CSE 574: Programming Assignment 1
Handwritten Digits Classification Using Neural Networks

Ajay Kumar Davuluri, UBIT ID: adavulur, Person No: 50168851
Lalith Vikram Natarajan, UBIT ID: lalithvi, Person No: 50169243
Srinath Goud Vanga, UBIT ID:srinathg , Person No:50169176

March 3, 2016

1 Objective

Implement a Multilayer Perceptron Neural Network and evaluate its performance in classifying handwritten digits to understand

- The working of a Neural Network.
- Use Feed Forward, Back Propagation to implement Neural Network.
- Setup a Machine Learning experiment on real data.
- Role of regularization in the bias-variance tradeoff.

2 Dataset

The MNIST dataset [1] consists of a training set of 60000 examples and test set of 10000 examples. All digits have been size-normalized and centered in a fixed image of 28 x 28 size. In original dataset, each pixel in the image is represented by an integer between 0 and 255, where 0 is black, 255 is white and anything between represents different shade of gray.

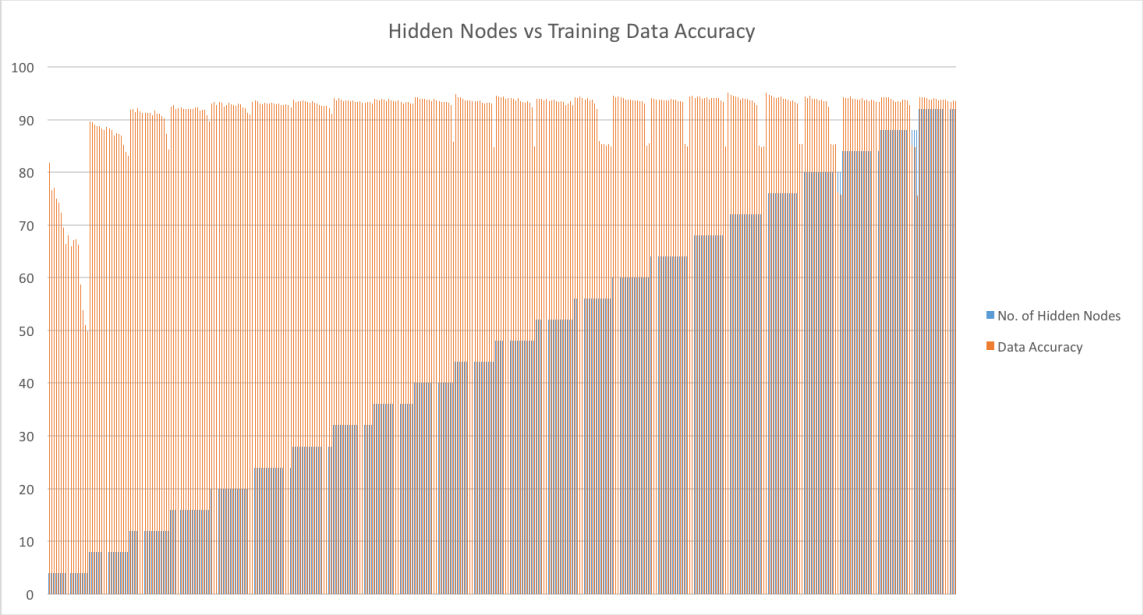
3 Implementation

- Preprocessing includes selecting features to reduce the computational overhead and shuffling data to ensure randomness in the experiment.
- nnObjFunction() is using back propagation to learn the weights. The error function that is used is squared loss error function.

- scipy.minimize is used to minimize the error and learn the weights (The return value of nnObjFunction is fed as an input to this function to help with the gradiance).
- In the nnPredict() function we are using the learned weights in the minimize function to predict the label. We obtained a 10 bit vector and the maximum value's index is chosen as the predicted label.

