

CSE 574 Introduction to Machine Learning
Programming Assignment 2
Handwritten Digits Classification

GROUP 71

SIVAKUMAR PASUPATHI 50366350
SAI PRASHANTH SELVAGANAPATHY 50419614
PARTHIBAN RAJENDRAN 50415960

Feature Selection

There are totally 717 feature that we can use to train our model. The remaining features are common among all the data points in the training sets and are excluded from the final list of features.

hyper-parameter for Neural Network

We varied the lambda values and the number of hidden unit node with the following values and compared the accuracy and choose the lambda which had the highest accuracy

Lambda - 0,5,10,15,20,25,30, 35,40,45,50, 55,60

Hidden nodes: - 4,8,12,16,20

We varied the lambda and the number of hidden nodes for all possible combinations and selected the one which had the highest test accuracy.

Output with various combination of lambda and hidden node values:

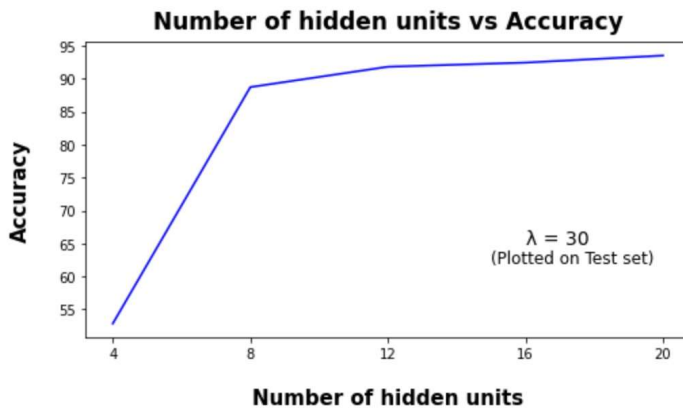
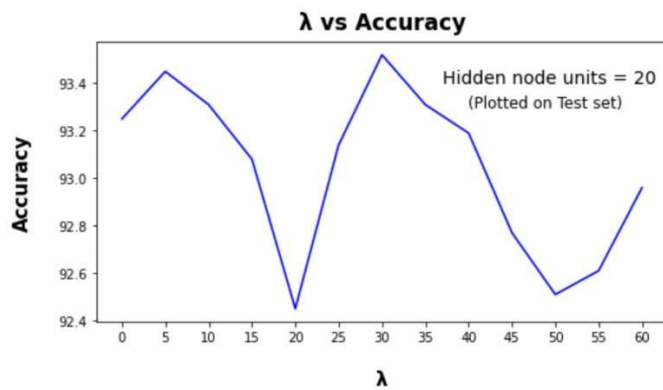
λ	M	Training set accuracy	Validation set accuracy	Test set accuracy	Training Time in secs
0	4	57.706	56.94	56.67	24.2
	8	84.894	84.13	85.43	27.67
	12	92.074	91.26	91.84	30.04
	16	93.58	92.9	93.35	31.6
	20	93.69	92.97	93.25	33.37
5	4	65.57	64.92	65.38	25.54
	8	87.55	87.03	86.58	25.33
	12	91.546	91.02	91.36	26.80
	16	92.79	91.9	92.35	32.49
	20	93.67	93.24	93.45	34.47
	4	59.80	59.01	59.76	25.65
	8	89.70	89.53	90.01	26.11

10	12	92.39	91.79	92.51	26.06
	16	92.53	92.2	92.77	31.76
	20	93.60	93.02	93.31	33.58
15	4	65.03	64.34	64.64	23.15
	8	87.83	87.44	87.46	25.55
	12	89.84	89.09	89.67	27.20
	16	92.96	92.29	92.84	31.19
	20	93.50	92.72	93.08	32.75
20	4	62.87	61.96	62.89	19.68
	8	88.88	88.42	89.01	24.10
	12	92.45	91.69	92.07	26.56
	16	93.29	92.35	92.97	29.76
	20	92.944	92.14	92.45	32.65
25	4	76.06	76.27	75.33	22.44
	8	89.55	89.07	89.56	25.57
	12	92.15	91.64	91.86	27.29
	16	92.68	92.15	92.94	31.08
	20	93.47	92.78	93.14	34.15
30	4	53.46	53.22	52.83	24.31
	8	88.92	88.33	88.73	25.02
	12	91.95	91.22	91.82	26.84
	16	92.81	92.30	92.45	31.58

	20	93.744	93.08	93.52	32.15
35	4	63.36	63.27	63.72	23.96
	8	87.07	97.31	86.98	24.66
	12	91.37	90.83	90.93	27.37
	16	93.07	92.33	92.77	30.61
	20	93.422	92.94	93.31	32.31
40	4	59.34	58.48	58.52	21.63
	8	88.58	88.12	89.01	24.36
	12	91.76	91.36	91.44	26.81
	16	91.87	91.35	92.15	31.25
	20	93.12	92.7	93.19	32.33
45	4	69.04	68.85	69.88	19.86
	8	85.60	85.11	85.55	24.17
	12	91.82	91.52	92.0	26.29
	16	92.75	92.07	92.55	30.94
	20	92.73	92.13	92.77	32.21
50	4	62.30	61.29	62.09	20.49
	8	89.11	88.57	89.28	24.02
	12	91.74	91.07	92.0	25.73
	16	92.02	91.34	91.74	30.68
	20	92.48	91.85	92.51	31.61
	4	72.70	72.65	73.49	19.45

