PA-1 Handwritten Digits Classification

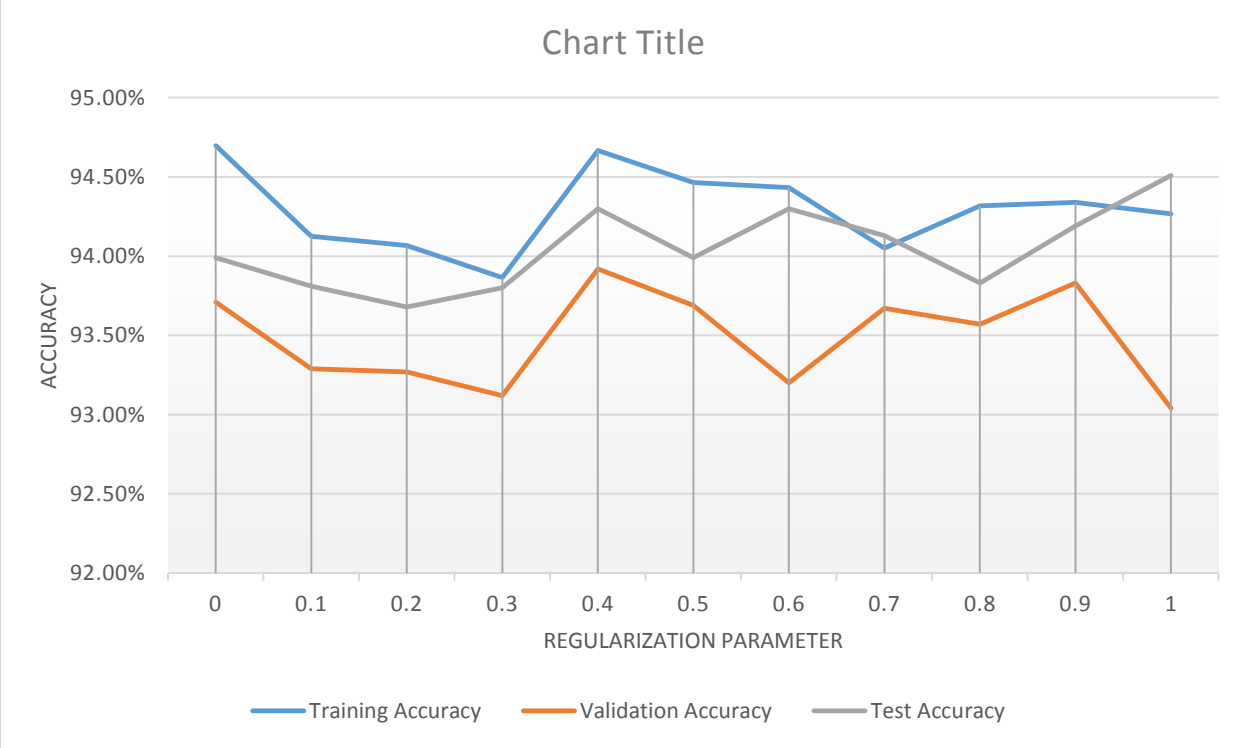
CSE 574 – INTRODUCTION TO MACHINE LEARNING

Susana Dsa, Nikhil Lopes, Avinash Pali (Team 8)
UNIVERSITY AT BUFFALO | CSE 574 INTRODUCTION TO MACHINE LEARNING

1. Experiment 1 - Finding optimal value of lambda (Regularization Parameter)

- Number of hidden units (constant): 50
- Range of lambda: 0 - 1
- Step Size: 0.1
- Data

Regularization Parameter	Training Accuracy	Validation Accuracy	Test Accuracy	Time Taken (Seconds)
0	94.696%	93.71%	93.99%	141.157
0.1	94.126%	93.29%	93.81%	133.987
0.2	94.066%	93.27%	93.68%	140.661
0.3	93.864%	93.12%	93.80%	140.219
0.4	94.666%	93.92%	94.30%	146.911
0.5	94.464%	93.69%	93.99%	147.381
0.6	94.432%	93.20%	94.30%	145.024
0.7	94.050%	93.67%	94.13%	132.056
0.8	94.318%	93.57%	93.83%	143.444
0.9	94.340%	93.83%	94.19%	136.387
1.0	94.266%	93.04%	94.51%	150.200

