

PA-1 Handwritten Digits Classification

CSE 574- Introduction to Machine Learning

Course Number : CSE 574

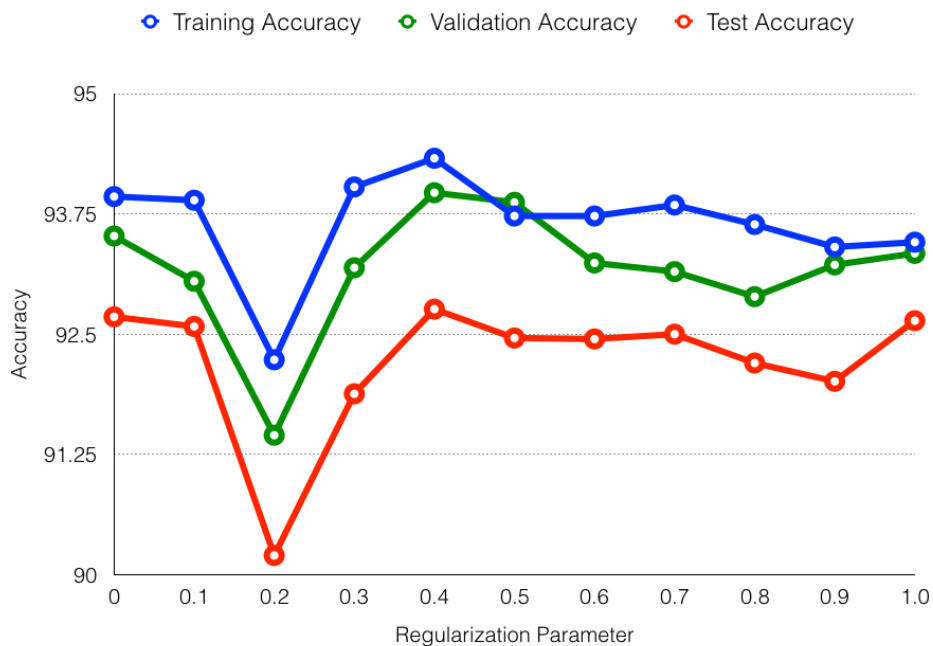
VINAY GOYAL

EXPERIMENT 1:

Finding Optimal value of Regularization Parameter(Lambda)

- Number of Hidden Units: 50
- Range of Lambda: 0-1
- For each step size of 0.1, we are calculating the Training Accuracy, Validation Accuracy, Test Accuracy and Time(seconds)

Regularization Parameter	Training Accuracy	Validation Accuracy	Test Accuracy	Time(Seconds)
0	93.93	93.52	93.68	86
0.1	93.892	93.05	92.58	81
0.2	92.234	91.45	90.2	78
0.3	94.03	93.19	91.88	77
0.4	94.328	93.97	92.76	81
0.5	93.728	93.87	92.46	80
0.6	93.728	93.24	92.45	85
0.7	93.842	93.15	92.5	87
0.8	93.64	92.59	92.2	86
0.9	93.406	93.22	92.01	81
1	93.456	93.34	92.64	78



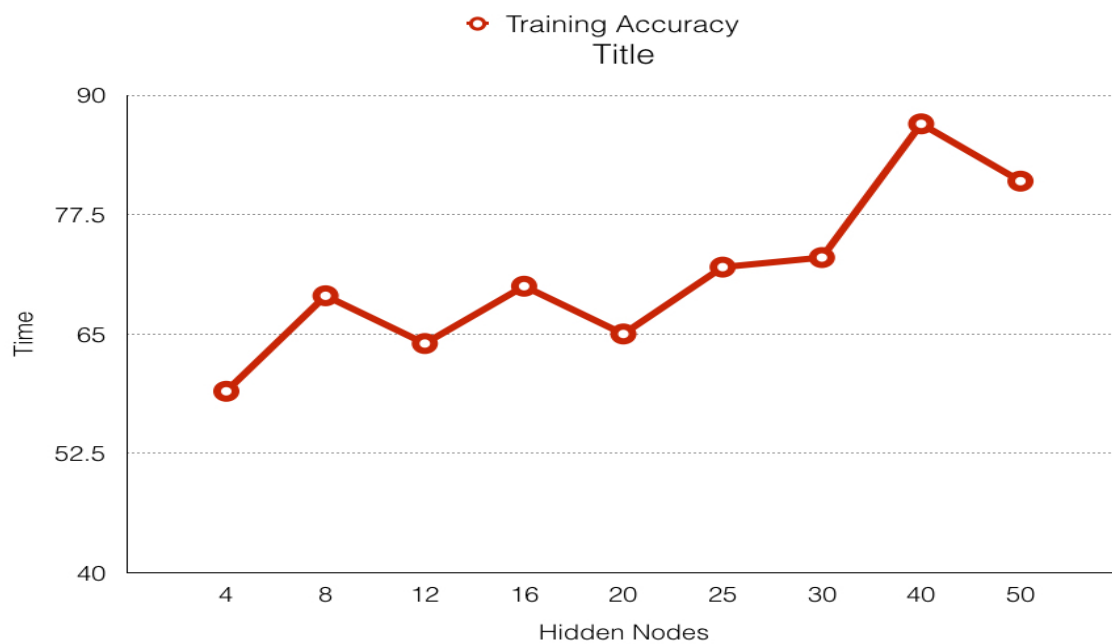
Observation: The accuracy of prediction gets lower with increasing values of the regularization parameter. This happens because higher lambda values give more importance to the weights at the expense of error function, while the reverse happens for lower values of regularization parameter.

