CSE 574

A PROJECT REPORT ON

Implementation of a Multilayer Perceptron Neural Network for Handwritten digits classification

Abdul Muqtadir Mohammed (5013-5042)

Sai Sri Harshita Vallurupalli (5013-4381)

Gitanjali Palwe (5013-4107)

Abstract:

The main objective of the project is to evaluate the performance of a Multilayer Perceptron Neural Network for classifying handwritten digits. It mainly covers implementation of a Multilayer Perceptron Neural Network using feed-forward propagation, gradient descent and Back-propagation of the error. This report evaluates performance of the neural network accuracy by regularization of neural network.

Implementation of the neural network:

Following steps were followed while implementing the neural network:

- **Dataset:** 60,000 training examples were loaded in a .mat file to train the neural network. Each training example is described by 784 features. 50,000 training examples were used for training and remaining 10,000 for validation.
- **Feature Selection:** We have done feature selection using standard deviation formula and included features that have standard deviation greater than 0.1.
- **Feed-forward pass:** The input features are applied to 50 hidden input nodes (where maximum accuracy of 96% was achieved) whose output goes to a sigmoid function. Maxiter value of 60 was used
- **Back-propagation:** The error in prediction was calculated and transmitted backwards by updating the weights.
- **Regularization:** To avoid over-fitting, regularization was incorporated with term lambda in error function.

Observations while choosing the hyper-parameters:

Accuracy Vs lambda

The value of lambda can significantly affect the results on the training and validation set. In particular, a model without regularization (lambda = 0) fits the training set well, but does not generalize. Conversely, a model with too much regularization i.e. high value of lambda does not fit the training set and testing set well. A good choice of lambda (e.g., lambda = 0.5) can provide a good fit to the data. Following graph demonstrated the change in accuracy of neural network as a function of lambda (λ) .We can observe that we get different accuracy values for the training, validation and test datasets at different values of lambda. As you can see from Fig 1.1, the lambda value=0.4 gives the highest accuracy for test data when compared to different values of lambda.

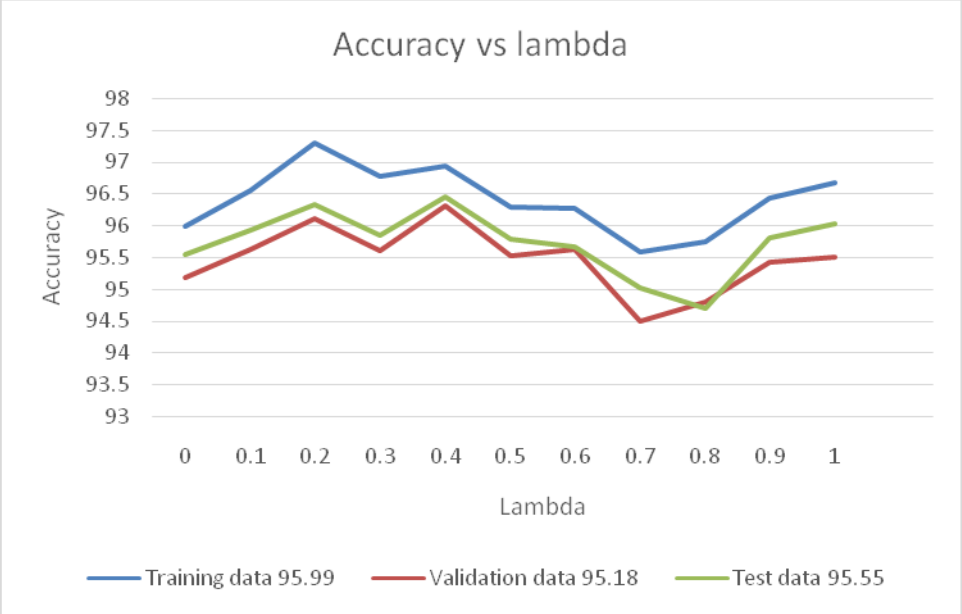

Fig1.1 Accuracy in classification VS lambda

Accuracy VS number of hidden units

We train the dataset taking different values of hidden nodes and choose wherever we get maximum accuracy. Following graph shows accuracy of neural network as a function of number of hidden units. As we vary the number of hidden nodes, we achieve maximum accuracy of **96% at 70 hidden nodes** across training data, validation data and test data. In general as we increase the number of nodes accuracy of network increases. As you can see, after 70 hidden nodes, the accuracy does not vary drastically and settles down.

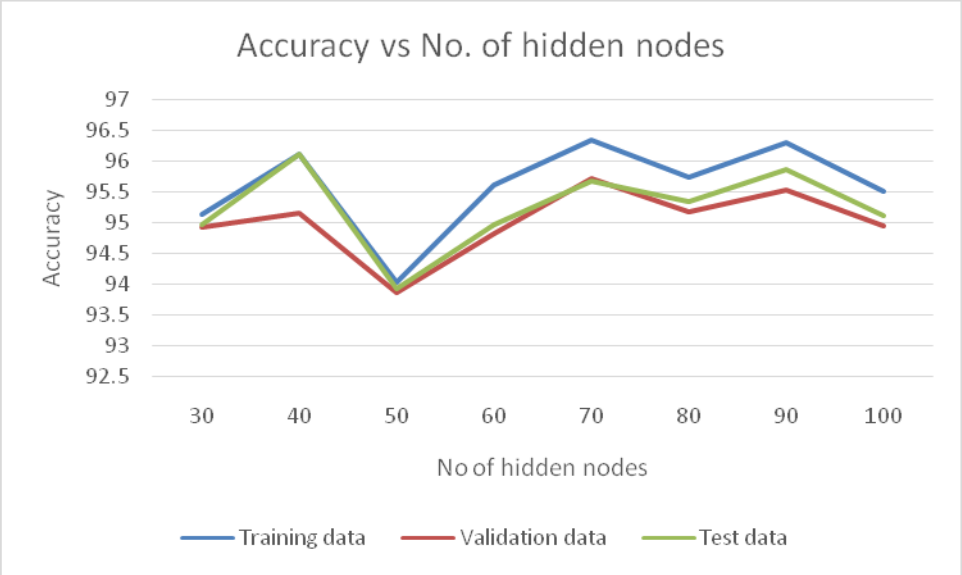

Fig1.2 Accuracy in classification vs Number of hidden nodes

Conclusion:

Thus we have successfully implemented the classification of handwritten digits by implementing the multi-layer perceptron neural network. We have evaluated the efficiency of the neural network by achieving an **accuracy of 96.45%** in classifying examples across the dataset (test data) using ***70 hidden nodes*** and $\lambda=0.4$. Also regularization of error function was carried out to achieve better accuracy.