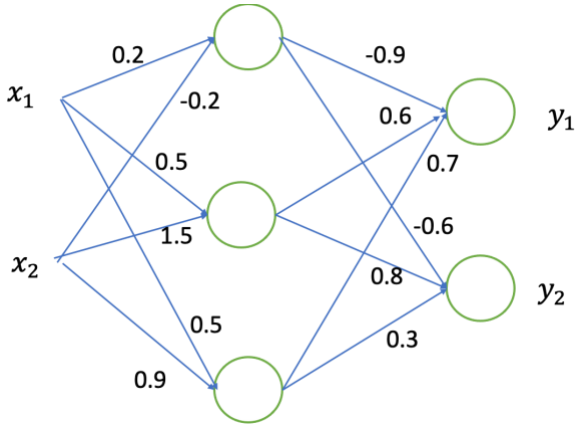
DSCI 6671: Deep Learning
Spring 2021

Assignment#1

This assignment consists of two parts: Part A requires written answers and Part B programming in Python.

Part A

Q1. The following is a fully connected neural network with biases 0.5, 0.3, 0.1 for neurons (from top to bottom) in the hidden layer and zeros for the output neurons. ReLU is the activation function for the hidden units and softmax for the output units.



Consider the input,

$$X = \begin{bmatrix} 0.1 & -0.5 \\ -0.7 & 1.3 \\ 0.5 & -1.5 \\ 1.3 & 0.8 \end{bmatrix},$$

where each row corresponds to an observation, and associated target labels, $\boldsymbol{y} = [0\ 0\ 1\ 1]^T$. Perform matrix computation for the following questions.

a) Arrange the weights of each layer in a matrix and biases in a vector.
b) Compute the affine transformation of the input using weights and biases in the hidden layer.
c) Compute the output of the hidden layer considering ReLu as the activation function.
d) Compute the affine transformation of the 1st layer output in the output layer.
e) The equation for softmax is $softmax(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$, Apply softmax to the output in (d) to compute the final output of the network.
f) Find the predicted class label, $\hat{y}$, for the input, X.
g) Use one-hot encoding to convert the target labels, y, to probabilities.
h) Compute accuracy.
i) Calculate the average cross-entropy loss between predicted and target class probabilities.

Part B

See the accompanied jupyter notebook