

# ELL-409 (Machine Intelligence and Learning)

*Assignment 1 - Back Propagation*

*Submitted by*

*Mihir Gupta, 2018PH10816*

## **Table of Contents**

- 1) **Effects of varying Learning rate (Eta)**
- 2) **Effects of varying Lambda (regularization parameter)**
- 3) **Effects of varying number of neurons in hidden layer**
- 4) **Effects of varying number of hidden layers**
- 5) **Effects of varying activation function (sigmoid vs tanh)**

**The plots for hyper-parameter tuning are generated during the training process using matplotlib**

### **1) Effect of varying Learning rate (Eta)**

It was observed that better accuracy could be obtained with lower learning rates in this dataset

```
Layer Structure = [784, 70, 10]  
Epochs = 100 , batchsize = 10 , lambda = 5 , varying eta
```

**eta = 0.001**

Epoch = 1 , accuracy = 11.952380952380953 %

Epoch = 2 , accuracy = 14.0 %

.....

Epoch = 100 , accuracy = 81.14285714285714 %

**eta = 0.003**

Epoch = 1 , accuracy = 17.333333333333332 %

Epoch = 2 , accuracy = 31.476190476190474 %

..... .

Epoch = 99 , accuracy = 87.0952380952381 %

Epoch = 100 , accuracy = 87.61904761904762 %

**eta = 0.009000000000000001**

Epoch = 1 , accuracy = 25.428571428571427 %

Epoch = 2 , accuracy = 30.238095238095237 %

.....

Epoch = 99 , accuracy = 83.76190476190476 %

Epoch = 100 , accuracy = 85.47619047619048 %

**eta = 0.027000000000000003**

Epoch = 1 , accuracy = 22.285714285714285 %

Epoch = 2 , accuracy = 31.571428571428573 %

.....

Epoch = 100 , accuracy = 67.80952380952381 %

**eta = 0.081000000000000002**

Epoch = 1 , accuracy = 11.619047619047619 %

Epoch = 2 , accuracy = 20.0 %

.....

Epoch = 100 , accuracy = 39.57142857142857 %

**eta = 0.243000000000000005**

Epoch = 1 , accuracy = 9.380952380952381 %

Epoch = 2 , accuracy = 9.666666666666666 %

..... .

Epoch = 100 , accuracy = 18.761904761904763 %

**eta = 0.729000000000000001**

Epoch = 1 , accuracy = 10.0 %

Epoch = 2 , accuracy = 9.619047619047619 %

..... .

Epoch = 100 , accuracy = 9.476190476190476 %

**eta = 2.1870000000000003**

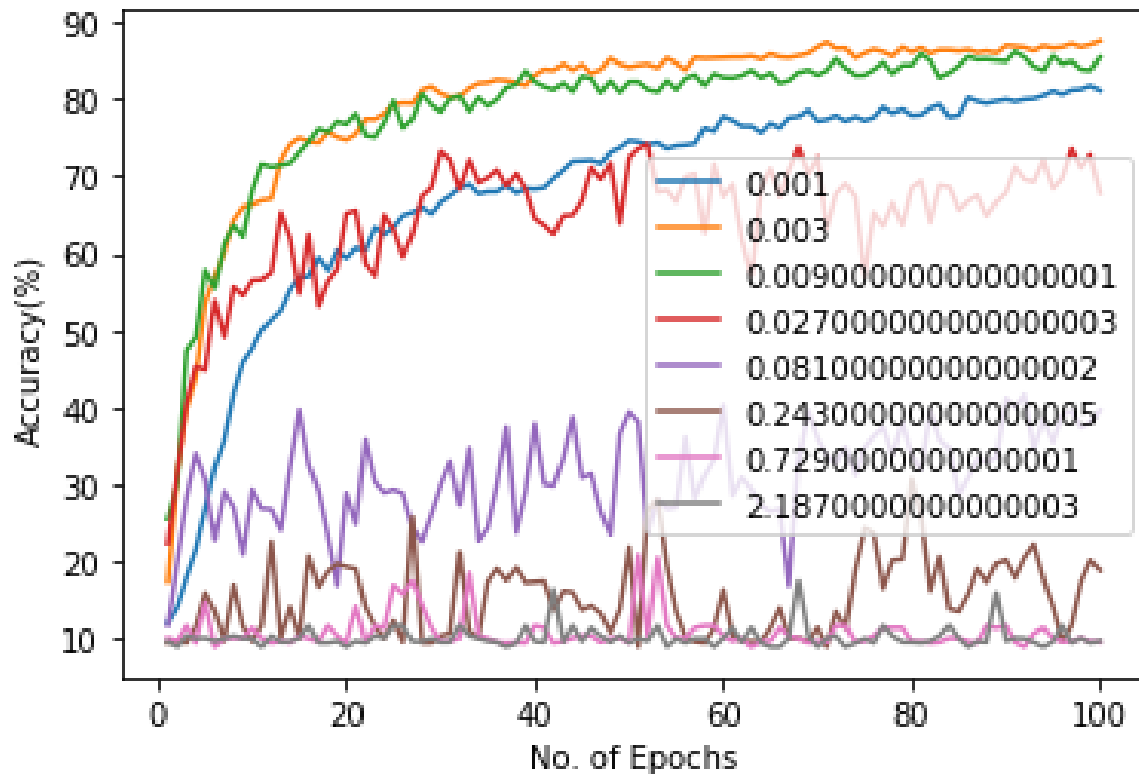
Epoch = 1 , accuracy = 9.571428571428571 %

