

# Deep Learning : Theory and Practice

*Homework # 1*

Due date: March 5, 2018

1. Show that

(a)

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^T \mathbf{A} \mathbf{x} = (\mathbf{A} + \mathbf{A}^T) \mathbf{x}$$

(Points 5)

(b)

$$\frac{\partial}{\partial \mathbf{A}} \text{tr}(\mathbf{A} \mathbf{B}) = \mathbf{B}^T$$

(Points 5)

2. What are the different approaches to Machine learning in terms of classification settings. Enumerate the difference between Generative modeling and discriminative modeling.  
(Points 5)

3. Find the maximum likelihood based update equations for logistic regression using Gradient Descent algorithm for both 2 class and for  $K > 2$  classes.  
(Points 10)

4. Implement the logistic regression in python on the MNIST dataset with the following changes

- Use the validation set from the training data set and not from the test set. The split should be 50000 for training, 10000 for validation. The original testset of 10000 samples will be used in the final testing only.
- Run for 15 epochs of SGD training and measure the performance on validation (for each iteration) and test (on the last iteration).

Based on the above modifications, investigate the following experiments

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

What is the influence of the above parameters on the validation and test accuracy.  
(Points 25)