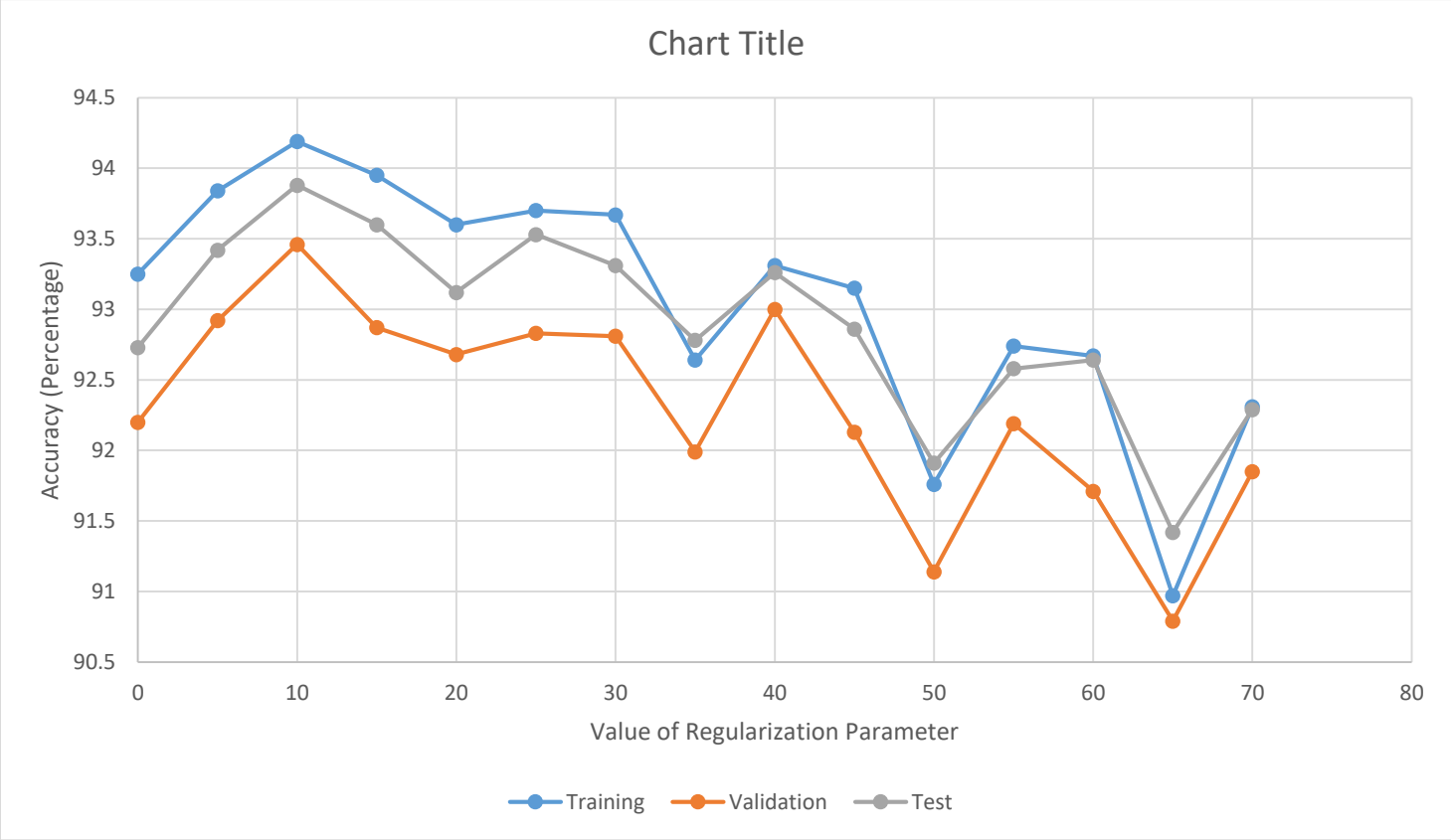
CSE 574
Group 55

Team Members: **Rohit Naresh Tarachandani**
Karan Gopal Bhatia
Jay Narendra Shah

Digits Dataset

Value of Regularization parameter (lambdaval) v/s Accuracy for 20 Hidden Nodes



Observations:

When regularization parameter is set to 0, we can clearly see that overfitting occurs as accuracy is high for training data but low for validation data and test data.

