CSE-574 Machine Learning

Programming Assignment 1

Handwritten Digits Classification

Spring 2017

University at Buffalo

Members:

Hrishikesh Saraf	5020 5927
Nikita Konda	5020 7717
Shreya Nimje	5020 6520

Overview:

In this assignment, we have implemented a Multilayer Perceptron Neural Network and evaluated its performance in classifying handwritten digits. The implementation involves 3 steps: Pre-processing, in which we select features from the training, validation, and test set; Feedforward and back propagation, which passes the assigned weights and the training data through an activation function and calculates the error in back propagation; and finally putting this calculated error again in the neural network via back propagation process, to learn the correct weights. Our implementation shows the behavior of neural network on changing the hyper-parameters. We have also analyzed a face dataset and compared the performance of the neural network against deep neural network using the TensorFlow library.

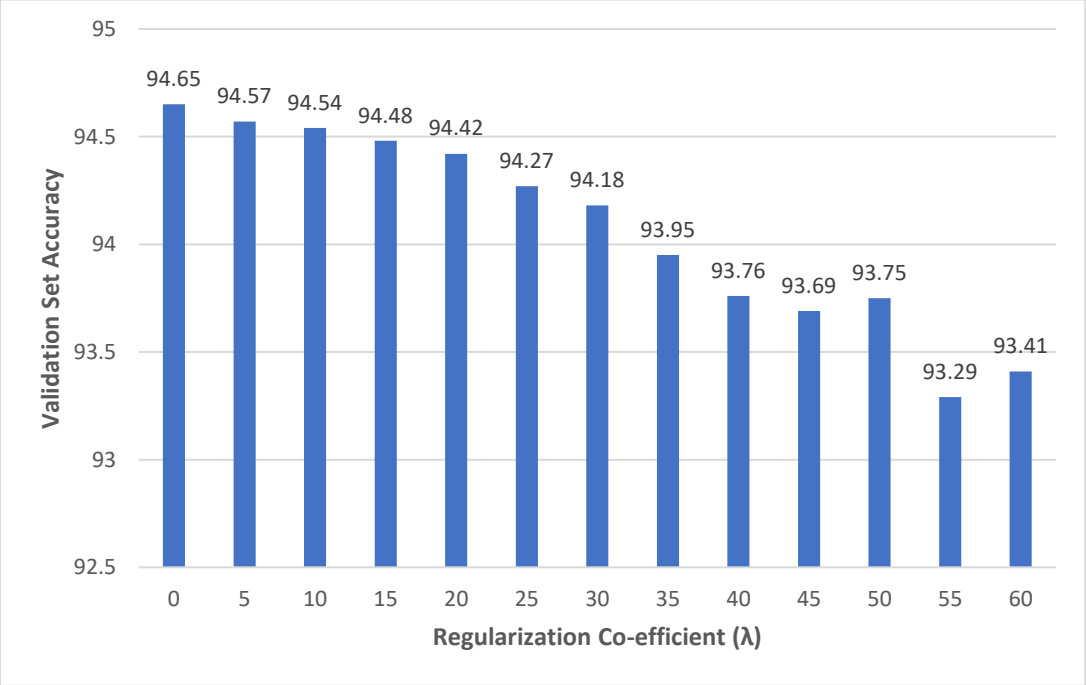
1. Choosing hyper-parameters for neural network:

1. Regularization Coefficient λ:

To choose the hyper–parameter λ we evaluated our neural network by varying the value of regularization coefficient λ against validation set accuracy.

