Homework #3

MTH 9899 Baruch College
DATA SCIENCE II: Machine Learning

Due: May 9, 2018 - 18:00

Notes

- Code for this **MUST** be written in **Python 3.x** as a Jupyter Notebook using PyTorch. Please only use CPU training to ensure it's easy for us to recreate.

Problem 1 In homework 2, we generated a sample dataset and built a regression tree. Take this same dataset and fit it with a neural network using the PyTorch library. Please choose reasonable values for the hyperparameters that aren't specified below.

1. Try using a single hidden layer with a tanh activation functions and a linear output activation. Use minibatches of 100 rows at a time and graph the training and test error across epochs.
2. Repeat the previous trial using ReLu instead of tanh for the hidden activation including graphs of training and test error. Which version is faster to converge?
3. Staying with ReLU as the hidden activation function, add a 2nd hidden layer and graph training and test errors across epochs. Does it take more or less time to converge and why? Does it produce a better result than a single layer and why?