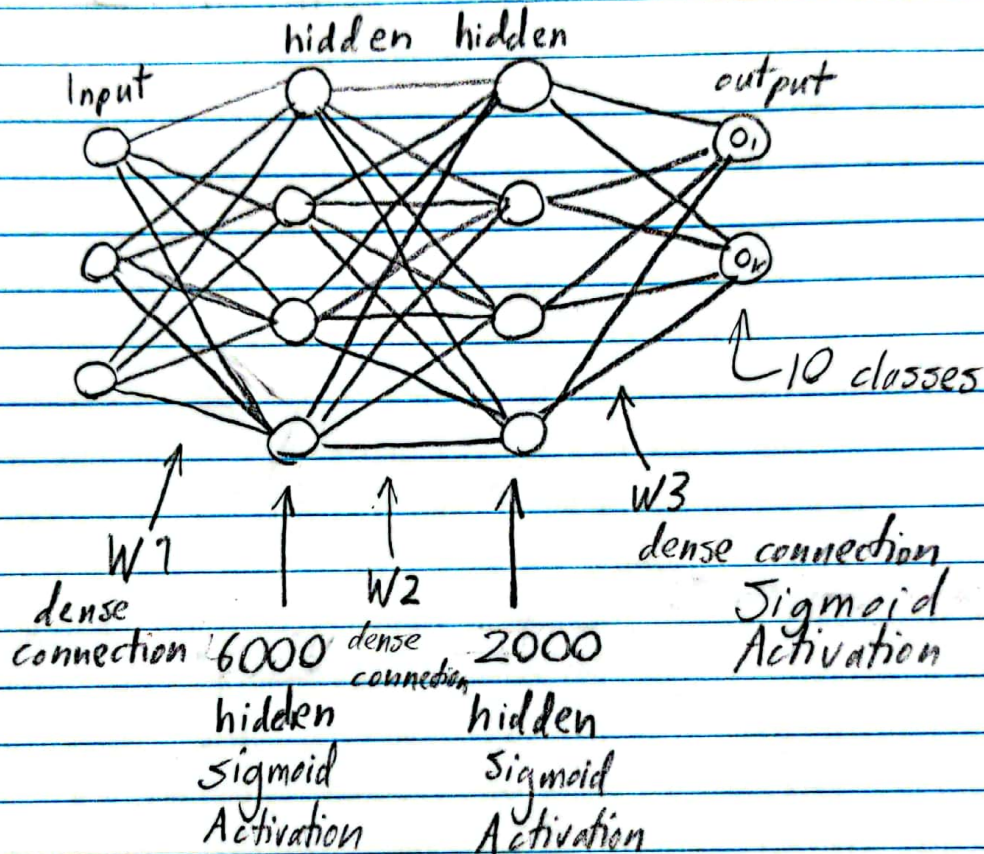


Problem 1 Code

Structure



Hyper parameters

learning Rate: 0.05
Optimizer: Gradient Descent
Epochs: 100
hidden layer 1: 6000
hidden layer 2: 2000
Activations: Sigmoid
Batchsize: 25
Random weight initialization
based on gaussian distribution

training loss: 0.790850934154814

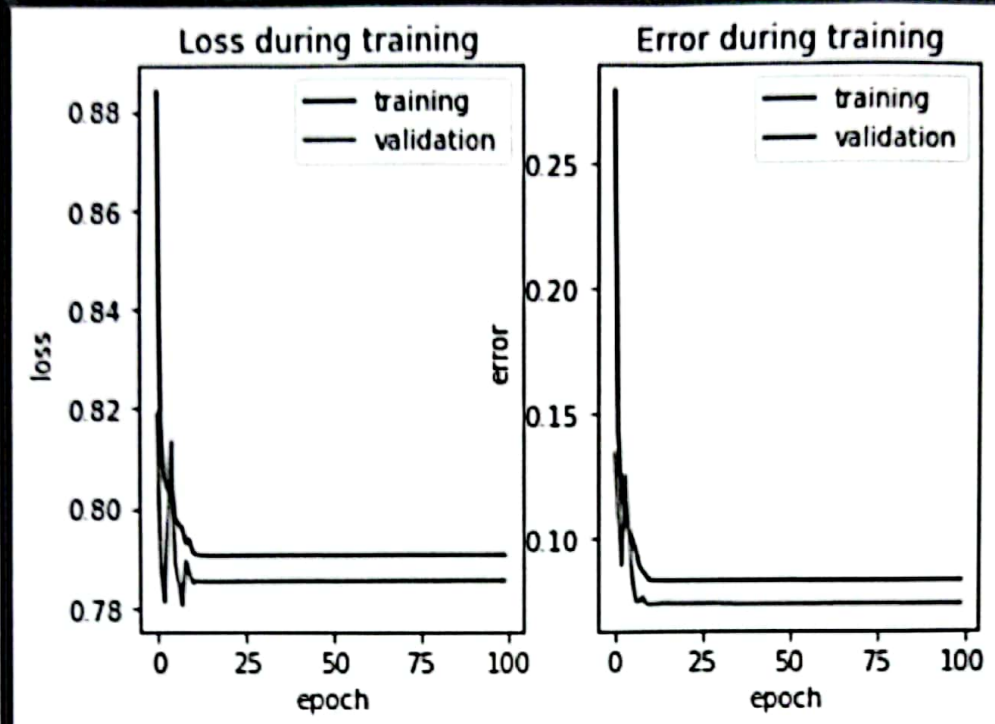
training error: 0.08298

validation loss: 0.7856630233304758

validation error: 0.0735

test loss: 0.7868565932217936

test error: 0.0738



✓ 8h 45m 38s completed at 8:18 AM

Final Accuracy & Performance

Test Accuracy: 7.38%

Test Loss: 0.7868

Iterations in development

- Iterated through learning rate with for loop from 0.01 to 0.1 and found 0.05 worked the best
- Iterated through hidden layer shapes & sizes finding large to small worked the best
- Final Model had accuracy of 7.38% in final iteration