Deep Learning: Theory and Practice

Homework # 1
Due Date: 24 February 2019

Submission Instructions:

- Solutions to question 1 and 2 can be done either on pen-paper, MS WORD or LaTeX. Submit the pdf file of the scanned papers/latex generated pdf or the WORD document.

- For question 3, document the results with plots, comments, etc, and submit it along with the python codes.

- Put all the files in a folder, tar/zip the folder with your credentials in folder name as: ***assignment#_FirstName_LastName_regID***, eg. assignment1_abc_xyz_123

- Upload the tar/zip folder in your Google drive, right click on the folder and share it with **deeplearning.cce2019@gmail.com**

- The assignment folder can be updated till the deadline by sharing the latest tar/zip folder (with exactly same folder name).

Questions:

1. Evaluate the matrix derivative

$$\frac{\partial}{\partial \boldsymbol{A}} \operatorname{tr}(\boldsymbol{AB})$$

   where $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices of shapes $n \times m$ and $m \times n$ respectively.

   (10 **points**)

2. Find the maximum likelihood based update equations for logistic regression using Gradient Descent algorithm for both 2 class and for $K > 2$ classes.

   (40 **points**)

3. The MNIST dataset consists of 60,000 training images of handwritten digits (0-9). The test data consists of 10,000 images of digits. All images are of resolution 28×28. Reshape the images into vectors of size 784, and implement the logistic regression in python for digit classification. Split the train dataset into 50,000 images for training and 10,000 images for validation. Run 15 epochs of Stochastic Gradient Descent training and plot the Log likelihood and classification accuracy on train and validation sets as a function of epochs. Report the test accuracy using the weights learnt after the last epoch.

   Comment on the influence of learning rate and batch size by performing the above experiments using:

   (a) Four different choices of learning rate 0.001, 0.01, 0.05, 0.1
   (b) Four different choices of batch size 1, 32, 128, 1024.

   (50 **points**)

   One link to download the dataset is :
   http://yann.lecun.com/exdb/mnist/