Name: Duy Nguyen

ROB 537 – HW 1

Using training set 1

A. Hidden Unit and Time training

I train the data in 500 epochs with the hidden unit goes from 0 to 90 with a 10 units increment. Learning rate is at 0.0002, momentum = 0.7, L2 penalty = 0.0002

As the training time increase, the accuracy also increases. However, when the training process reaches a maximal point, the accuracy stops increasing and becomes stables. For the hidden units, from 0 to around 50 hidden units, the rate of going to the stable accuracy increases. However, after 50 hidden units, the rate of going to the stable point starts to decrease. It happens to all 3-test data. However, looking at the result when there is no hidden unit (the program picks the same value for all the answer), I can see that the test 2 and test 3 are extremely bias and in different direction.

B. Learning Rate

I train the data in 500 epochs with the learning rate goes from 0 to 0.0018 with a 0.0002 increment. Hidden unit is 50, momentum = 0.7, L2 penalty = 0.0002

As the learning rate increases, the converging tends to increase. However, the learning rate does not change that much after 0.0006. I guess that my learning rate increment is a little small to see any significant changes in different learning rate. Test 2 and test 3 fluctuate more around the stable accuracy (85%) before converging to its than test 1. I guess that is because test 2 and 3 are more bias than test 1

C. Momentum

I train the data in 500 epochs with the momentum goes from 0 to 0.9 with a 0.1 increment. Hidden unit is 50, learning rate = 0.0008, L2 penalty = 0.0002

In general, I can see that the converging rate increases as the momentum increases. However, that increment becomes insignificant after 0.6. Test 1 tends to converge faster than test 2 and test 3. Test 2 and test 3 fluctuate around the converging point more than test 1. Again, the reason could be because of the bias of test 2 and test 3

D. L2 Penalty

I train the data in 500 epochs with the L2 penalty goes from 0 to 0.0018 with a 0.0002 increment. Hidden unit is 50, learning rate = 0.0008, momentum = 0.7

I think there is an increment in the converging rate as L2 Penalty increase. However, that increment is not that much. Test 1 tends to converge faster than test 2 and test 3.

Using training set 2

A. Hidden Unit and Time training

I train the data in 500 epochs with the hidden unit goes from 0 to 90 with a 10 units increment. Learning rate is at 0.0002, momentum = 0.7, L2 penalty = 0.0002

As the training time increase, the accuracy also increases. However, in test 1 and test 3 the algorithm takes a lot longer to learn than training set 1. In the contract, test 2 takes almost no time to learn the correct function. The rate of going to the stable point increase around 40 hidden units for test 1 and 10 hidden units for test 3. The rate of going to the stable points almost the same for all different hidden units in test 2.

B. Learning Rate

I train the data in 500 epochs with the learning rate goes from 0 to 0.0018 with a 0.0002 increment. Hidden unit is 50, momentum = 0.7, L2 penalty = 0.0002

For test 1, I can see clearly that as the learning rate increases, the rate of converging increases. The converging point is around 65 % accuracy. For test 2, there is no effect of the learning rate on the data, it starts off with the really high accuracy and the stable point is around 91% accuracy. For test 3, the learning rate have similar influence on the converging rate as test 1, however, the converging accuracy is around 40%. I think is because of the bias in training and test data sets

C. Momentum

I train the data in 500 epochs with the momentum goes from 0 to 0.9 with a 0.1 increment. Hidden unit is 50, learning rate = 0.0008. L2 penalty = 0.0002

There is a clear increment in the converging rate of test 1 and test 3 when the momentum increases. There is no effect of momentum on the converging for test 2. Test 2 has the highest stable accuracy (around 91%), test 1 is the second (around 65 %), and test 3 has the lowest stable accuracy (around 40). Again, I think it is due to the bias of the training and test data. I think is because of the bias in training and test data sets

D. L2 Penalty

I train the data in 500 epochs with the L2 penalty goes from 0 to 0.0018 with a 0.0002 increment. Hidden unit is 50, learning rate = 0.0008, momentum = 0.7