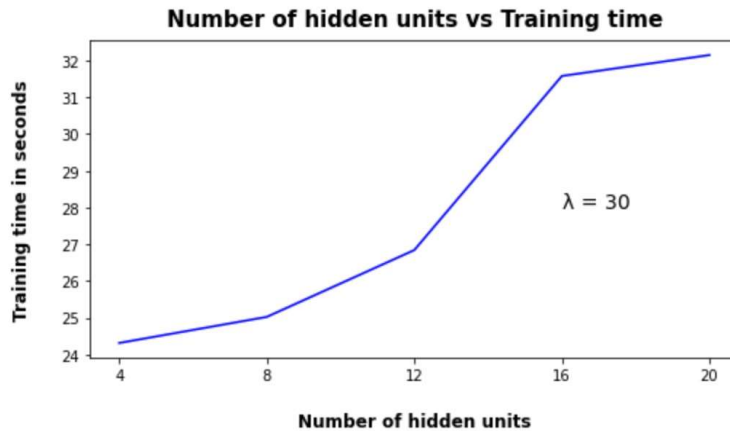
55	8	89.88	89.48	89.9	24.43
	12	91.42	91.4	91.38	25.62
	16	92.52	92.19	92.61	30.77
	20	92.96	92.69	92.61	30.69
60	4	65.01	64.28	65.14	19.89
	8	86.56	86.18	86.68	22.30
	12	88.58	87.82	88.7	25.51
	16	92.07	91.68	92.03	30.53
	20	92.77	92.30	92.96	32.08

From the above table we can see that the best test accuracy was achieved with for the combination of $\lambda= 30$ and **m (hidden node) = 20**, which is the optimal value of λ and m for the mnist dataset.



Accuracy of Handwritten Digit Test:

We Recorded the maximum Accuracy of 93.52% in the test set for lambda value of 30 and hidden node count of 20. The accuracy generally increases with the increase in the number of hidden units, but this comes at the cost of increase in training time (example: from 24 sec to 32 secs for the optimal lambda value)



```
Activities Terminal Nov 14 12:46 PM
salprashanth@salprashanth: ~/Desktop/Assignment2/Assignment2/basecode

Validation set Accuracy:92.72%

Test set Accuracy:93.14%
Total Time: 34.15244197845459

LamdaVal: 30 , hidden_val: 4

Training set Accuracy:53.464%
Validation set Accuracy:53.22%

Test set Accuracy:52.839999999999996%
Total Time: 24.318821668624876

LamdaVal: 30 , hidden_val: 8

Training set Accuracy:88.928%
Validation set Accuracy:88.33%

Test set Accuracy:88.73%
Total Time: 25.022607803344727

LamdaVal: 30 , hidden_val: 12

Training set Accuracy:91.95%
Validation set Accuracy:91.22%

Test set Accuracy:91.82000000000001%
Total Time: 26.044977140426536

LamdaVal: 30 , hidden_val: 16

Training set Accuracy:92.81%
Validation set Accuracy:92.38000000000001%

Test set Accuracy:92.45%
Total Time: 31.58514952659607

LamdaVal: 30 , hidden_val: 20

Training set Accuracy:93.744%
Validation set Accuracy:93.08%

Test set Accuracy:93.51%
Total Time: 32.157980640930176

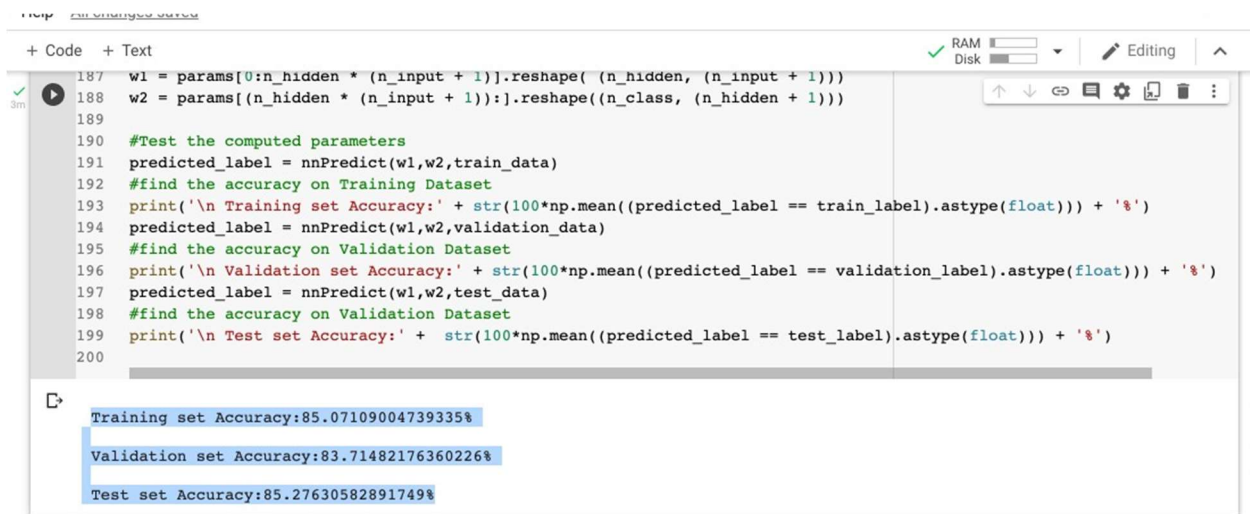
LamdaVal: 35 , hidden_val: 4

Training set Accuracy:63.336000000000006%
Validation set Accuracy:63.27%
```

