MA 721: Project 1 (due Monday Oct. 30, 2017)

Write a Keras/Theano/Tensorflow program to implement a fully connected MLP to solve the classification problem for the modifications of the MNIST dataset listed below. The original MNIST dataset contains 60000 training and 10000 testing images. (Some may divide the 60000 training set as 50000 training and 10000 validation set). For each of the modified datasets, train your neural network and present your best results in a table that includes the minibatch size, the number of hidden layers, the numbers of hidden units given as a listing, the total number of parameters, the optimization methods used, the learning rate, the initializer, the regularization if used, the number of training epochs, the final training accuracy, the final testing accuracy, and the ratio of the test accuracy to the 10th root of the product of the total number of parameters and the number of epochs. This ratio will be used to score your results. Also, plot the convergence curve of the loss function, the training accuracy and the testing accuracy against the training epochs.

1. Training set: the first 5,000 of the original training images; Testing set: same.

2. Use the full MNIST training and testing sets, but add 745 to all pixel values of the images. (The resulting data has entries ranging from 745 to 1000 and, after scaling, may be thought of as brighten images.)

3. Training set: take the first 40,000, the next 40,000 starting from the 10,001st, and the last 40,000 of the original training images, line them up in three rows, and then concatenate three images in each column to make 40,000 three digit images; Testing set: concatenate in the same way the first 8,000, the next 8,000 starting from the 1,001st, and the last 9,000 of the original testing images. The output will be ordered three digits that may be labeled as numbers between 0 and 999.

4. Training set: take the first 50,000 of the original training images and line up with the last 50,000 training images in two rows; then add the pixel values of two images in every column to make 50,000 samples; Testing set: add in the same way the first 9,000 of the original testing images with the last 9,000 testing images. The output will be unordered two digits that may be labeled as numbers between 0 and 54.

Notes:
1. No convolution layer is allowed for this project.
3. Consult Section 11.4 for some practical advices.
2. Submit your codes and results through canvas in a single PDF file.