

CSE 574
INTRODUCTION TO MACHINE LEARNING
PROJECT ASSIGNMENT – 1
HANDWRITTEN DIGIT CLASSIFICATION

Report done by
CSE 574 Group 14
Anuradha Ashavatha Rao(aashavat)
Nagadeesh Nagaraja(nagadees)
Nagesh Rao(nageshra)

Selection of hyperparameters:

We initially had 50 hidden layers while building the neural network and had set the lambda value to 0. Once the neural network was complete we observed that the accuracy was 94%.

We iterated through various combinations of lambda and number of hidden layers and got the results as tabulated below. Our observations are as follows:

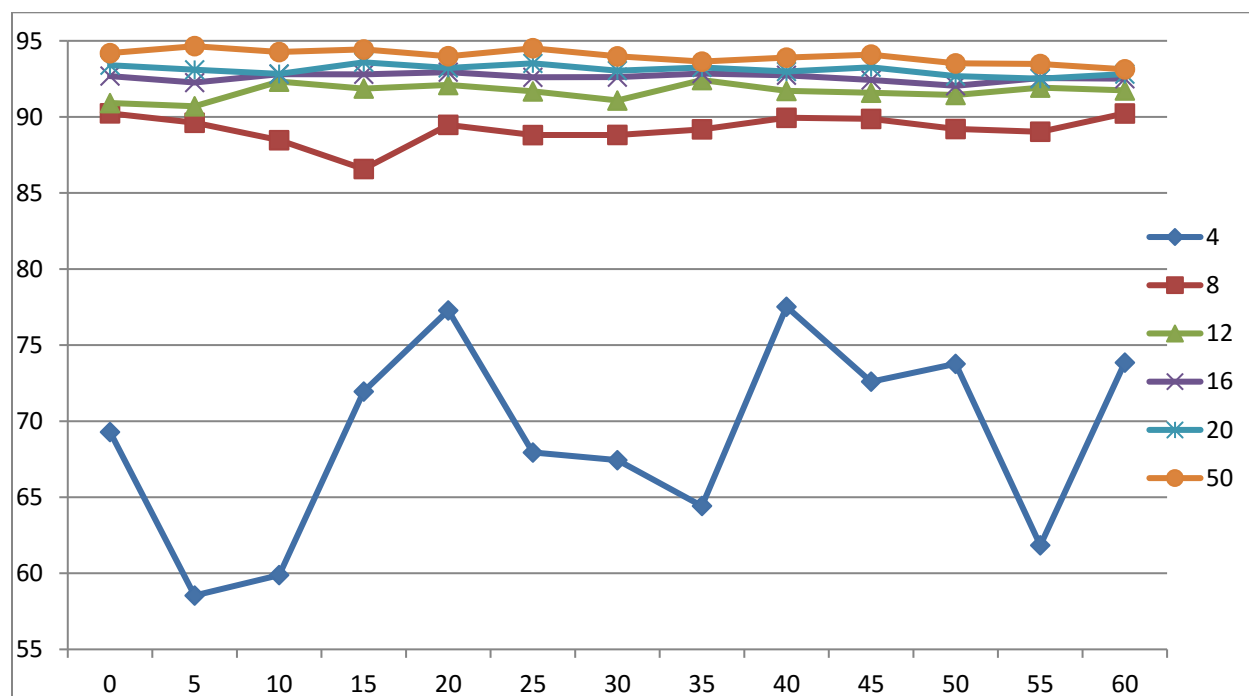
Training data:

- If the hidden layers are less(4, in this case) and we vary only the lambda value, we get a poor accuracy, but on increasing the lambda value, we observe a pattern where the value falls drastically followed by a gradual increase. The accuracy did not exceed 75%
- If the hidden layers are more(50, in this case) and we vary only the lambda value, we observe that the accuracy follows the same pattern, but the amount of variance is subdued as compared to the previous case. This shows that the more hidden layers there are, the more accurate the predictions get. From the table below, we observe that for lambda = 5 and number of hidden layers = 50, we get the maximum accuracy i.e. 94.64%
- Another observation made was that as the number of hidden layers increase, the variance of accuracy of the neural networks decreases, for a given lambda value. Ex: For lambda value = 5, we find the accuracy of the neural network with 4 hidden layers at 58. On increasing the number of hidden layers to 8 the change in accuracy is about 31.08%. But, on increasing the number of hidden layers from 8 to 12, the change in accuracy falls to just about 1.08%, and remains about the same till we get the best result of 94.64%

