

CSE 574 - Introduction to Machine Learning

Programming Assignment 1 *Handwritten Digits Classification*

Submitted by :

- 1) Madhur Gupta (UBIT Name : madhurgu)
- 2) Mohit Arora (UBIT Name : marora3)
- 3) Abhijit Pattanaik (UBIT Name : abhijitp)

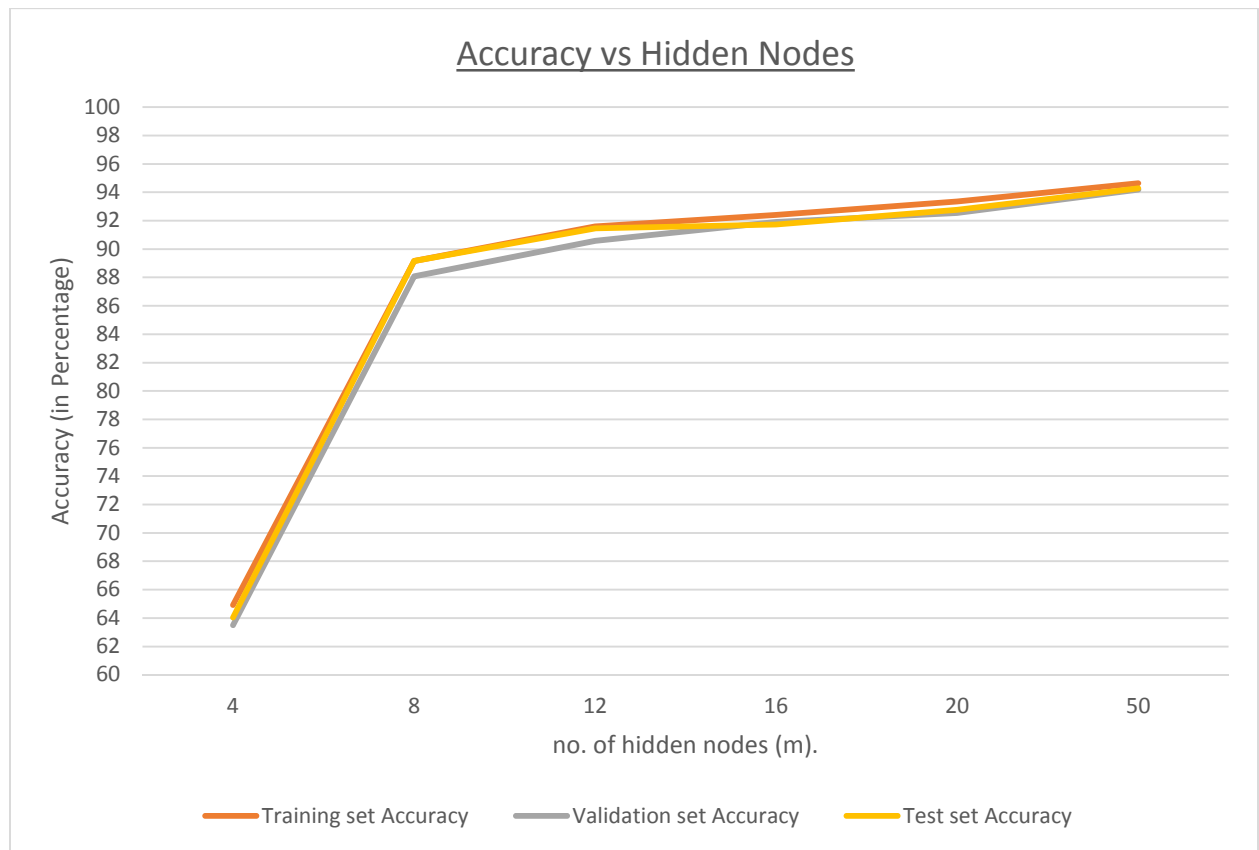
For Multilayer Perceptron Neural Network, we performed following two experiments for measuring accuracy for various data:

1) Accuracy vs Hidden Nodes (for fixed $\lambda=0.5$) :

From various runs on the given “mnist” data, we received following test results:

m	Training set Accuracy	Validation set Accuracy	Test set Accuracy
4	64.928	63.49	64.03
8	89.166	88.07	89.17
12	91.604	90.59	91.46
16	92.396	91.91	91.74
20	93.364	92.53	92.77
50	94.628	94.18	94.27

Using above test results below graph was plotted



From the above graph (or the table), we can see that we obtained maximum value of accuracy for all types of data (i.e. training, validation and test) was obtained for $m = 50$.

