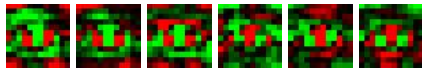


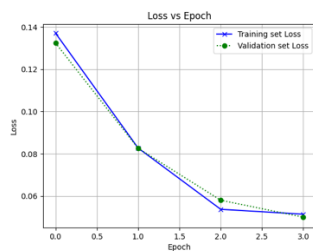
2 point - your source code for neural network training. Make us it's easy for the TA to find the key components: forward propagaion, backward propagaion, weight updates, loss calculation.

See Attached File

1 point - A visualization of the weights you learn for the neurons in the hidden layer of your tuned single-hidden-layer network using the framework function 'VisualizeWeights'



2 points - Tune the parameters until you have a single layer network with validation set accuracy greater than 85%. Report on the parameters you used. Include a chart with training and validation set loss on the y-axis vs epoch nuber on the x-axis. In 1-2 sentences indicate the approach you took to finding hyperparameters that work. Include a brief log of the Hyperparameters you tried.



Best parameters for single layer network:

Step size : 1.0

Convergence: 0.005

Hidden layer structure : [6]

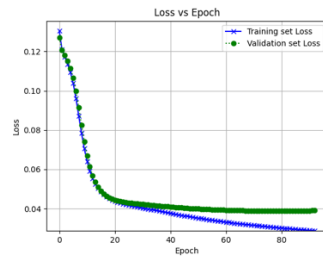
Validation set accuracy: 87.19%

My approach was to tune the hyperparameter hidden structure first to eliminate bias issues and then tune the hyperparameters convergence and stepsize to so that I could get the best accuracy. I did not consider the confidence intervals as part of this exercise.

Hyper Parameters tried:

1. sweeps['hiddenStructure'] = [[5], [6], [7], [8], [9], [10], [11], [12]]
2. sweeps['convergence'] = [0.1, 0.5, 0.01, 0.05, 0.001, 0.005, 0.0001, 0.0005]
3. sweeps['stepSize'] = [4.0, 3.0, 2.0, 1.0, 0.5, 0.1, 0.05, 0.01]

2 points - Tune the parameters until you have a two layer network with validation set accuracy greater than 89%. Report on the parameters you used. Include a chart with training and validation set loss on the y-axis vs epoch number on the x-axis. In 1-2 sentences indicate the approach you took to finding hyperparameters that work. Include a brief log of the Hyperparameters you tried.



Best parameters for double layer network:

Step size : 0.01

Convergence: 0.0001

Hidden layer structure : [7, 2]

Validation set accuracy: 89.66%

My approach was to tune the hyperparameter hidden structure first to eliminate bias issues and then tune the hyperparameters convergence and stepsize to so that I could get the best accuracy. I did not consider the confidence intervals as part of this exercise.

Hyper Parameters tried:

1. sweeps['hiddenStructure'] = [[2,2], [3,2], [4,2], [5,2], [6,2], [7,2], [8,2], [9,2]]
2. sweeps['convergence'] = [0.5, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001]
3. sweeps['stepSize'] = [4.0, 3.0, 2.0, 1.0, 0.5, 0.1, 0.05, 0.01]