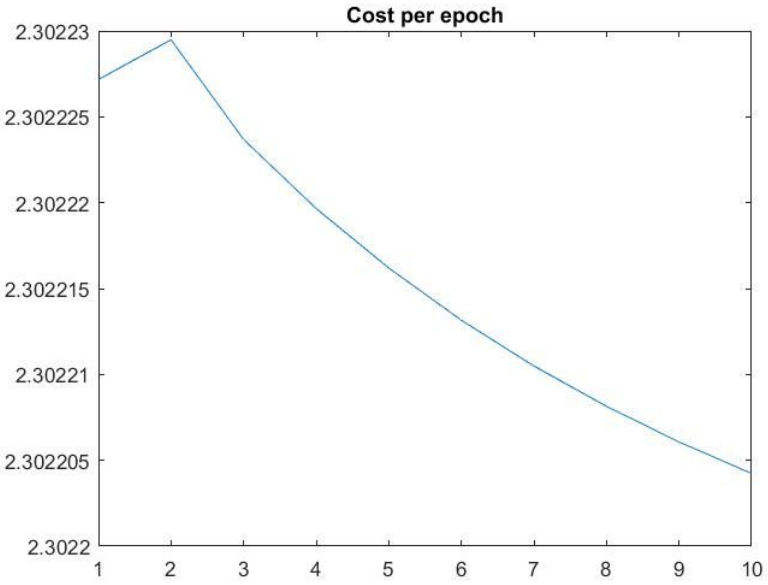
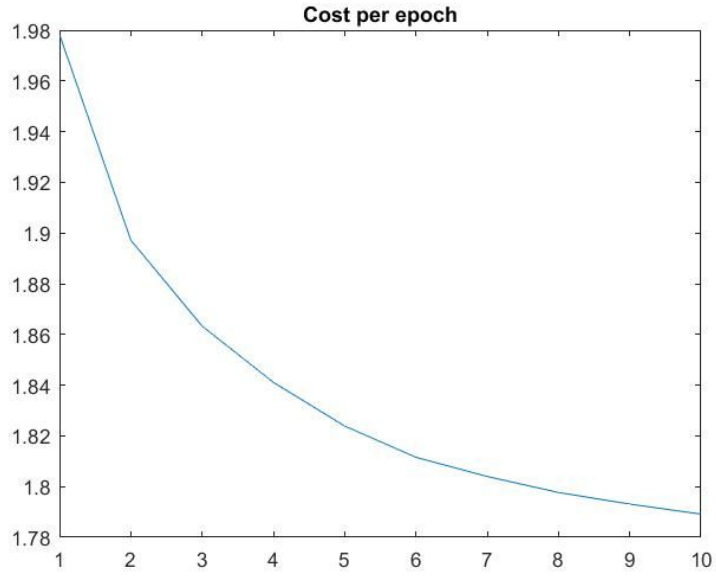
Deep Learning Assignment 3
Arsalan Syed 9611176331
CDATE 3
DD2424

The algorithm used to calculate the gradient of the cost function was successful since the difference the difference in the analytic result and the numerical result was on average 1e-6.

To test the effects of batch normalization on a 3 layer network, the network was trained two times using the same parameters. The only difference that in the second test, the batch normalization functions for the forward and backward pass were used. On the left is the cost function when no batch normalization was used whereas the diagram on the right uses batch normalization.