Now as we increase the regularization parameter, we observe that the accuracy gap between the training data and validation data starts to vary. The accuracy does improve and remains somewhere at the top till value of regularization is 20. However, as we increase lambda further, the accuracy starts decreasing for all data sets. This happens due to under-fitting the data set too much.

Strategy for selection of optimal value of regularization parameter:

We select that value as optimal for which the accuracy is considerably high for validation data and test data. Also, the accuracy gap between the training data and validation data should be low. This ensures that our algorithm performs uniformly across a wide variety of datasets.

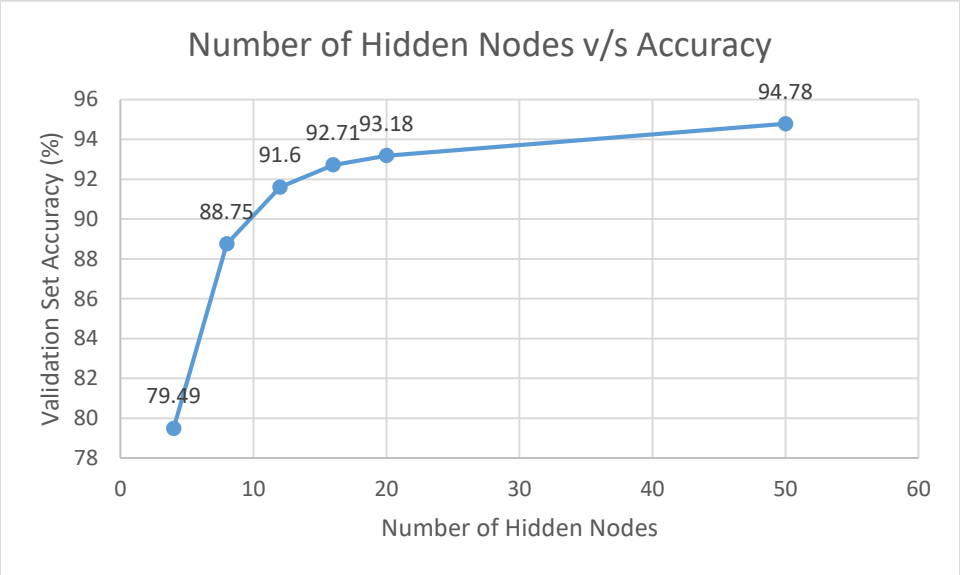
Results:

Optimal regularization parameter value: 10

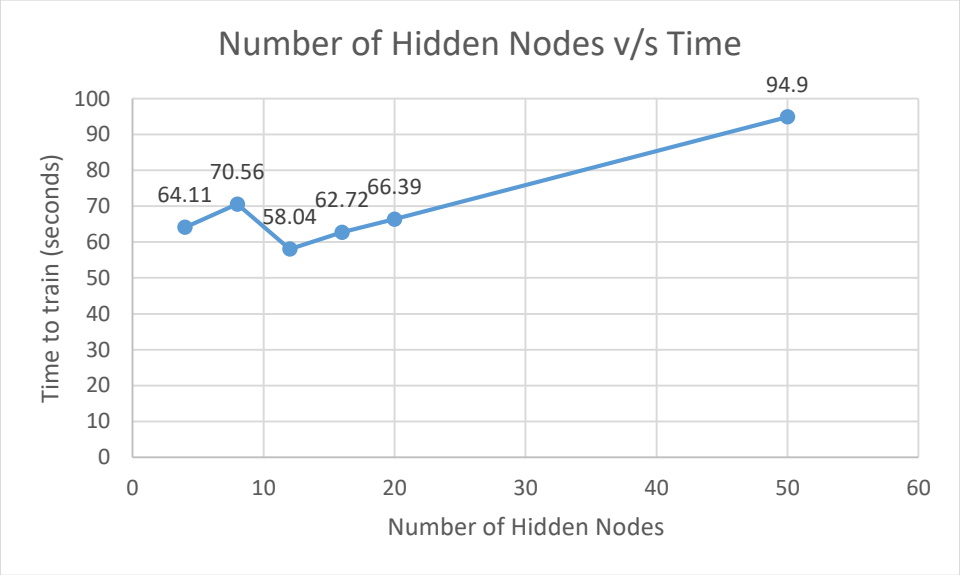
Accuracy:

