Neural Networks HW-2B Report

line 1: 1st Sai Krishna Wupadrashta
line 2: M.S., EE -*NC State Univerisity*

***Abstract*—This is the HW -2B report for ECE 542(Neural Networks), which implements the back propagation algorithm for the Hand written digit recognition using the MNIST dataset provided.**

I. Final Neural Network Structure

The final neural network structure or the best structure after tweaking the hyperparameters consists of:

i) One input layer- taking in 784 neurons for the 28x28 pixel for the handwritten digits datasets.

ii) One hidden layer- consisting of 30 neurons.

iii) One Output layer- that consists of 10 outputs Yi, which gives the output in the form of one-hot encoding to detect what the handwritten digit is from 0 to 9.

II. Choice of Hyperparameters

The hyperparameters considered are the epochs, learning rate, number of hidden layers, number of neurons in each hidden layer, testing accuracy, training accuracy, validation accuracy.

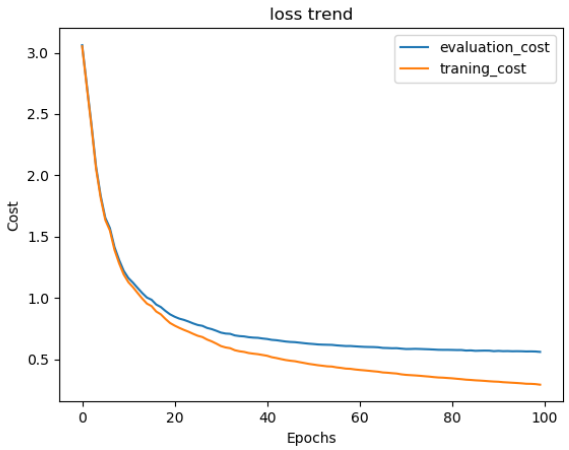
The following picture below shows the variations in the hyperparameters considered, with the best-case choice of hyperparameters highlighted in yellow:

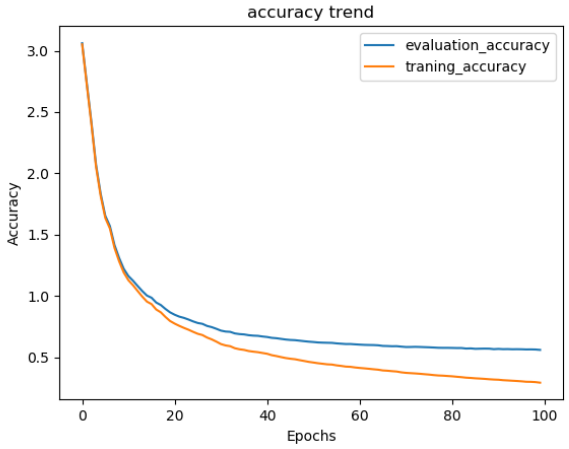
epochs	learning rate	no. of hidden layers	no. of neurons in each hidden layer	testing accuracy	training accuracy	validation accuracy
batch size=128						
100	0.001	1	20	90.16%	2897/3000	9074/10000
100	0.01	1	20	83.80%	2845/3000	8459/10000
100	0.0001	1	20	83.61%	2563/3000	8486/10000
150	0.0001	1	20	87.27%	2705/3000	8838/10000
50	0.0001	1	20	73.99%	2257/3000	7531/10000
100	0.001	2	20	89.37%	2899/3000	8983/10000
100	0.001	1	30	91.25%	2911/3000	9142/10000
batch size=64						
100	0.001	1	20	90.20%	2947/3000	9047/10000
batch size=258						
100	0.001	1	20	89.14%	2814/3000	8975/10000
batch size=64						
100	0.001	2	30	87.86%	27645/3000	8870/10000

Thus, the best possible choice of hyperparameters are [784,30,10], with a batch size of 128 and a learning rate of 0.001.

III. Learning Curves for Training and Validation Set

The learning curves for the training and validation set for the best possible case is as below:






```
Epoch 95 training complete
[training loss]: 0.30525459463603016
[training accuracy]: 2906 / 3000
[Validation loss]: 0.5664795602585229
[Validation accuracy]: 9137 / 10000
Epoch 96 training complete
[training loss]: 0.3017088244163112
[training accuracy]: 2907 / 3000
[Validation loss]: 0.5650416141999838
[Validation accuracy]: 9140 / 10000
Epoch 97 training complete
[training loss]: 0.3010394496333674
[training accuracy]: 2909 / 3000
[Validation loss]: 0.5654801861612055
[Validation accuracy]: 9139 / 10000
Epoch 98 training complete
[training loss]: 0.29888250878997724
[training accuracy]: 2911 / 3000
[Validation loss]: 0.5645484326814849
[Validation accuracy]: 9142 / 10000
Epoch 99 training complete
[training loss]: 0.2941922964801225
[training accuracy]: 2911 / 3000
[Validation loss]: 0.5614567747261844
[Validation accuracy]: 9142 / 10000
[Testing Accuracy]: 9125 / 10000
(base) PS C:\Users\saikr\Desktop\hw02b\assignment\experiment>
```

Thus, the loss and accuracy values obtained are:

i) Validation loss=0.5614

ii) Training Loss=0.2914

iii) Validation Accuracy=9142/10000

iv) Training Accuracy=2911/3000

And the Final Accuracy on the Test Set is 9125/10000 or 91.25%