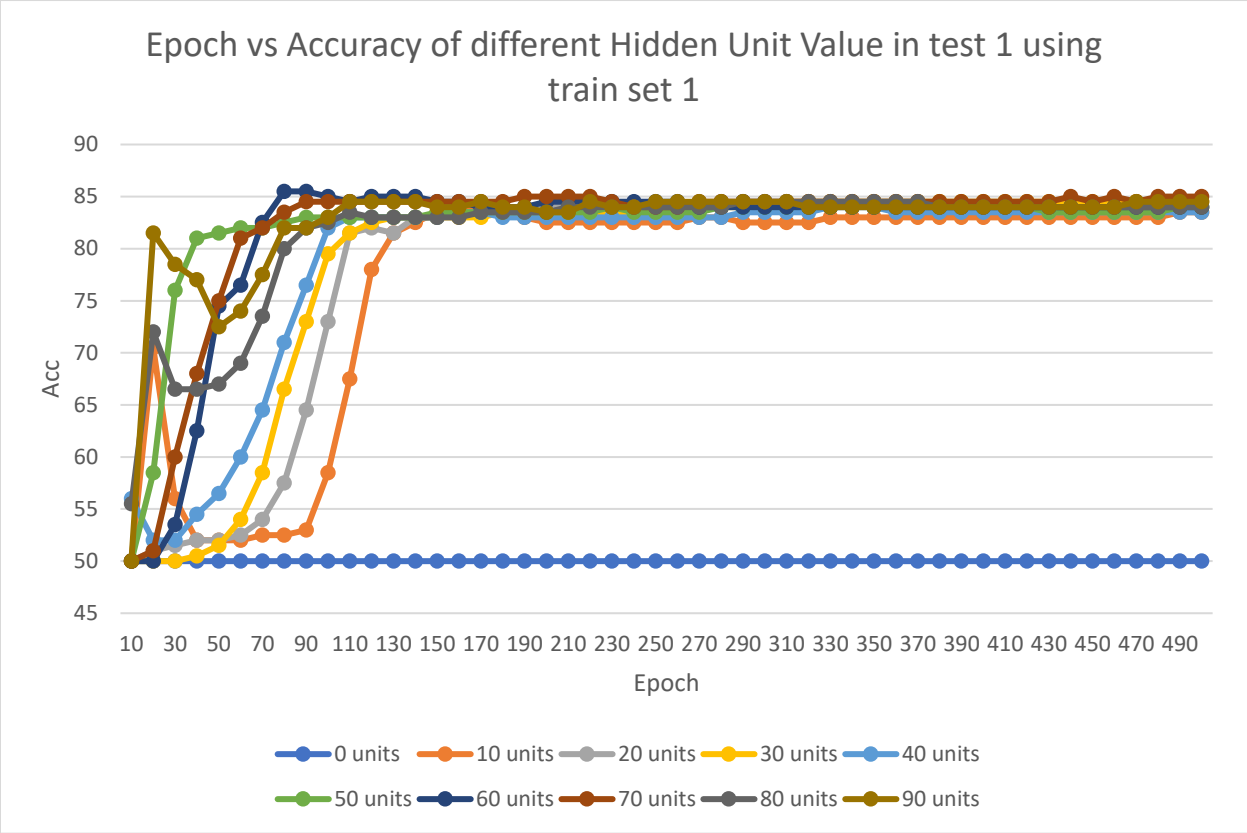
L2 Penalty has a lot influences on test 1 and test 3 result as there is big improvement in the converging rate as L2 increases. L2 has no influence on the converging rate of test 2. Test 2 has the highest stable accuracy (around 91%), test 1 is the second (around 65 %), and test 3 has the lowest stable accuracy (around 40). Again, I think it is due to the bias of the training and test data. I think is because of the bias in training and test data sets