1. Training dataset - 94.19%
2. Validation dataset – 93.46%
3. Test data set – 93.88%

Number of Hidden Nodes (Value of Regularization: 10) v/s Validation Set Accuracy



Number of Hidden Nodes (Value of Regularization: 10) v/s Time to Train (in seconds)



Observations:

We observe that as we increase the number of nodes, the validation set accuracy increases as well as the time required to train the data. This is quite expected behavior since higher number of nodes require more computations and hence the increase in time.

Strategy for selection of optimal value of Hidden Nodes:

We select that value for hidden nodes for which the accuracy is very good as well as time taken for training is relatively not very high. We observe that accuracy difference between 50 hidden nodes and 20 hidden nodes is not very high but the training time gap is relatively high. Hence, it is better to go for 20 hidden nodes as it offers very good accuracy without taking a hit on the training time.

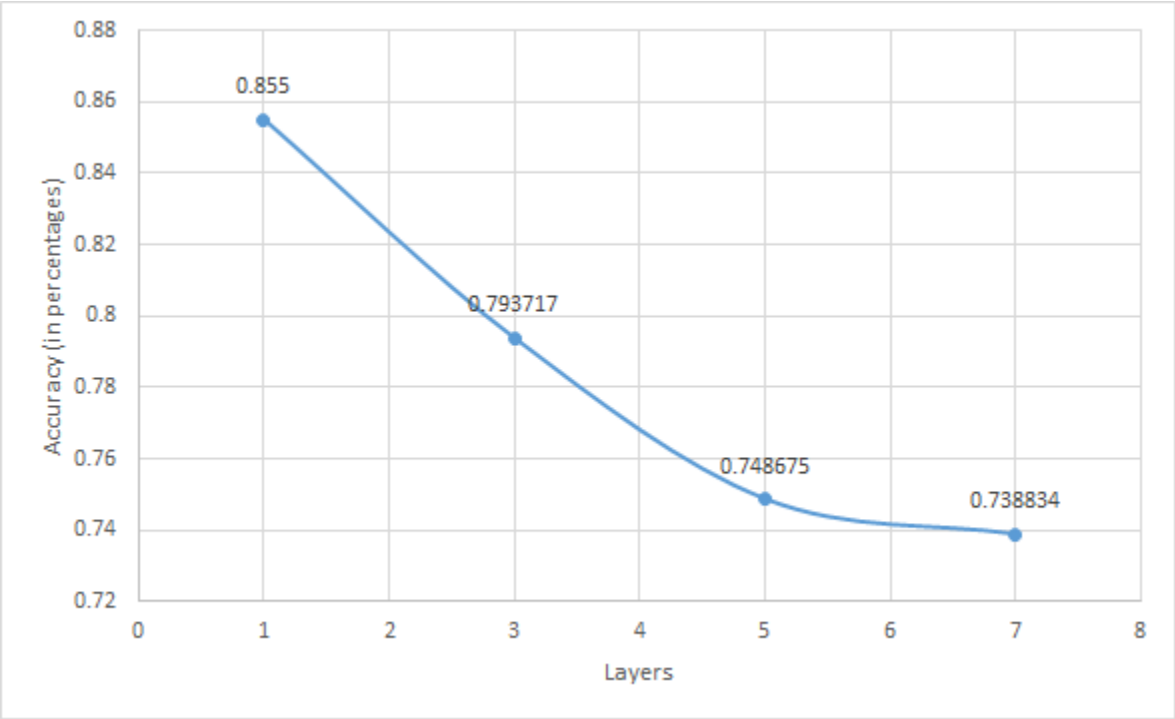
Time taken to train data with 20 Nodes and value of regularization as 10 is 66.39 seconds

