ECSE 4965/6965

Introduction to Deep Learning

Program Assignment #1,

Due date: 4pm, Feb. 9th

In this programming assignment, you will implement techniques to learn a linear regressor that takes an input vector **x** of 10 dimensions and outputs a scalar *y*. Specifically, this assignment involves learning the regression parameters $\Theta=(\mathbf{w},w_0)^t$, where **w** is a 10x1 vector and w_0 is a scalar, given the training data $\mathbf{D}=\{\mathbf{x}[m], y[m]\}$, m=1,2, …, 10,000. **D** is stored in a Mx11 matrix (data.txt), where the first 10 columns represent input **x** and the last column represent *y*. You will implement the following methods to learn the regression parameters Θ based on the mean squared loss function.

a. Implement in Tensorflow the closed form gradient-based solution to solve Θ analytically. Write down the equation used as well as the computed values for Θ.

b. Implement in Tensorflow the gradient descent method to solve for Θ iteratively using all data in **D**. Initialize Θ to 0.01 for all elements of Θ. Try different learning rates between 0.001 an 0.005. Write down the estimated values for Θ as well as plot the loss function value change as a function of iteration number *t*.

c. For graduate students taking this class at 6000 level, implement in Tensorflow the Stochastic Gradient Descent method with a batch size set between to 20 and 10. Give the estimated values for Θ, the selected batch size, and plot the loss function value as a function of iteration number *t*.

Do not use Tensorflow's existing gradient descent and stochastic gradient descent functions for tasks b and c. Submit your Tensorflow code via LMS, along with the required outputs as specified above.