due Dec 13, 2019)

Code and results must be uploaded in Moodle by midnight of the day it is due. If the question requires a textual response, you can create a PDF and upload that. The PDF might be generated from MS-WORD, LaTeX, the image of a handwritten response, or using any other mechanism. Late HW carries a penalty of 25% per day. Do not copy or use unfair practices

---

To get started, we will do an assignment that will refresh some of the things you may have already done before We will classify handwritten digits using the classic MNIST dataset for the purpose.

We will use three different approaches. In the first, you will use a random forest (of decision trees); in the second you will use a multi-layered feedforward neural network with 1 hidden layer (MLFFNN), and in the third you will use a convolutional neural network (CNN).

(30 points) Tabulate and provide the result you obtain with each approach (you can tweak the hyper-parameters so that you get the best possible results from each approach)

(20 points) Comment on the performance (training and generalization) performance you get from each of the methods

(20 points) Generate the confusion matrix to ensure there is no structure in the error that is being made

(30 points) For the two feed-forward nerual networks (MLFFNN and CNN), visualize the embedding as a "digit cloud" that you obtain using tSNE or PCA to project to 2-dimensions.