Randy Delacruz CISC 3440

Project 3

1. There is 1 neurons in the hidden layer. There is 1 neuron in the output later. The dimension is [0] for both the hidden and output layer. Network diagram located in Network\_Diagram.png file.

2) h1 = (x1 \* wh1) + (x2 \* wh1) + (x1 \* wh2) + (x2 \* wh2) + bh

3)o(h) = h1 \* w0 + b0

4)gradient\_

gradient\_bo =

gradient \_wh1

gradient\_wh2

gradient\_bh

5)wh1 =

wh2 =

wo=

bh=

bo=

6) Even with 300 epoch training, there was a difference in predictions for the middle samples [0,1] and [1,0]. The model predicted a value of 0.25 for the middle samples. There was also a difference in the final sample [1,1], with a value of 0.49. I believe this happens because we aren’t using activation values for the hidden and output layers.

7) I think adding activation functions to the hidden and output layer could result in better predictions. A step function on the output layer that changes predictions close to 0, to 0 and predictions close to 1 to 1 can work.

Task 6

The maximum reductions in latent dimensionality I achieved was 10. I also change the batch size to 200. With these 2 changes, I received an error score of 2.84078 on the test data. This error score was in the same range of score that I achieved with no changes to any of the hyperparameters. No changes to the epochs or learning rate was necessary because the error loss was near 0 very early on in the training and perhaps because there was sufficient training data (237,395 samples). There was no difference between the original construction of hand models and the reconstructed.