Experiment 2:

Finding optimal number of hidden units

- Regularization parameter(constant): 0.4
- Range of Hidden Units: 4-50

Hidden Units	Training Accuracy(%)	Validation Accuracy(%)	Test Accuracy(%)	Time(Seconds)
4	25.902	25.91	19.03	59
8	83.696	83.96	83.36	69
12	91.214	90.52	88.75	64
16	92.612	91.86	88.86	70
20	92.34	91.60	88.7	65
25	92.508	91.84	90.42	72
30	93.89	93.48	91.37	73
40	93.068	92.52	91.4	87
50	94.328	93.97	92.76	81

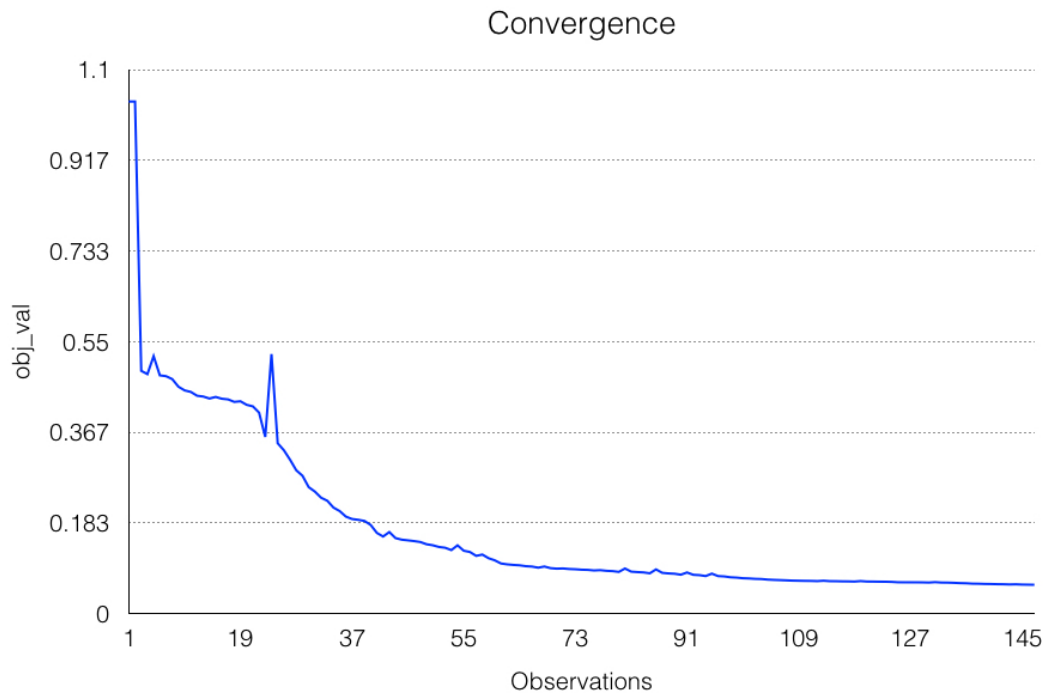


Observation: It can be observed that the time taken to train our neural network increases as the number of hidden units are increased. On the similar lines, the training accuracy also increases proportionally with the increase in the number of hidden units.

Experiment 3:

Observe the value of error function as the network gets trained.

- Regularization parameter(constant): 0.4
- Hidden Units: 50



Observation: The value of the error function is clearly decreasing as the number of iterations increases and the neural network gets trained.

Conclusion:

The number of hidden units and the regularization parameter have been chosen so as to achieve a balance between the training accuracy and the time taken by the neural network.