Epoch = 2 , accuracy = 8.857142857142858 %

.....

Epoch = 99 , accuracy = 9.380952380952381 %

Epoch = 100 , accuracy = 9.619047619047619 %



**Plot1 : Effect of varying Learning Rate (Eta)**

## **2) Effect of varying Lambda (regularization parameter)**

Layer Structure = [784, 70, 10]

Epochs = 100 , batchsize = 10 , eta = 0.005 , varying lambda

**lambda = 0.1**

Epoch = 1 , accuracy = 22.666666666666668 %

Epoch = 2 , accuracy = 34.285714285714285 %

.....

Epoch = 99 , accuracy = 86.47619047619048 %  
Epoch = 100 , accuracy = 85.66666666666667 %

**lamba = 1**

Epoch = 1 , accuracy = 22.428571428571427 %  
Epoch = 2 , accuracy = 35.095238095238095 %  
..... .  
Epoch = 98 , accuracy = 86.0 %  
Epoch = 99 , accuracy = 86.80952380952381 %  
Epoch = 100 , accuracy = 87.14285714285714 %

**lamba = 5**

Epoch = 1 , accuracy = 18.952380952380953 %  
Epoch = 2 , accuracy = 34.142857142857146 %  
Epoch = 3 , accuracy = 47.142857142857146 %  
Epoch = 4 , accuracy = 44.142857142857146 %  
..... .  
Epoch = 99 , accuracy = 87.42857142857143 %  
Epoch = 100 , accuracy = 87.47619047619048 %

**lamba = 10**

Epoch = 1 , accuracy = 22.857142857142858 %  
Epoch = 2 , accuracy = 35.19047619047619 %  
Epoch = 3 , accuracy = 44.57142857142857 %  
.....  
Epoch = 100 , accuracy = 85.80952380952381 %