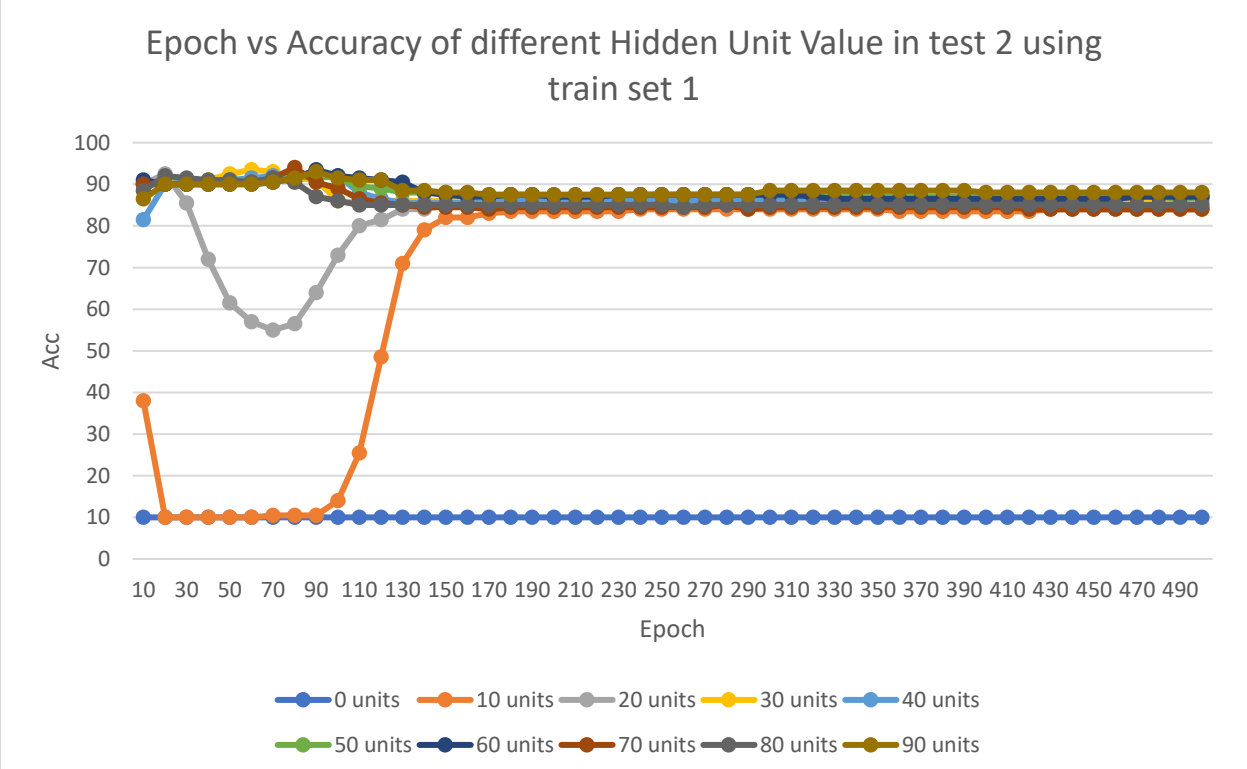
Conclusion

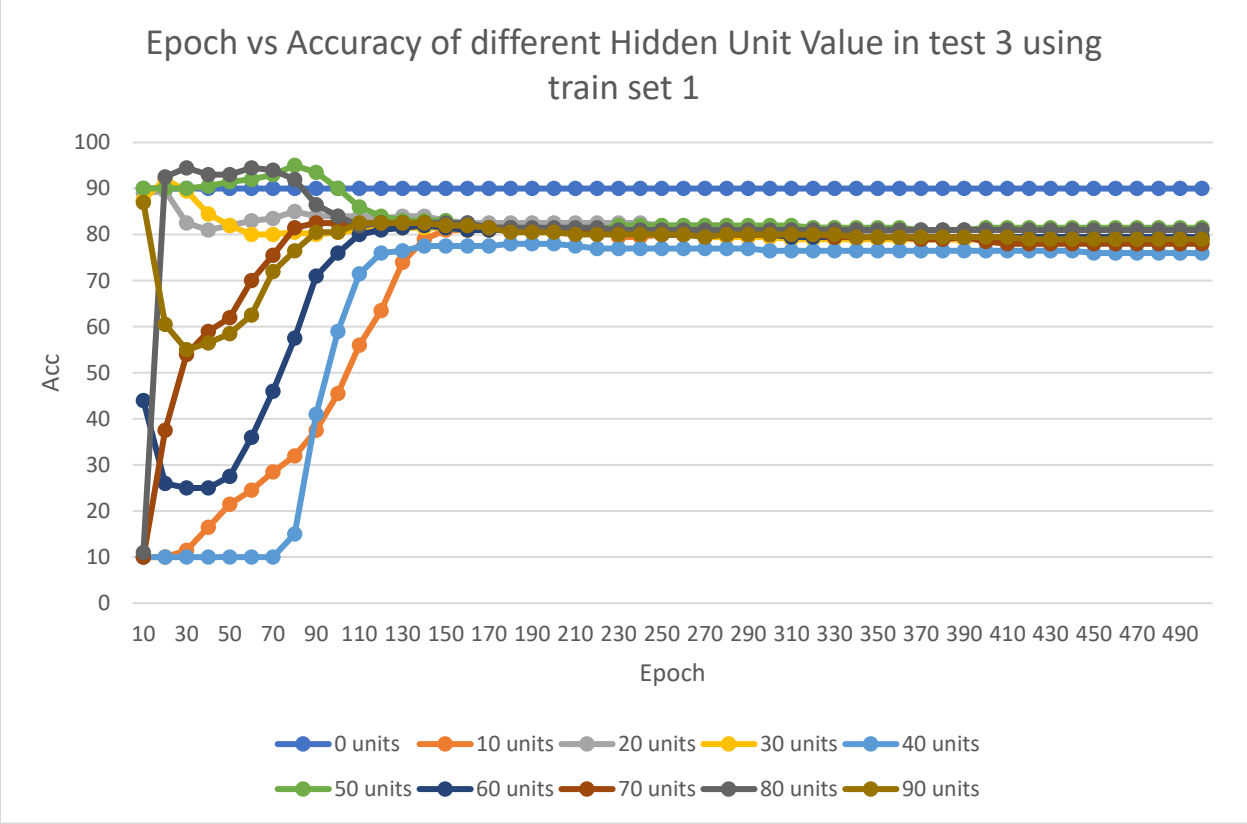
I can conclude that the training set 2 is bias and its bias is similar to the test 2 because it takes extremely long time to train in a balanced data set and extremely fast in a similar bias data. When the training data is not bias, there is not a lot influences from Momentum and L2 Penalty, however, when the training data is bias, Momentum and L2 Penalty improve the performance of the algorithm a lot.

