

Machine learning Assignment – I

CSE 574

Submitted by

Kaushikraj Palanichamy (Kaushpal)

Sankara Vadivel Dhandapani (sankarav)

Sathish Kumar Deivasigmani(sdeivasi)

Table of Contents

Preprocessing:.....	3
Feature selection:	3
Choosing the right hidden node:	3
Training accuracy vs hidden node vs lambda	4
Test accuracy vs hidden node vs lambda.....	5
Training time vs hidden nodes vs lambda.....	6
Choosing the right regularization parameter:	6

Preprocessing:

In preprocessing train data is divided into two batches: train data and validating data. Train data and validation data have around 50,000 and 10,000 samples respectively with their truth values.

Feature selection:

- 1) In the 28x28 image, pixels that are same across all the 60K samples were removed.
- 2) Maximum dimension of such image was found out. In our minst data, it was 20x20
- 3) All the other images have been resized to 20x20 using **bilinear** image resize
- 4) All the image points have been resized from 0-255 scale to 0-1 scale.
- 5) At the end the feature set had 400 values (20x20 excluding input bias)

Choosing the right hidden node:

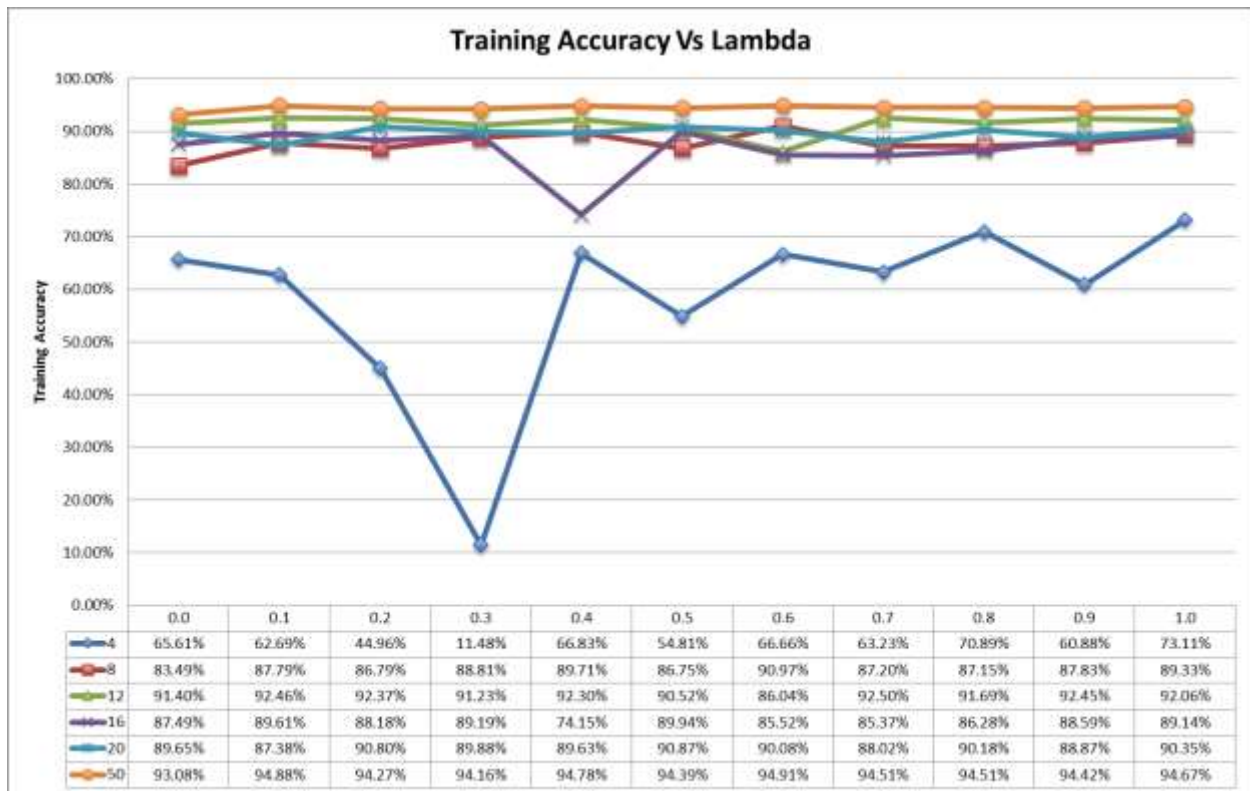
After running algorithms for hidden nodes 4,8,12,16,20,50,100, we found out that in the when we increase the number of hidden nodes slightly there is no visible change in the accuracy or time taken to converge. But when we jumped from 20 to 50, the average accuracy for all λ got increased significantly.