Regularization Co-efficient v/s Validation Set Accuracy

No. of hidden units = 50 and iterations = 50

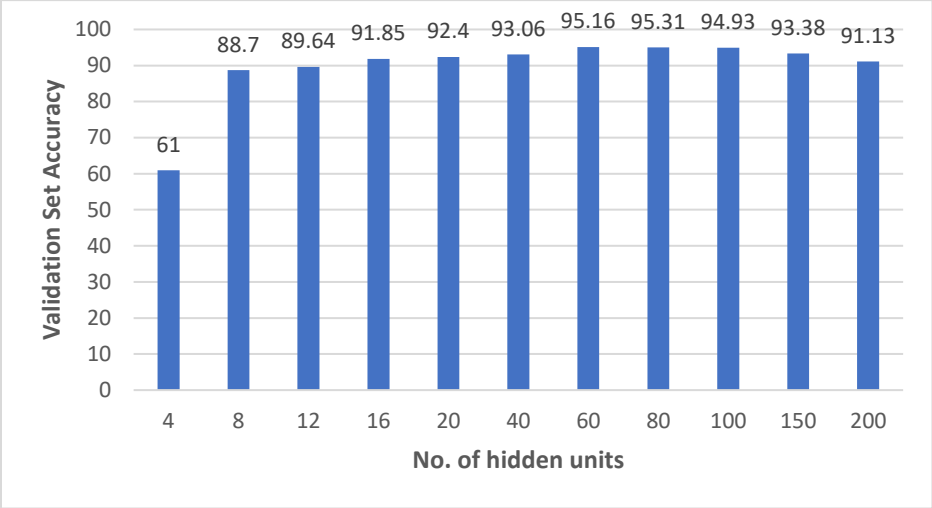


After evaluation, we observed that the validation set accuracy was best obtained at λ values 0 - 20. As the regularization, co-efficient increased, there was a decrease in the accuracy percentage. This is mainly because as we increase the regularization coefficient λ, over-fitting decreases and a more generalized model is formed. But in our case, since the testing set is similar to training set the generalization harms the accuracy.

2. Number of Hidden Units:

Number of Hidden Units v/s Validation Set Accuracy

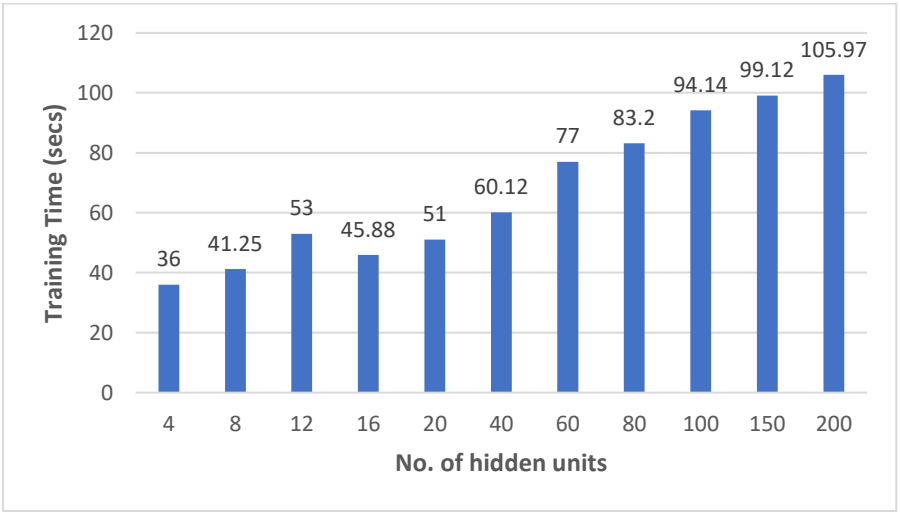
$\lambda = 0$ and iterations = 50



The above graph shows result of validation set accuracy on changing the number of hidden units. The best accuracy of 95.31% was obtained when the number of hidden units was 80. We observed that the accuracy initially increased but after a certain number of hidden units (i.e. 80 in our implementation) there was a very minute or not much change in the accuracy result.

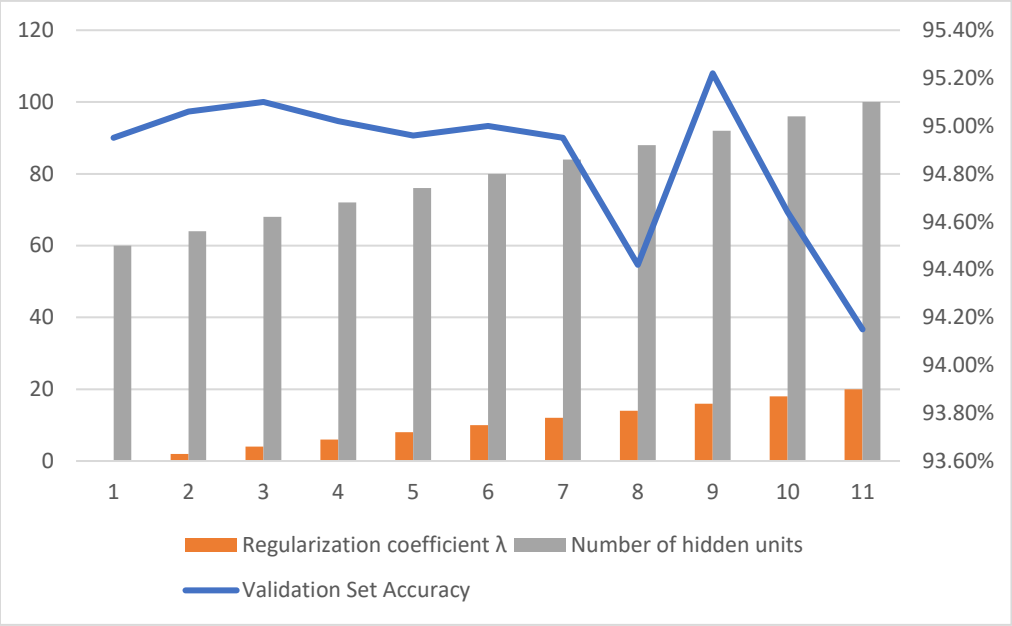
Number of hidden Units v/s Training Time

