CSE 574 INTRODUCTION TO MACHINE LEARNING

PROGRAMMING ASSIGNMENT - 1

HANDWRITTEN DIGITS CLASSIFICATION USING NEURAL NETWORKS

Group 31

Anudeep Bulla 50168954

Vallabhaneni Susmitha Chowdary 50169236

Narmadha Viswanathan 50169758

INTRODUCTION

The objective of this project is to implement a Multilayer Perceptron Neural Network and evaluate its performance in classifying handwritten digits. Feed forward propagation, Backward propagation were used for implementation of the neural network. We have evaluated the performance of the neural network by changing the values of hidden nodes, regularization coefficients.

IMPLEMENTATION

The dataset considered consists of a training set of 60,000 examples and a test set of 10,000 examples. 50,000 training examples were used for training and the rest were used for validation. Feature selection was performed to remove features with values that are exactly same for all the data points in the dataset. Feed forward propagation is performed by feeding the input feature to the hidden units and the outputs are given to the sigmoid function. The error was calculated and transmitted backwards by performing backward propagation. Regularization is done to avoid overfitting problem.

OBSERVATION

The graphs are plotted for different values of regularization coefficients and hidden nodes.

The accuracy changes considerably with change in lambda value and the number of hidden nodes. The lambda values can be varied between zero which indicates that there is no regularization to 1 in steps of 0.1, 0.2 and so on and the changes in accuracy can be observed. An optimal value of lambda which gives a good accuracy for the training as well as the testing set has to be chosen. It can be observed that the accuracy values are different for the training, testing and validation data for different values chosen for lambda and hidden nodes.

PARAMETERS :

After observing with several values of lambda and hidden nodes count, few inferences could be made. Firstly, having small hidden node count i.e 0-10 the accuracy changes sharply also, with a count of 10-20 we can observe a steady increase in the accuracy. By having the hidden node count above 20 we can observe that the accuracy changes slowly to reach a point where the accuracy for different values like 40,50 42 is around 93%. Any further increase in the hidden node count doesn't seem to improve the accuracy by an amount justified by the increased training time. Thus, by varying the count between 20-40 we chose the optimal count to be **38.**

In order to chose the lambda, for any particular hidden node count the accuracy is found to be similar between 0 and 0.4. Any further change in lambda is degrading the accuracy. Thus, by checking within these ranges we chose the optimal lambda value to be **0.3 for hidden nodes 38**.

In Fig 1.1 we have obtained the results for different values of hidden nodes and lambda values.

NUMBER OF HIDDEN NODES	λ(LAMBDA VALUES)	ACCURACY			
		TRAINING DATA	TEST DATA	VALIDATION DATA	AVERAGE
4	0.1	70.90	70.65	70.45	70.66
8	0.1	89.46	89.34	89.51	89.43
12	0.1	91.35	91.13	91.19	91.22
16	0.1	92.46	92.3	91.53	92.13
20	0.1	93.23	93.26	92.78	93.09
24	0.2	93.22	93.16	92.80	93.06
24	0.3	93.76	93.67	93.23	93.55
28	0.1	92.91	92.79	92.64	92.78
28	0.2	94.08	93.75	93.33	93.72
28	0.3	94.022	93.61	93.277	93.63
30	0.3	93.72	93.7	93.377	93.59
38	0.3	94.65	94.32	93.84	94.27
40	0.3	94.070	94.0	93.74	93.93
42	0.3	94.10	93.99	93.48	93.85

Fig 1.1

VISUALIZATION - GRAPHS

Fig 2.1 - Accuracy vs (# Hidden nodes, Lambda)

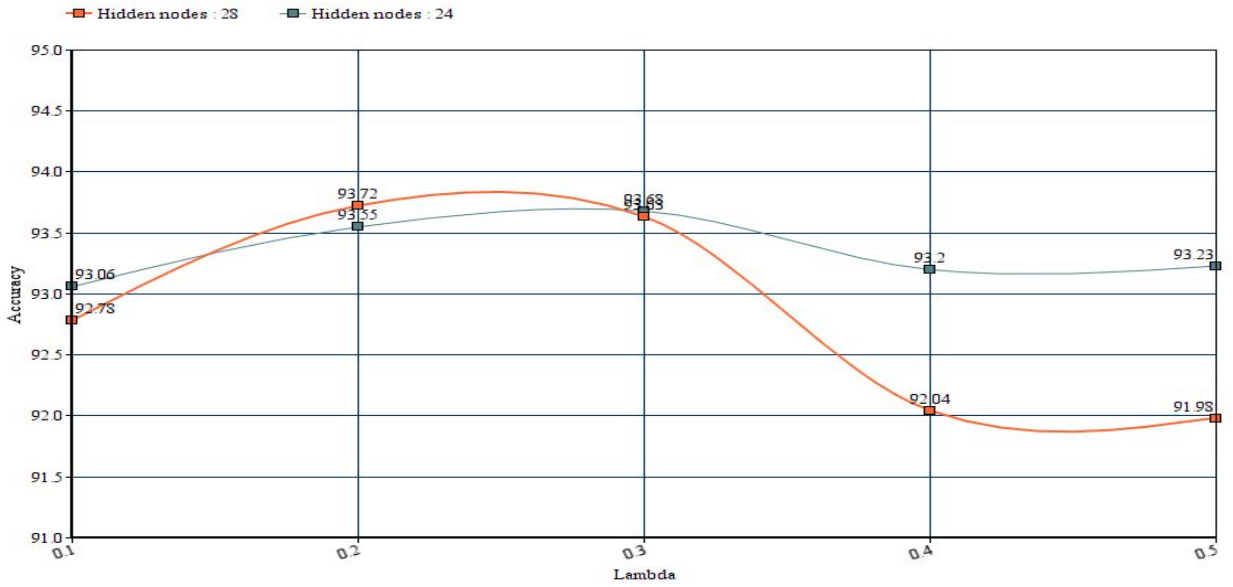

Fig 2.2 - Accuracy vs Lambda (Constant Hidden node count)

ANALYSIS AND TAKE-AWAYS

- Accuracy of the network with hidden node below 10 drops below the 90% mark, and any number of hidden nodes greater than 10 throw up an accuracy of at least 90%, strongly suggesting that *the network needs at least as many hidden nodes as there are output nodes*.
- When regularization is applied, the resulting *accuracy usually peaks within the 0-0.3 levels for the lambda value*, and then drops off.
- The *time to train the neural network is proportional to the number of nodes in the hidden layer,* as can be expected due to the calculations over increasing sizes of matrices. The lambda value doesn't seem to affect the time taken.

CONCLUSION

Multilayer perceptron network to classify handwritten digits was implemented and the results were obtained. The performance of the network was evaluated and the maximum accuracy of **94.65** was achieved by setting the number of **hidden nodes** as **38** and the **lambda (λ) value** as **0.3**.