Table Training data:

	no.of Hidden layer					
lamda	4	8	12	16	20	50
0	69.29	90.232	90.912	92.684	93.398	94.192
5	58.538	89.618	90.7	92.256	93.098	94.648
10	59.88	88.47	92.32	92.812	92.826	94.268
15	71.948	86.578	91.862	92.802	93.588	94.434
20	77.282	89.474	92.098	92.94	93.218	93.988
25	67.944	88.81	91.68	92.606	93.52	94.514
30	67.438	88.814	91.08	92.618	93.042	93.984
35	64.422	89.184	92.424	92.846	93.25	93.634
40	77.52	89.942	91.712	92.738	92.986	93.9
45	72.594	89.87	91.59	92.424	93.264	94.092
50	73.766	89.21	91.444	92.044	92.686	93.522
55	61.838	89.026	91.924	92.552	92.512	93.48
60	73.85	90.23	91.744	92.51	92.814	93.128

Chart Training data:



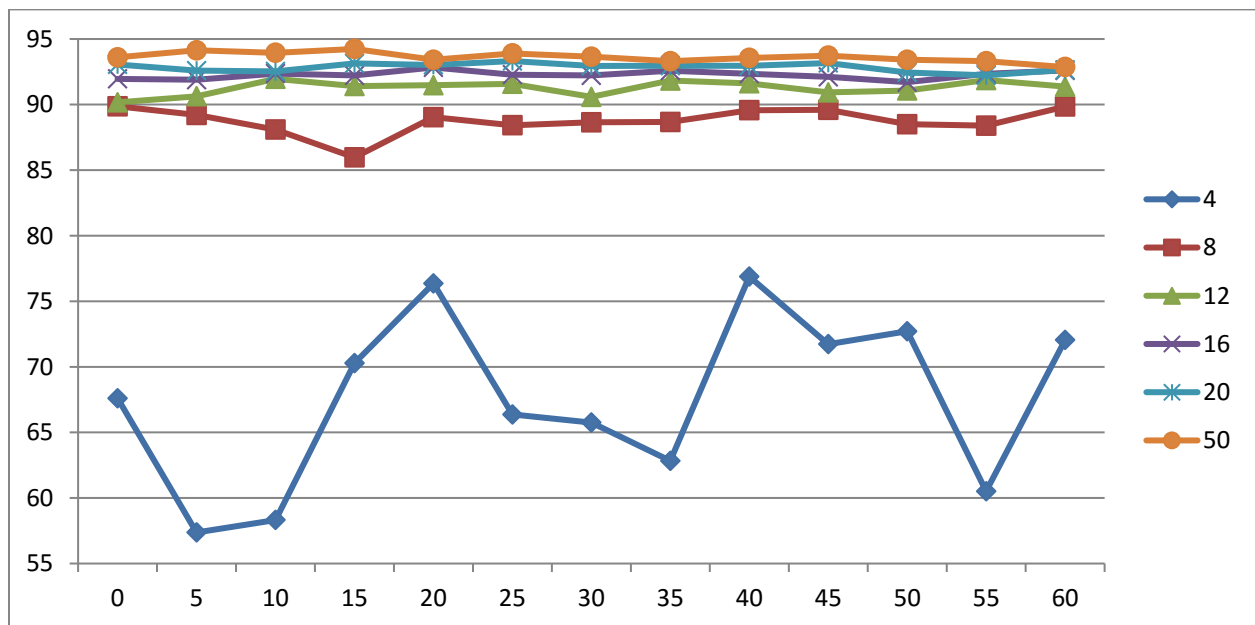
Validation data:

- With validation data, we found that the accuracy fell down by a small margin to 94.14%.
- Most of the trends observed were similar to that observed on the training data.
- We observed that the maximum accuracy was observed when the lambda value was set to 5 and number of hidden nodes was 50.

Table Validation data:

	no.of Hidden layer					
lamda	4	8	12	16	20	50
0	67.6	89.87	90.17	91.96	93.06	93.6
5	57.37	89.21	90.63	91.89	92.59	94.14
10	58.32	88.09	91.96	92.36	92.53	93.95
15	70.28	85.97	91.41	92.22	93.14	94.24
20	76.36	89.04	91.48	92.83	93.02	93.42
25	66.37	88.42	91.58	92.28	93.32	93.9
30	65.75	88.65	90.59	92.22	92.95	93.65
35	62.83	88.67	91.83	92.58	92.96	93.31
40	76.88	89.57	91.62	92.35	92.96	93.56
45	71.73	89.6	90.93	92.11	93.17	93.73
50	72.71	88.5	91.07	91.71	92.45	93.42
55	60.51	88.39	91.87	92.35	92.23	93.31
60	72.05	89.85	91.37	92.6	92.65	92.86