$\lambda = 0$ and iterations = 50



The above graph shows variation of training time with an increase in the number of hidden units. Our implementation shows that with an increasing number of hidden units, the time taken to train the data also increases. So, based on the above results, best accuracy and training time was observed to be in the range of hidden units 60 – 100.

Now based on the above 2 optimal range of λ (0 - 20) and hidden units (60 - 100), we compare both against the validation set accuracy to determine their best possible values. As the number of hidden units increase, the overfitting problem arises hence we need to add regularization coefficient to avoid the problem. So, we increase the hidden units and λ gradually to determine the best validation set accuracy.



After evaluation of the results, we observed a peak of **95.22% accuracy** for the value of hyper-parameters:

Regularization coefficient λ = 16

Number of hidden units = 92

Accuracy results for the peak value were as follows:

Training set Accuracy: 95.894%

Validation set Accuracy: 95.22%

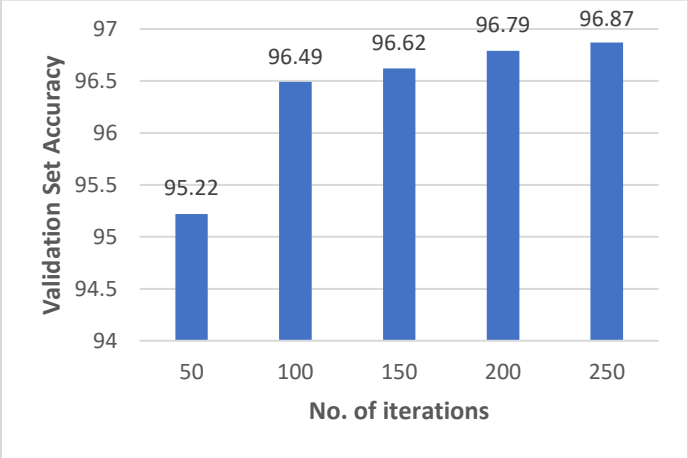
Test set accuracy: 95.47%

2. Accuracy of classification method on the handwritten digits test data:

The accuracy of our classification method on the handwritten digits test data with hyper-parameters λ = 16 and number of hidden units = 92 is 95.47%

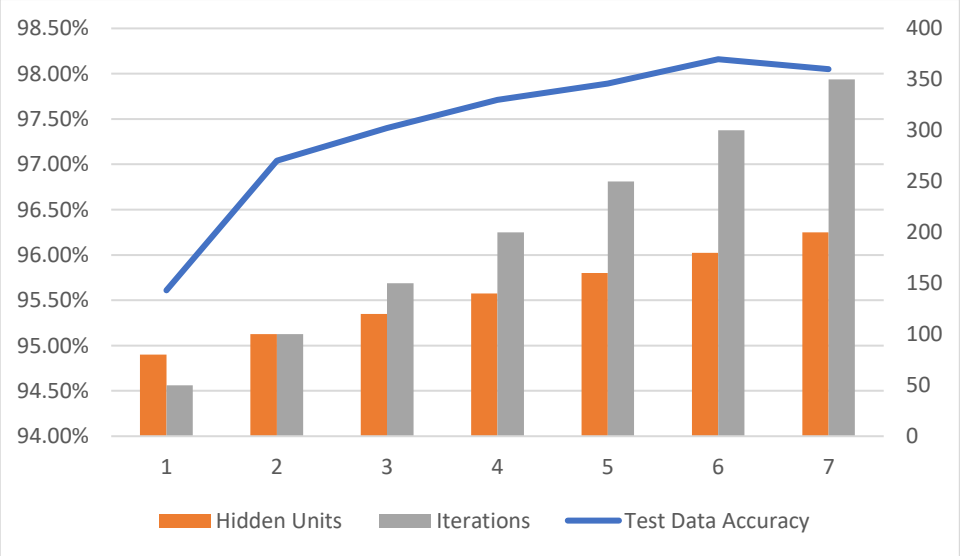
Other Observations:

Number of Iterations v/s Validation Set Accuracy



The above graph represents how the change in number of iterations affects the validation set accuracy. We observed that, with an increase in the number of iterations, there was also a prominent increase in the validation set accuracy.