Results:

Optimal Number of Hidden Nodes: 20

CelebA Dataset Results for Deep Neural Network and Normal Neural Network:

Number of Neural Network Hidden Layers v/s Accuracy

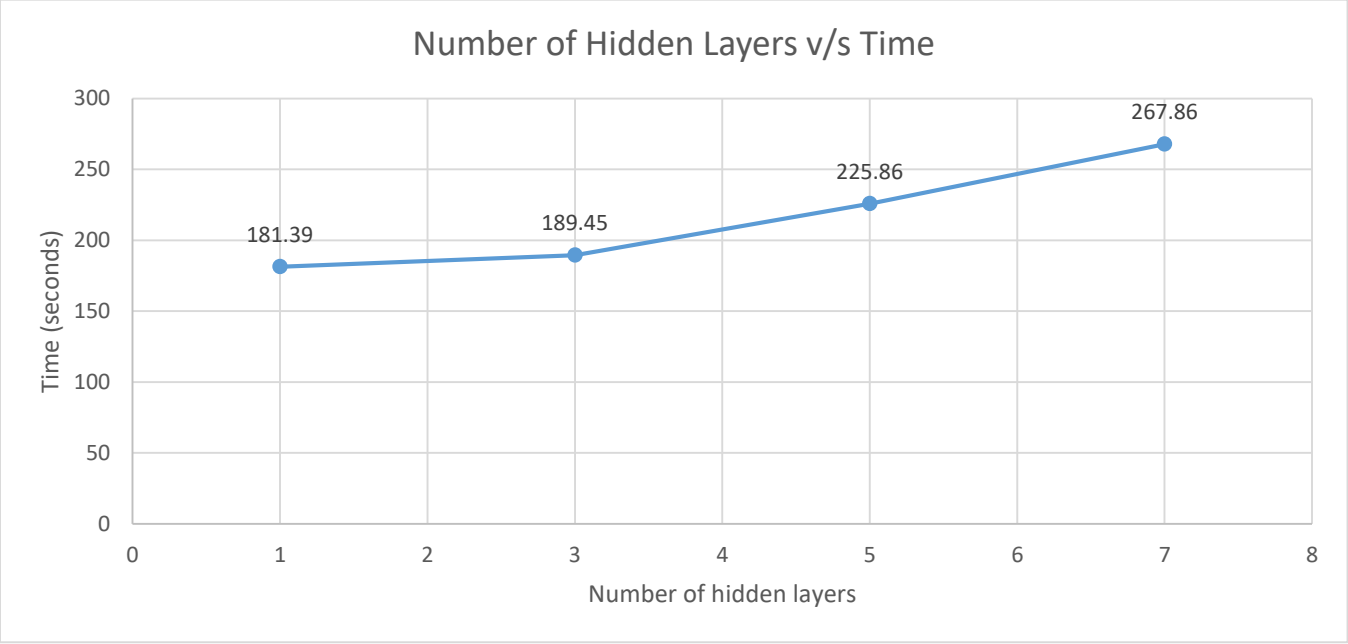


We observe that as the number of hidden layers increase, accuracy decreases. This is because the architecture becomes more and more complex and ends up not producing good results like the one produced by a single hidden layer.

Results:

Accuracy on the CelebA data sets: 1 hidden layer: 85.5%, 3 hidden layers: 79.37%, 5 hidden layers: 74.87%
7 hidden layers: 73.88%

Neural Network Hidden Layers v/s Time



We observe that the time taken to train the data increases with the increase in number of layers. This is expected behavior as number of computations and complexity increases with increase in the number of hidden layers, hence the increase in time.