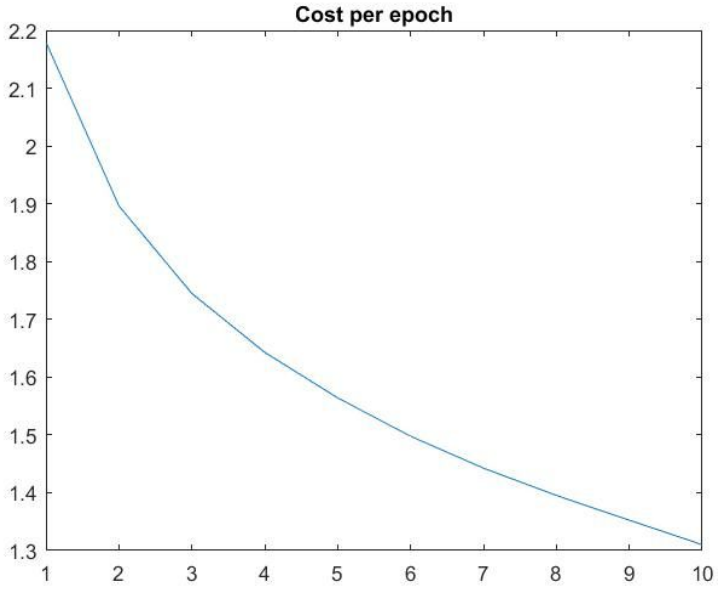


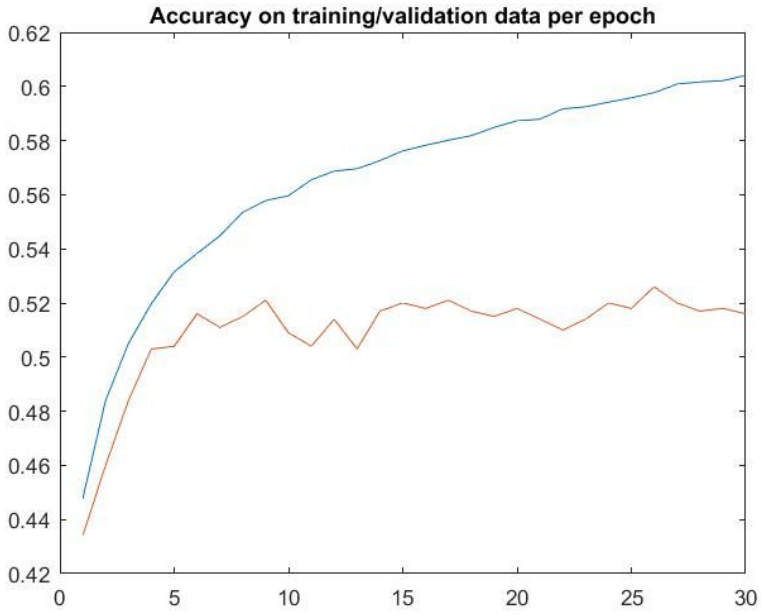


When searching for optimal parameters, 50 random values for an eta and lambda were generated and used for training over 10 epochs. Eta was in the range [0.00001,0.1] and lambda was in the range [0.000001,0.001]. From this the best 4 parameter settings were chosen.

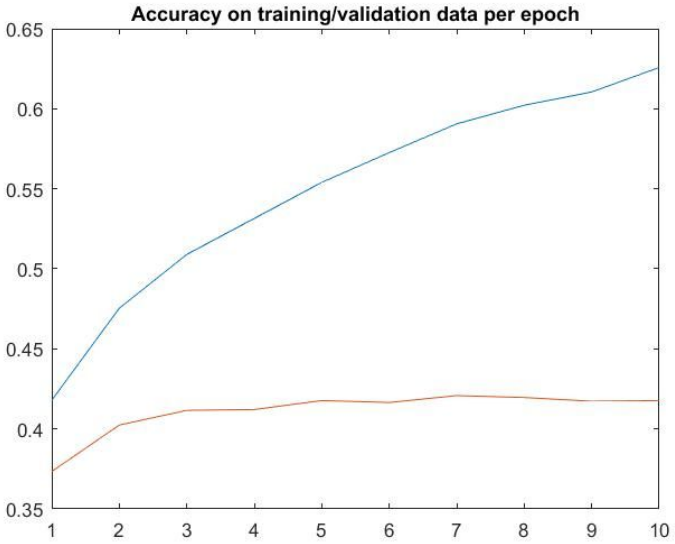
Lambda	Eta	Acc
7.8256e-04	0.0592	0.4442
2.3187e-04	**0.0498**	**0.4438**
3.7402e-04	0.0548	0.4435
4.0975e-04	0.0498	0.4457