The below graph in a gist shows our analysis of the behavior of the neural network on increasing the number of units along with the number of iterations.



We see that the test data accuracy increases with an increase in the number of hidden units and the number of iterations. So, it gives the best accuracy of 98.16% for test-set 6 in the graph, which has 180 hidden units and 300 number of iterations.

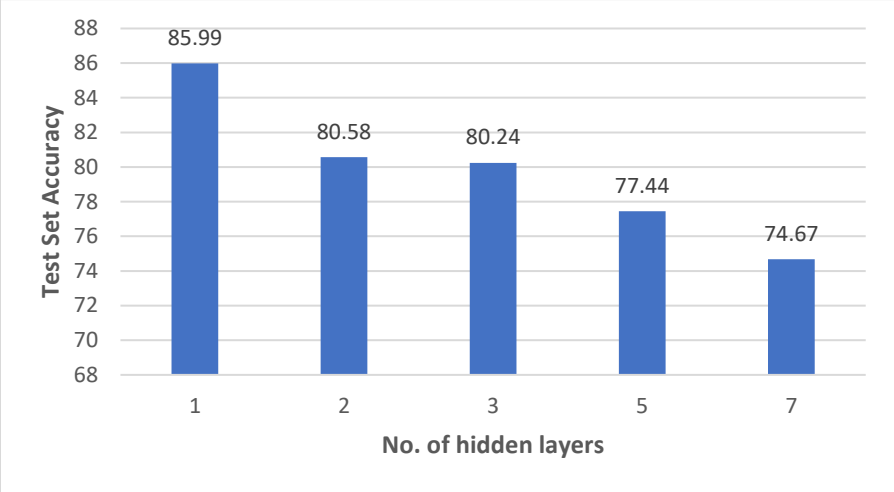
3. Accuracy of classification method on the CelebA data set:

The test set accuracy of our classification method on the CelebA data set is 85.99%.

4. Comparison of neural network with deep neural network:

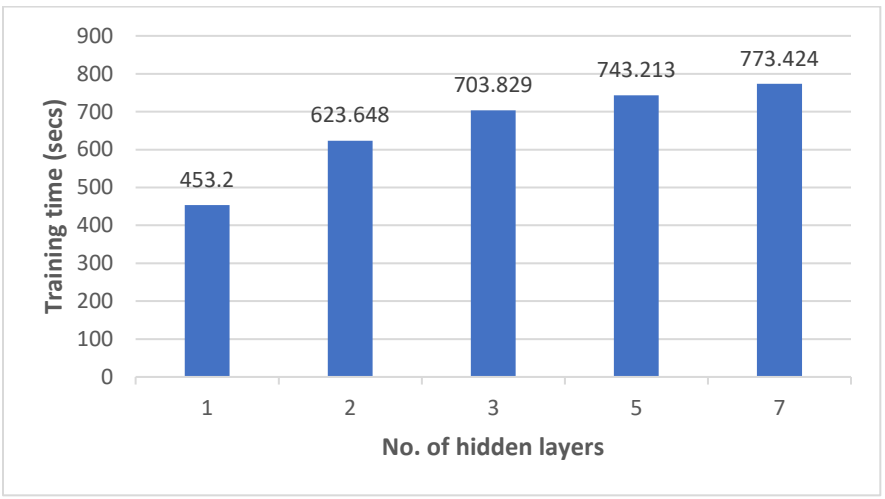
We used the CelebA dataset and implemented neural network with one hidden layer and deep neural networks with 2, 3, 5 and 7 hidden layers.

1. Number of Hidden Layers v/s Test Set Accuracy:



The graph shows how the test set accuracy varies with increasing number of hidden layers. After observation, we concluded that the test set accuracy decreases as the number of hidden layers increase in this case. This is because as we increase the number of hidden layers, overfitting increases which increases the accuracy of training set but decreases the accuracy on test set.

2. Number of hidden layers v/s Training time:



The graph shows how the training time varies with increasing number of hidden layers. We observed that as the number of hidden layers increase, there was a significant increase in the training time.