**lamba = 20**

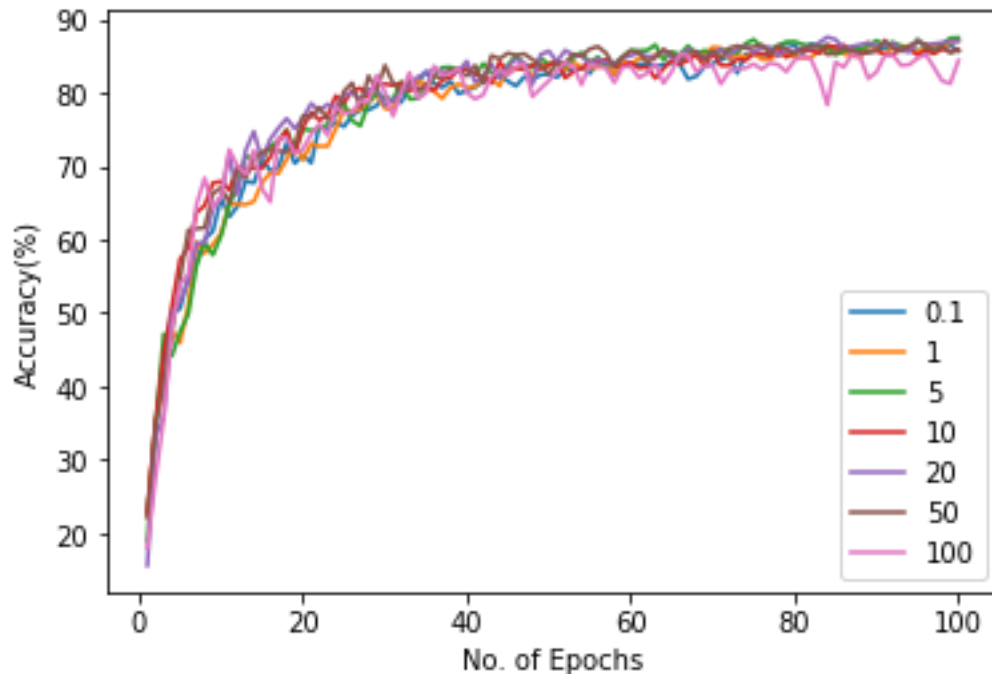
Epoch = 1 , accuracy = 15.619047619047619 %  
Epoch = 2 , accuracy = 33.142857142857146 %  
..... .  
Epoch = 100 , accuracy = 87.0 %

**lamba = 50**

Epoch = 1 , accuracy = 22.142857142857142 %  
Epoch = 2 , accuracy = 34.857142857142854 %  
..... .  
Epoch = 100 , accuracy = 85.80952380952381 %

**lamba = 100**

Epoch = 1 , accuracy = 18.0 %  
Epoch = 2 , accuracy = 27.523809523809526 %  
.....  
Epoch = 100 , accuracy = 84.38095238095238 %



**Plot2 : Effect of varying Lambda**

### **3) Effect of varying number of neurons in hidden layer**

It is seen that accuracy increases with number of neuros. Though, As no. of neurons increase , runtime becomes very high . Also, more neurons may lead to overfitting.

using 1 hidden layer and varying no of neurons

Epochs = 100 , eta = 0.005 , lambda = 7 ,batchsize = 10

#### **No. of neurons in hidden layer = 10**

Epoch = 1 , accuracy = 9.380952380952381 %

Epoch = 2 , accuracy = 9.380952380952381 %

Epoch = 3 , accuracy = 9.380952380952381 %

.....

Epoch = 99 , accuracy = 43.04761904761905 %

Epoch = 100 , accuracy = 43.285714285714285 %

#### **No. of neurons in hidden layer = 30**

Epoch = 1 , accuracy = 12.428571428571429 %

Epoch = 2 , accuracy = 15.095238095238095 %

.....

Epoch = 100 , accuracy = 78.76190476190476 %

#### **No. of neurons in hidden layer = 70**

Epoch = 1 , accuracy = 21.952380952380953 %

Epoch = 2 , accuracy = 35.095238095238095 %  
 .....  
 Epoch = 100 , accuracy = 85.23809523809524 %