Chart validation data:



Testing data:

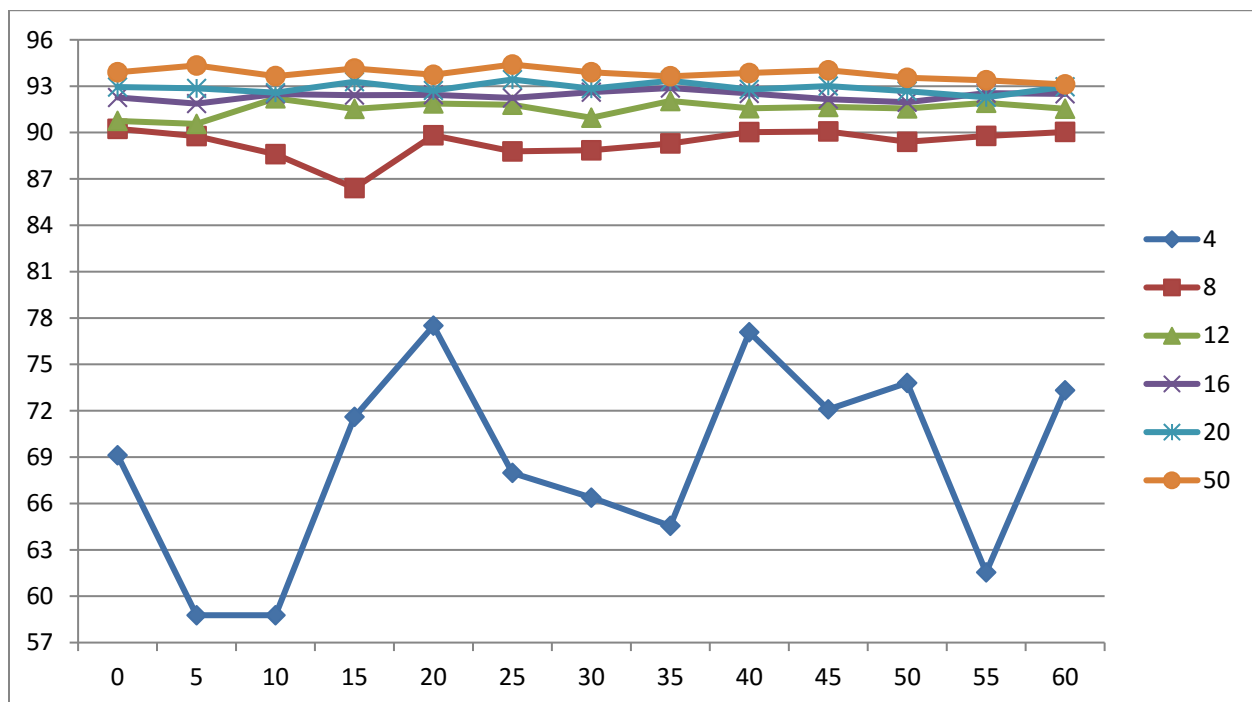
- This testing dataset was used to test the accuracy of the neural network on new data that it has not seen before and analyze its accuracy.
- We observed that all trends observed in the training dataset occurred in the testing dataset too.

- The accuracy, however, was maximum when the lambda value was 25. This could be attributed to the fact that the images sent in as input was random and would vary for every new dataset.
- The maximum accuracy obtained was 94.4%(lambda = 25, number of hidden nodes = 50)

Table testing data:

	no.of Hidden layer					
lamda	4	8	12	16	20	50
0	69.12	90.24	90.75	92.27	92.94	93.9
5	58.77	89.77	90.56	91.87	92.86	94.34
10	58.77	88.6	92.21	92.49	92.58	93.65
15	71.6	86.41	91.53	92.42	93.29	94.14
20	77.5	89.82	91.88	92.44	92.72	93.74
25	67.98	88.78	91.8	92.24	93.43	94.4
30	66.38	88.85	90.96	92.61	92.84	93.9
35	64.55	89.29	92.04	92.88	93.35	93.64
40	77.08	90.03	91.57	92.53	92.79	93.85
45	72.09	90.07	91.66	92.16	93.01	94.03
50	73.8	89.4	91.56	91.96	92.68	93.54
55	61.54	89.78	91.92	92.55	92.28	93.38
60	73.32	90.04	91.54	92.5	92.96	93.11

Chart testing data:



