

Columbia University  
IEOR4742 – Deep Learning for OR & FE (Hirsa)  
Assignment 2 – Due 8:40 am on Tuesday Oct 15th, 2019

**Problem 1 (Impact of different activation functions and optimization on learning):** In the code *example\_5\_hidden\_layer.jpynb*

- (a) what is the exact number of parameters we are trying to learn?
- (b) use different activation functions in your architecture
  - (i) 1st layer: sigmoid, 2nd layer: tanh, 3rd layer: reLU, 4th layer: sigmoid, 5th layer: leaky reLU
  - (ii) 1st layer: tanh, 2nd layer: reLU, 3rd layer: reLU, 4th layer: tanh, 5th layer: leaky reLU
  - (ii) architecture of your choiceand assess its impact on accuracy?
- (c) for part (b) keep the same architecture, just use different optimization routine and assess its impact on accuracy?

**Problem 2 (visualization of the lost function ):** Use sample code *example\_5\_layer\_interpolation.jpynb* and architecture and optimization routine in parts (b) & (c) of **Problem 1** to assess the loss function by interpolation, namely

- (a) impact of different architecture on the loss function surface
- (b) assessing the path traveled through the loss function having same architecture but using different optimization routine

**Problem 3 (CIFAR-10):** Repeat **Problem 1** for CIFAR-10 dataset.