When number of hidden nodes got increased from 50 to 100, there was a small increase in the accuracy as shown in the figure but the time taken to converge was increased by an order of magnitude.

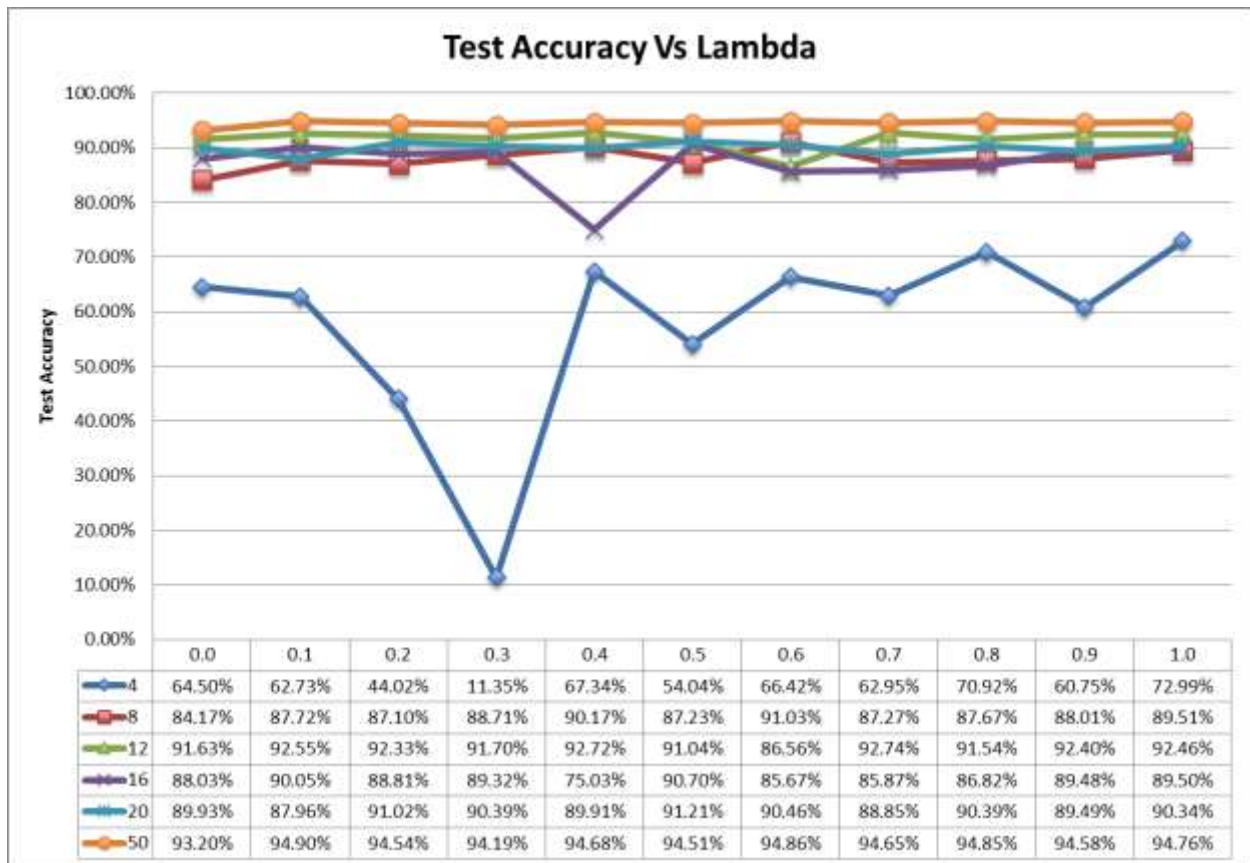
If training time is a constraint, we can choose number of hidden nodes = 12 and $\lambda = 0.4$ as it clearly outperforms its peers when comparing the time taken and also nearly as accurate as results of higher hidden nodes. If training time is not a constraint, then 50 hidden nodes and $\lambda = 0.1$ gives us the best results in test data.