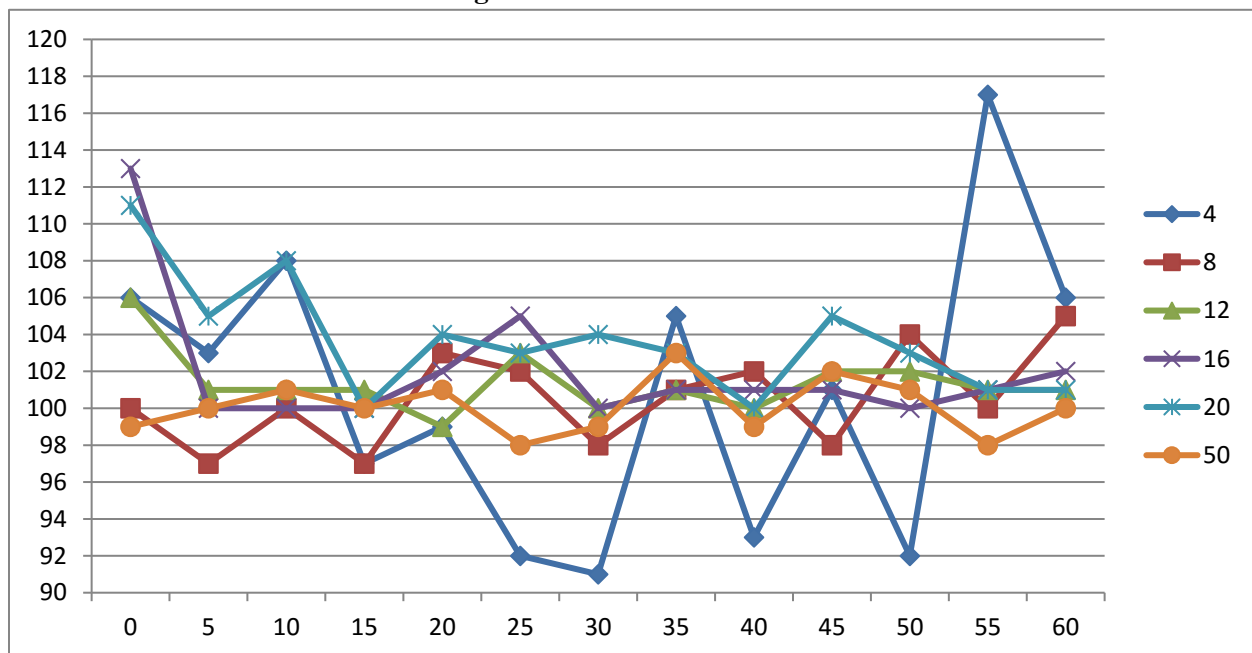
Convergence data:

- This data was used to give us an idea on how many iterations were required to converge.
- We observed that the lesser the number of hidden layers, the fluctuation in convergence value will be high.
- For our testing set, where $\lambda=25$ and number of hidden layers=50, we observe the number of iterations required is 98

Table number of iteration for convergence:

	no.of Hidden layer					
lamda	4	8	12	16	20	50
0	106	100	106	113	111	99
5	103	97	101	100	105	100
10	108	100	101	100	108	101
15	97	97	101	100	100	100
20	99	103	99	102	104	101
25	92	102	103	105	103	98
30	91	98	100	100	104	99
35	105	101	101	101	103	103
40	93	102	100	101	100	99
45	101	98	102	101	105	102
50	92	104	102	100	103	101
55	117	100	101	101	101	98
60	106	105	101	102	101	100

Chart number of iteration for convergence:



Running Neural Network on CelebA Data and Comparison with Deep Neural Network using TensorFlow Library:

We apply our neural network on the CelebA dataset. We used regularization parameters ($\lambda = 25$ and number of hidden nodes = 50) obtained from our neural network. On running the neural network, we obtain an accuracy of 84% on the test data.

On running the deep neural network with 2 hidden layers, we get an accuracy of 80% on the test dataset. We increased the number of hidden layers to observe the accuracy. We notice that the values obtained are 77.94%, 76.64% and 73.01% for 3, 5 and 7 layers respectively. This goes on to prove that it is not necessarily true that an increase in hidden layers will increase the accuracy. Even though deep neural network has capabilities to learn complex concepts, and might be capable to give accurate results for training set, when the test dataset is passed through the neural network, this will cause overfitting of the data and thus, causes the accuracy to fall.

Number of Hidden Layers	Accuracy
1	84%
2	80.20%
3	77.94%
5	76.64%
7	73.01%