Final Report -Multilayered Perceptron (ECE 542-HW02b)

Abhihsek Ranjan Singh
Electrical and Computer Engineer
NC State University
Raleigh, US
arsingh3@ncsu.edu

I. Final Network Structure

After considering numerous combinations of hyperparameters and structures I finally settled on a structure with one hidden layer containing 50 Neurons. The network was trained with the learning rate(eta) of 1e-3 for 30 epochs. In this Network I haven't used any type of normalization and hence the value of lambda is set to zero. The details of the performance of the Network are shown in TABLE I.

The learning curves for the validation cost and training cost is shown in Figure 1.

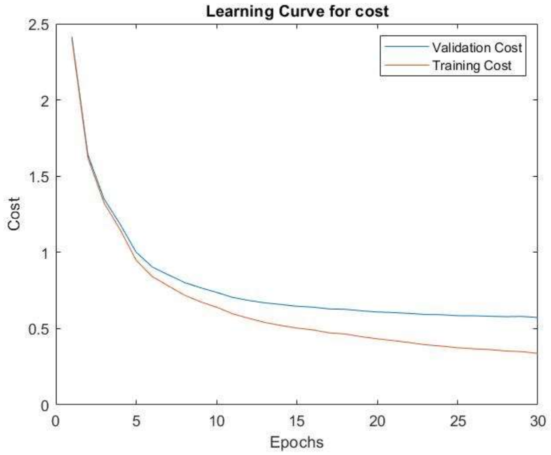

Figure 1: Learning Curve for Validation and Training Cost

Similarly, the learning curve for validation accuracy and training accuracy of the model is shown in Fig 2.

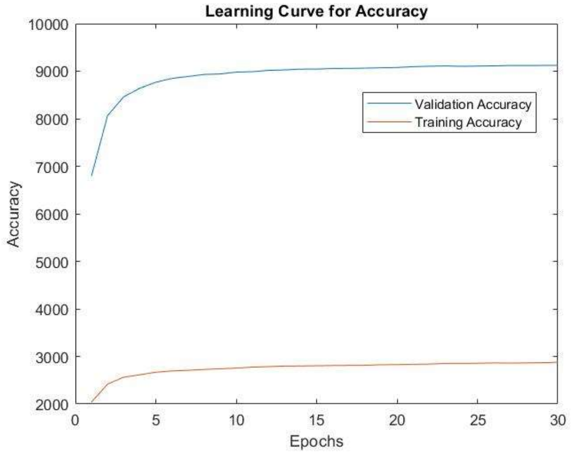

Figure 2: Learning curve of validation and training accuracy

II. Result

TABLE I. Training Details

Training (for final epoch)		**Validation (for final epoch)**		**Test**	
Accuracy (Out of 3000)	***Cost***	***Accuracy (Out of 10000)***	***Cost***	***Accuracy (Out of 10000)***	***Cost***
2877	0.3372	9123	0.5727	9132	0.5684

III. Derivative of Loss Function

The Cross-entropy loss function we are using is given below.

$$L_y = \sum(-y\log(a) - (1-y)\log(1-a))$$

After differentiate the above-mentioned equation with respect to a, we will get

$$\frac{\partial Ly}{\partial a} = \left(\frac{-y}{a} + \frac{1-y}{1-a}\right)\frac{\partial y}{\partial a}$$

Also, we know that

$$y = \sigma(a)$$

$$=> \frac{\partial y}{\partial a} = a(1-a)$$

This will give the final expression as

$$\frac{\partial Ly}{\partial a} = a - y$$

References

[1] http://neuralnetworksanddeeplearning.com/
[2] Lecture Notes ECE-542 Spring 2020