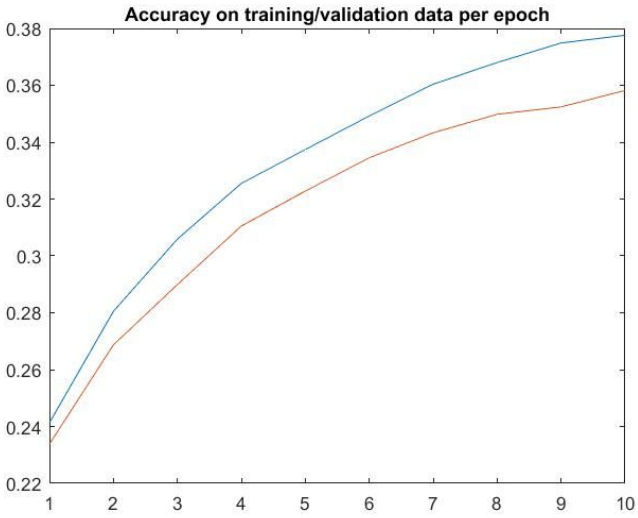
The best result was obtained by using the parameter settings in bold, using all available data and letting it run for 30 epochs. A test accuracy of 51.67% was obtained. Below are the plots obtained by this model.

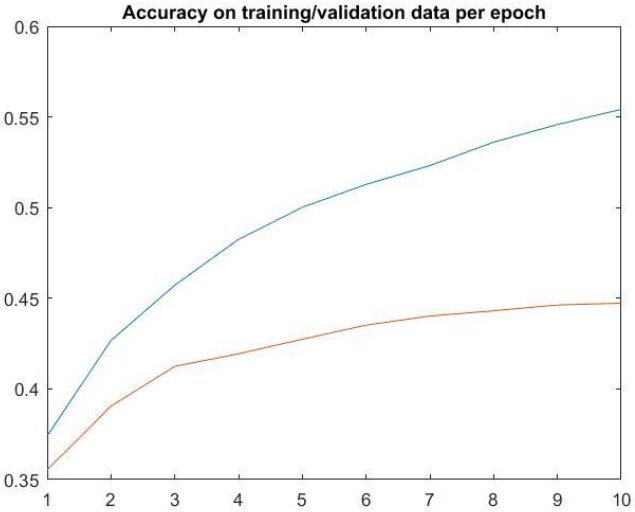


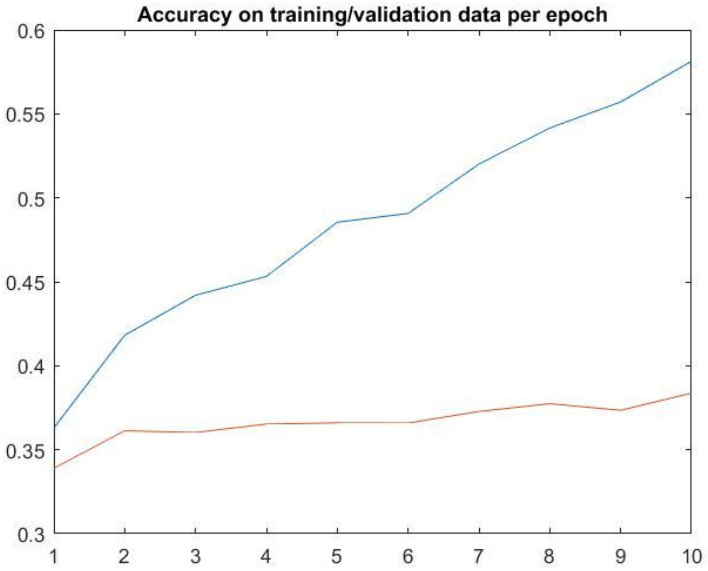


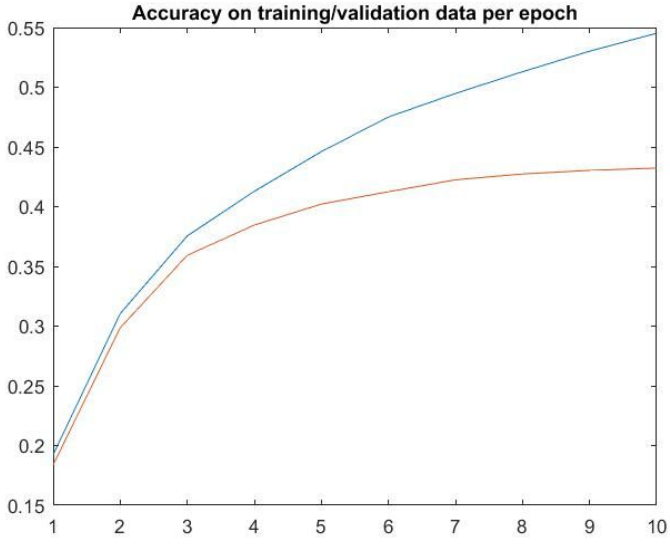
To test the effects of batch normalization on a 2 layer network, 3 different learning rates were used which were considered high, medium and low (0.1,0.01,0.001). The first 3 graphs are with batch normalization and the last 3 are without it.

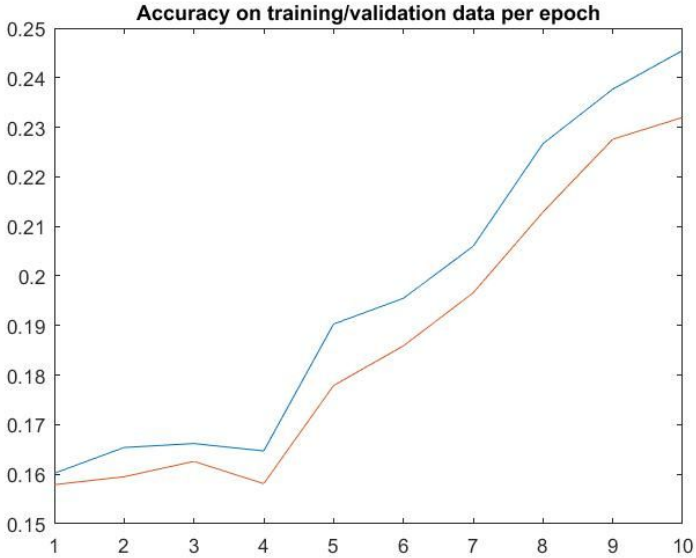

Accuracy on training/validation data per epoch
0.65
0.6
0.55
0.5
0.45
0.4
0.35
1
2
3
4
5
6
7
8
9
10


Accuracy on training/validation data per epoch
0.38
0.36
0.34
0.32
0.3
0.28
0.26
0.24
0.22
1
2
3
4
5
6
7
8
9
10


Accuracy on training/validation data per epoch
0.6
0.55
0.5
0.45
0.4
0.35
1
2
3
4
5
6
7
8
9
10


Accuracy on training/validation data per epoch
0.6
0.55
0.5
0.45
0.4
0.35
0.3
1
2
3
4
5
6
7
8
9
10


Accuracy on training/validation data per epoch
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
1
2
3
4
5
6
7
8
9
10


Accuracy on training/validation data per epoch
0.25
0.24
0.23
0.22
0.21
0.2
0.19
0.18
0.17
0.16
0.15
1
2
3
4
5
6
7
8
9
10