Accuracy of classification method on the CelebA data set:

We achieved the following accuracy when running the faceNN script

The accuracy of the model is **85.27%**



```
187 w1 = params[0:n_hidden * (n_input + 1)].reshape((n_hidden, (n_input + 1)))
188 w2 = params[(n_hidden * (n_input + 1)):(n_class * (n_hidden + 1))].reshape((n_class, (n_hidden + 1)))
189
190 #Test the computed parameters
191 predicted_label = nnPredict(w1,w2,train_data)
192 #find the accuracy on Training Dataset
193 print('\n Training set Accuracy:' + str(100*np.mean((predicted_label == train_label).astype(float))) + '%')
194 predicted_label = nnPredict(w1,w2,validation_data)
195 #find the accuracy on Validation Dataset
196 print('\n Validation set Accuracy:' + str(100*np.mean((predicted_label == validation_label).astype(float))) + '%')
197 predicted_label = nnPredict(w1,w2,test_data)
198 #find the accuracy on Validation Dataset
199 print('\n Test set Accuracy:' + str(100*np.mean((predicted_label == test_label).astype(float))) + '%')
200
```

Training set Accuracy:85.07109004739335%

Validation set Accuracy:83.71482176360226%

Test set Accuracy:85.27630582891749%

Comparison of your neural network with a deep neural network (using TensorFlow):

We ran the single and deep neural network on the celeb face data set and got the following results:

- As expected, increasing the number of layers increases the training time
- The accuracy decreases with increase in layers also due to overfitting of data

Script	Layers	Accuracy	Time in seconds
FaceNN	1	85.27	141
DeepNN	3	82.059	185
	5	80.961	192
	7	76.21	197

Convolutional Neural Network:

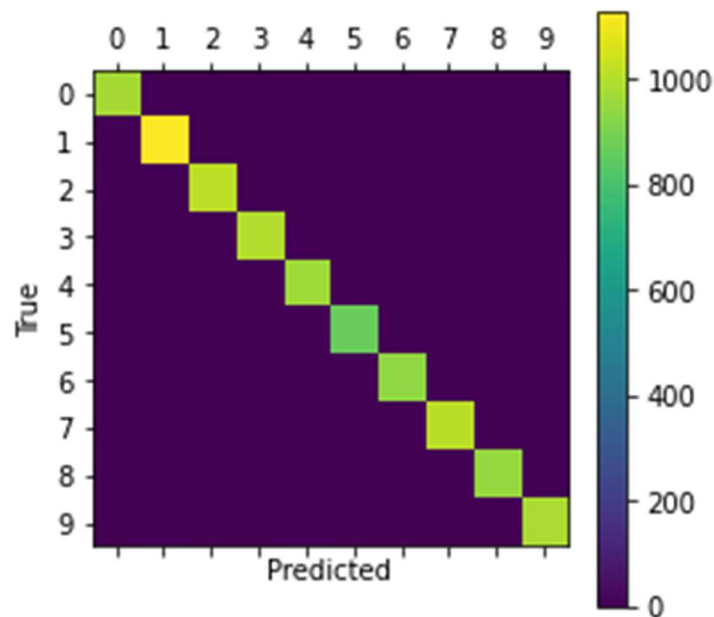
On running the cnsyscript.py for the given data set we get the confusion matrix as given below

Time usage: 0:02:15

Accuracy on Test-Set: 98.6% (9857 / 10000)

Confusion Matrix:

```
[[ 976   1   0   0   0   0   1   1   1   0]
 [   0 1130   2   0   0   0   1   1   1   0]
 [   6   1 1018   0   1   0   0   4   2   0]
 [   1   1   0 998   0   0   0   4   3   3]
 [   1   0   1   0 969   0   1   4   1   5]
 [   2   0   1   7   0 872   4   1   2   3]
 [   6   3   0   0   1   3 944   0   1   0]
 [   1   1   8   2   1   0   0 1013   1   1]
 [   5   1   5   2   2   1   1   4 950   3]
 [   4   5   0   1   5   1   0   4   2 987]]
```



We observe that the accuracy of convoluted neural network is higher than other neural networks and the accuracy is **98.6%**. The Training time for CNN is much higher at 135 secs. Convolutional Neural Network has the ability to view image as data which is better for classification.