**No. of neurons in hidden layer = 100**

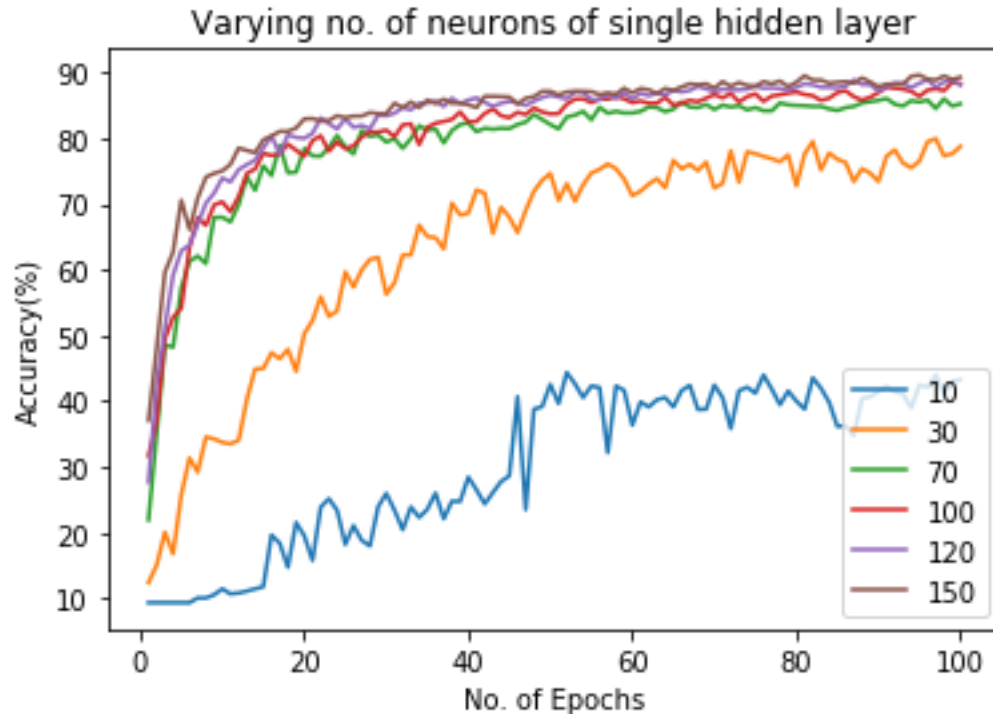
Epoch = 1 , accuracy = 31.714285714285715 %  
 Epoch = 2 , accuracy = 36.80952380952381 %  
 .....  
 Epoch = 100 , accuracy = 88.28571428571429 %

**No. of neurons in hidden layer = 120**

Epoch = 1 , accuracy = 27.714285714285715 %  
 Epoch = 2 , accuracy = 42.714285714285715 %  
 .....  
 Epoch = 100 , accuracy = 88.14285714285714 %

**No. of neurons in hidden layer = 150**

Epoch = 1 , accuracy = 37.095238095238095 %  
 Epoch = 2 , accuracy = 48.095238095238095 %  
 .....  
 Epoch = 100 , accuracy = 89.28571428571429 %



**Plot4 : Effect of varying number of neurons**

**4) Effect of varying number of hidden layers**

For this dataset best accuracy was obtained with a single hidden layer.

varying hidden layers

Epochs = 100 , eta = 0.005 , lambda = 7 ,batchsize = 10

**No. of hidden layers = 1**

Epoch = 1 , accuracy = 13.666666666666666 %

Epoch = 2 , accuracy = 17.428571428571427 %

.....

Epoch = 100 , accuracy = 85.71428571428571 %

**No. of hidden layers = 2**

Epoch = 1 , accuracy = 10.619047619047619 %

Epoch = 2 , accuracy = 9.523809523809524 %

.....