	Hidden nodes	Lambda
With time constraint for training	12	0.4
Without time constraint	50	0.1

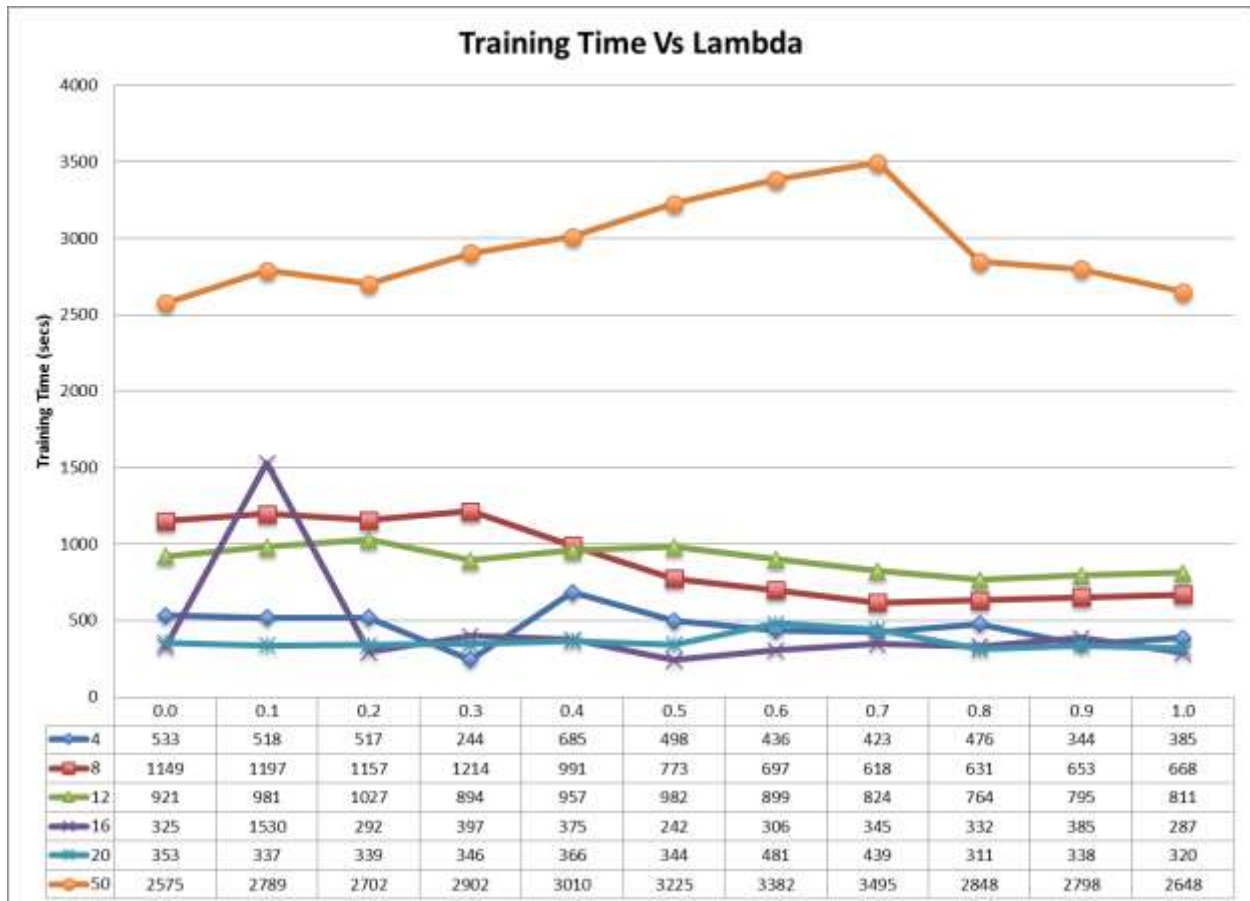
Training accuracy vs hidden node vs lambda



Test accuracy vs hidden node vs lambda



Training time vs hidden nodes vs lambda



We chose the number of hidden nodes as 50

Choosing the right regularization parameter:

We choose $\lambda = 0.1$ as it clearly gives higher accuracy in both testing, training and validation data.