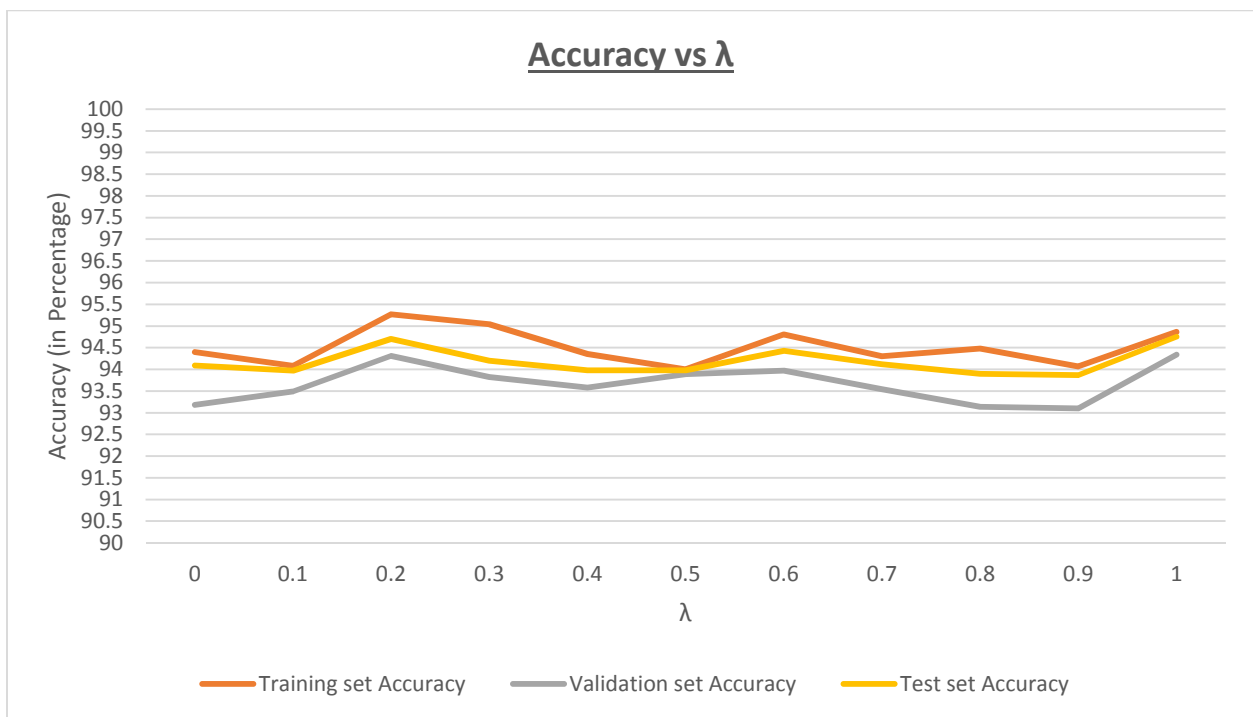
We also observe from above graph that after certain value of m , change in accuracy is not very huge as we increase m ,

2) Accuracy vs λ (for fixed hidden nodes = 50)

From various runs on the given “mnist” data, we received following test results:

λ	Training set Accuracy	Validation set Accuracy	Test set Accuracy
0	94.396	93.18	94.09
0.1	94.084	93.49	93.97
0.2	95.268	94.31	94.7
0.3	95.042	93.82	94.2
0.4	94.352	93.58	93.98
0.5	94.004	93.89	93.98
0.6	94.802	93.97	94.43
0.7	94.306	93.54	94.12
0.8	94.48	93.14	93.9
0.9	94.07	93.1	93.87
1	95.124	94.15	94.43

Using above test results below graph was plotted



From above graph, we chose $\lambda = 1.0$ At $\lambda=1.0$ we got the maximum accuracy for all types of data (i.e. training, validation and test).

We incremented the value of λ from 0.0 to 0.1 at interval of 0.1 and we got the optimum accuracy at $\lambda=1.0$ for all types of data . Based on the output of the gradient, λ corrects itself to avoid problem of overfitting.