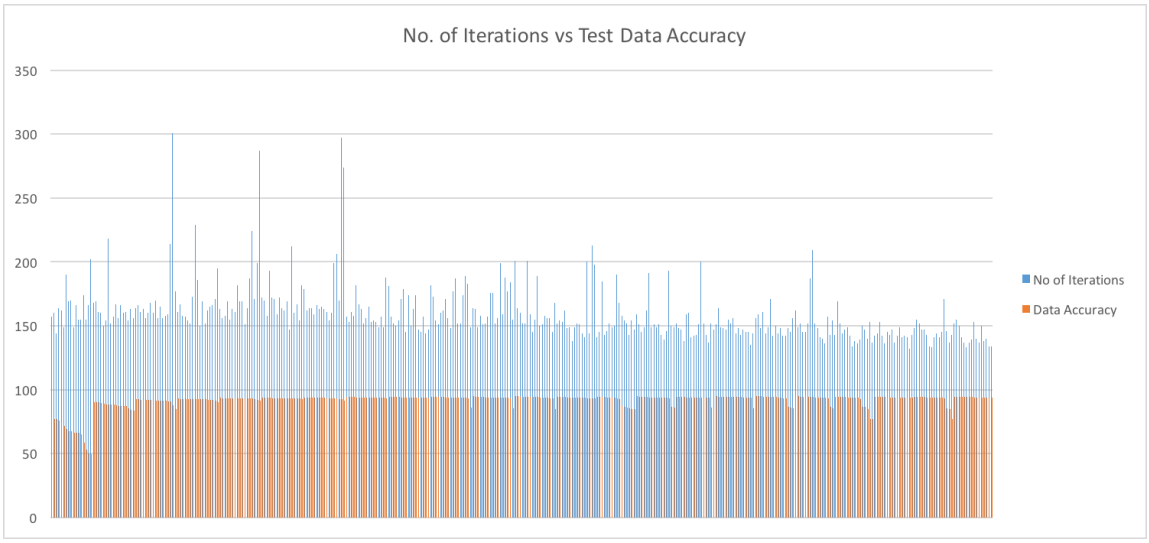
4 Observations

4.1 Hidden Nodes and Accuracy



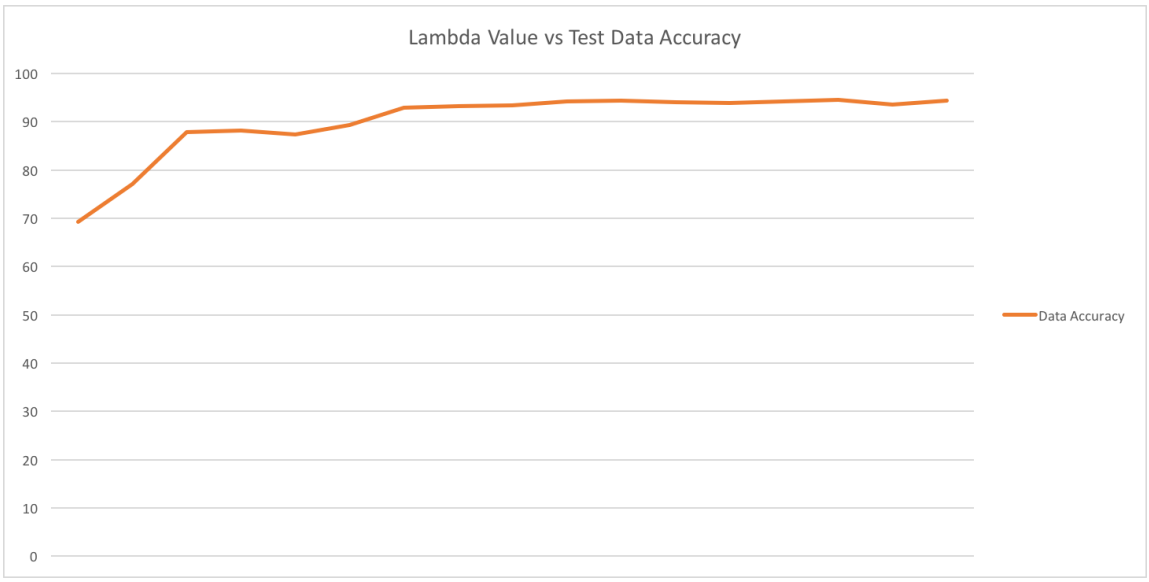
The above graphs show us that as the no. of hidden nodes increases, we see a marginal rise in the accuracy of the data sets. This is more pronounced when the no. of hidden nodes are greater than 32. The optimal no. of hidden nodes that gave a good accuracy was 76 nodes.

4.2 Iterations and Accuracy



The above graph shows us that as the no. of iterations is not really a factor in increasing the accuracy of the data sets. They vary wildly.

4.3 Lambda Value and Accuracy



The above graph shows us that there is a direct proportion between λ and the accuracy of the data sets for values less than 0.5 but as the λ rose, the changes were not really evident. The observed optimal value of lambda that gave the best accuracy for this dataset is 0.7

4.4 Role of Feature Selection

Using feature selection, we were able to remove 67 rows from the data set which were redundant, this helped us achieve reduced running time.