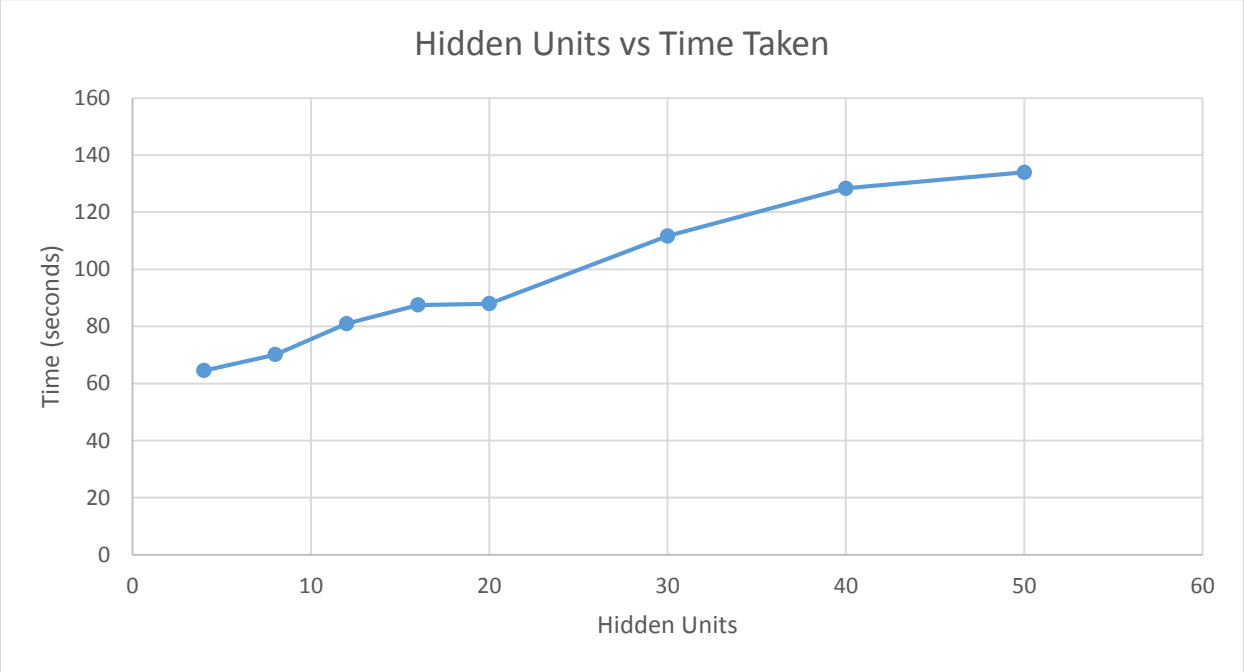


- Inference –The prediction accuracies get lower with increasing values of the regularization parameter due to the under-fitting problem because higher lambda values give more importance to the weights at the expense of the error function while the opposite follows for lesser vales of regularization parameter.

2. Experiment 2 – Finding optimal number of hidden units

- Regularization Parameter (constant): 0.1
- Range of hidden units: 4 - 50
- Data

Hidden Units	Training Accuracy	Validation Accuracy	Test Accuracy	Time Taken (Seconds)
4	64.90%	63.91%	64.75%	64.513
8	89.688%	88.72%	89.04%	70.121
12	91.450%	91.11%	90.66%	80.967
16	92.500%	91.21%	92.10%	87.508
20	93.956%	93.41%	93.40%	87.952
30	93.962%	93.32%	93.71%	111.675
40	94.178%	93.20%	93.96%	128.358
50	94.126%	93.29%	93.81%	133.987



- Inference – We can infer that the time taken to train increases as the hidden units are increased while also increasing the training accuracy.

3. Conclusion – We can infer that the hidden units and regularization parameters are chosen such that a balance is reached between the time taken (resources) and the training accuracy (results) reached.