Neural Networks, Back propagation
Indian Institute of Technology Hyderabad
Monday June 10th 2019.

1 Theory

1. Formulate the supervised learning problem for a multilayer perceptron model using the notation discussed in class for the sum of squared error loss function. Derive the locally optimal solution and describe the *backpropagation* algorithm.
2. Repeat the above for the cross entropy loss function.

2 Programming

1. Implement the back-propagation algorithm to learn the weights of a perceptron with 2 input nodes, 2 hidden nodes and 1 output node. Train your network to learn the following binary operations:
   (a) XOR
   (b) AND
   (c) OR

   Experiment with the number of training samples N and see how its affects performance. Add noise to the labels to generate more samples. Your code should make the number of nodes a configurable parameter.
2. A *multilayer perceptron (MLP) function* that accepts as input a vector, the number of hidden layers, the size of each hidden layer, the non-linear function, and the size of the output layer. This function should generate an output vector of the specified size. Generate the output with and without the *softmax* function applied to the output layer.
   (a) Classification: MNIST Digit Classification using Quadratic loss
   (b) Classification: MNIST Digit Classification using Cross entropy loss
3. Demonstrate that Neural Network Error surface with respect to weights is non-convex.