Epoch = 100 , accuracy = 43.23809523809524 %

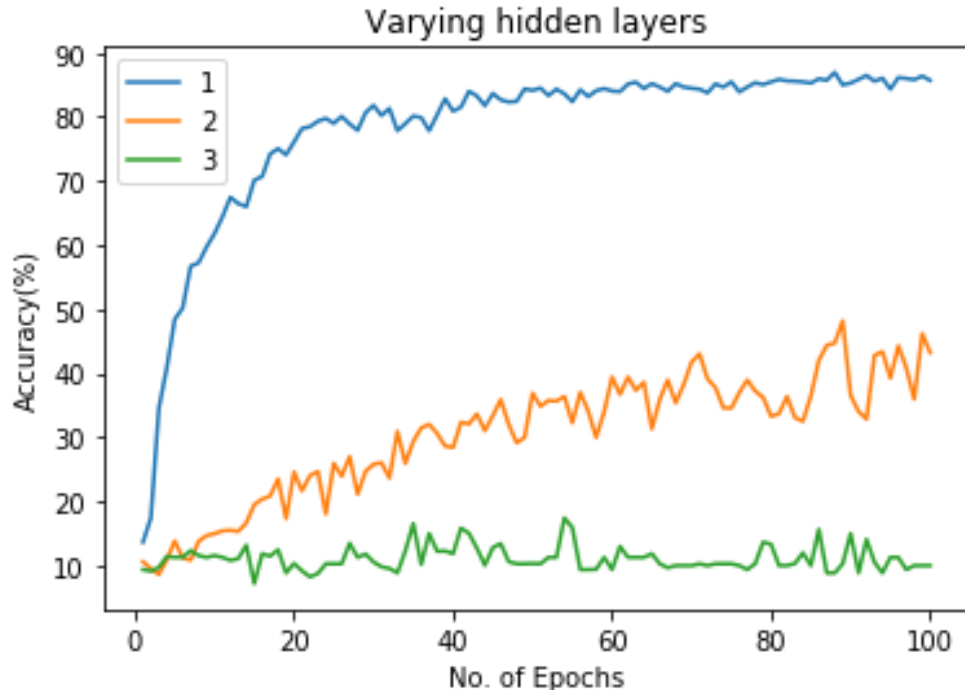
**No. of hidden layers = 3**

Epoch = 1 , accuracy = 9.380952380952381 %

Epoch = 2 , accuracy = 9.142857142857142 %

.....

Epoch = 100 , accuracy = 10.0 %



**Plot5 : Effect of varying number of hidden layers**

**5) Effects of varying activation function (sigmoid vs tanh)**

Similar accuracy was obtained using both sigmoid and tanh

Comparing sigmoid and tanh

Epochs = 100 , eta = 0.005 , lambda = 7 ,batchsize = 10

**using sigmoid**

Epoch = 1 , accuracy = 21.666666666666668 %

Epoch = 2 , accuracy = 35.714285714285715 %

Epoch = 3 , accuracy = 38.904761904761905 %

.....

Epoch = 99 , accuracy = 85.57142857142857 %

Epoch = 100 , accuracy = 86.38095238095238 %

**using tanh**

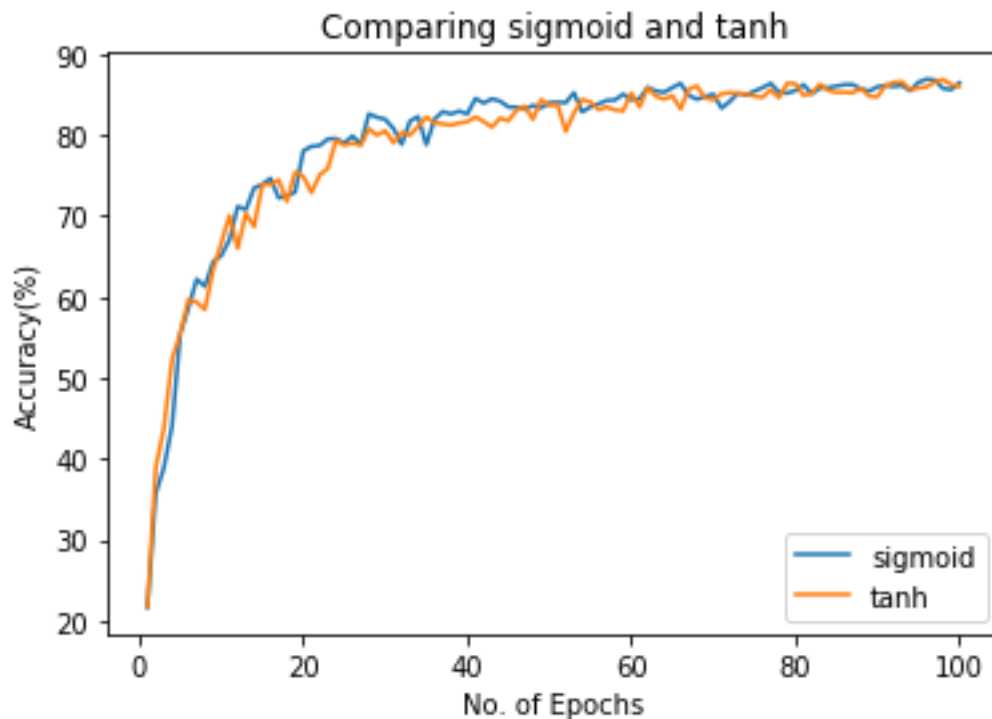
Epoch = 1 , accuracy = 21.80952380952381 %

Epoch = 2 , accuracy = 38.857142857142854 %

.....

Epoch = 99 , accuracy = 86.19047619047619 %

Epoch = 100 , accuracy = 85.95238095238095 %



**Plot6 : Effect of varying number of hidden layers**