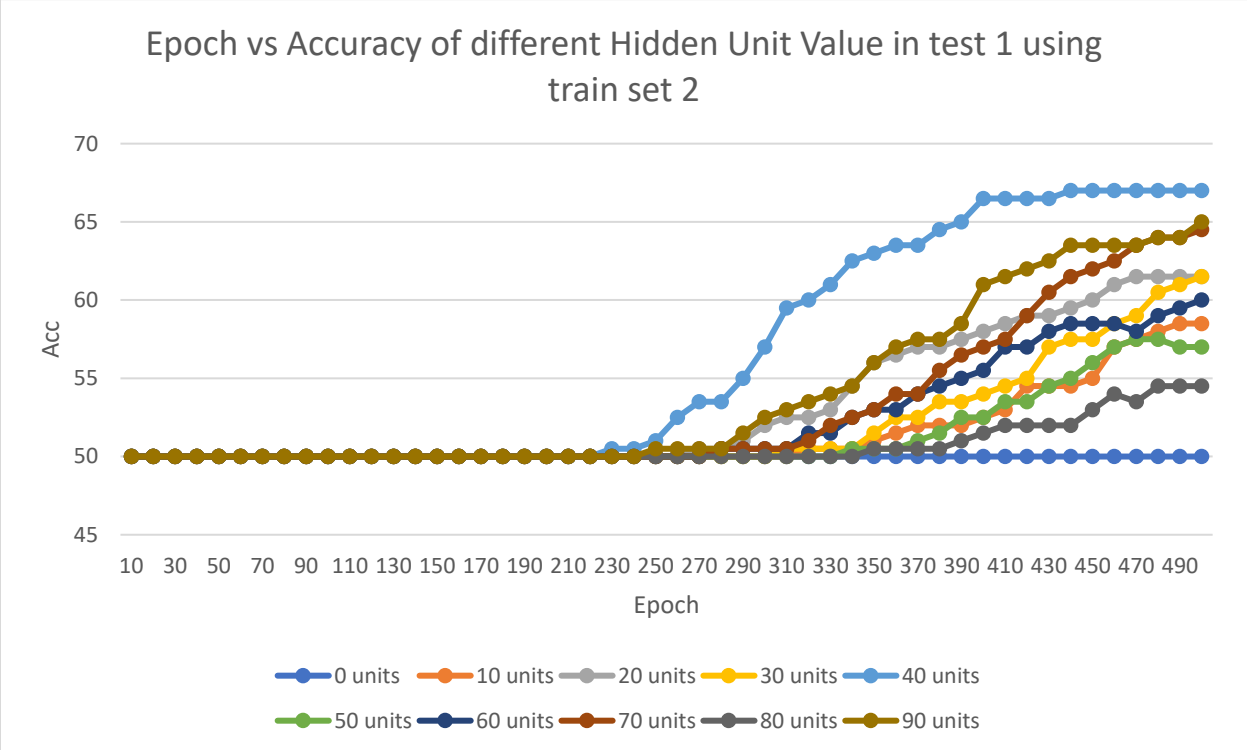
I think that one way to fit the problem is to combine the two-training data set together and flip the other of them every time we run a new epoch. Additionally, we can select small batch of that combined data by randomly pick with replacement some element in the training set and do the training with its.

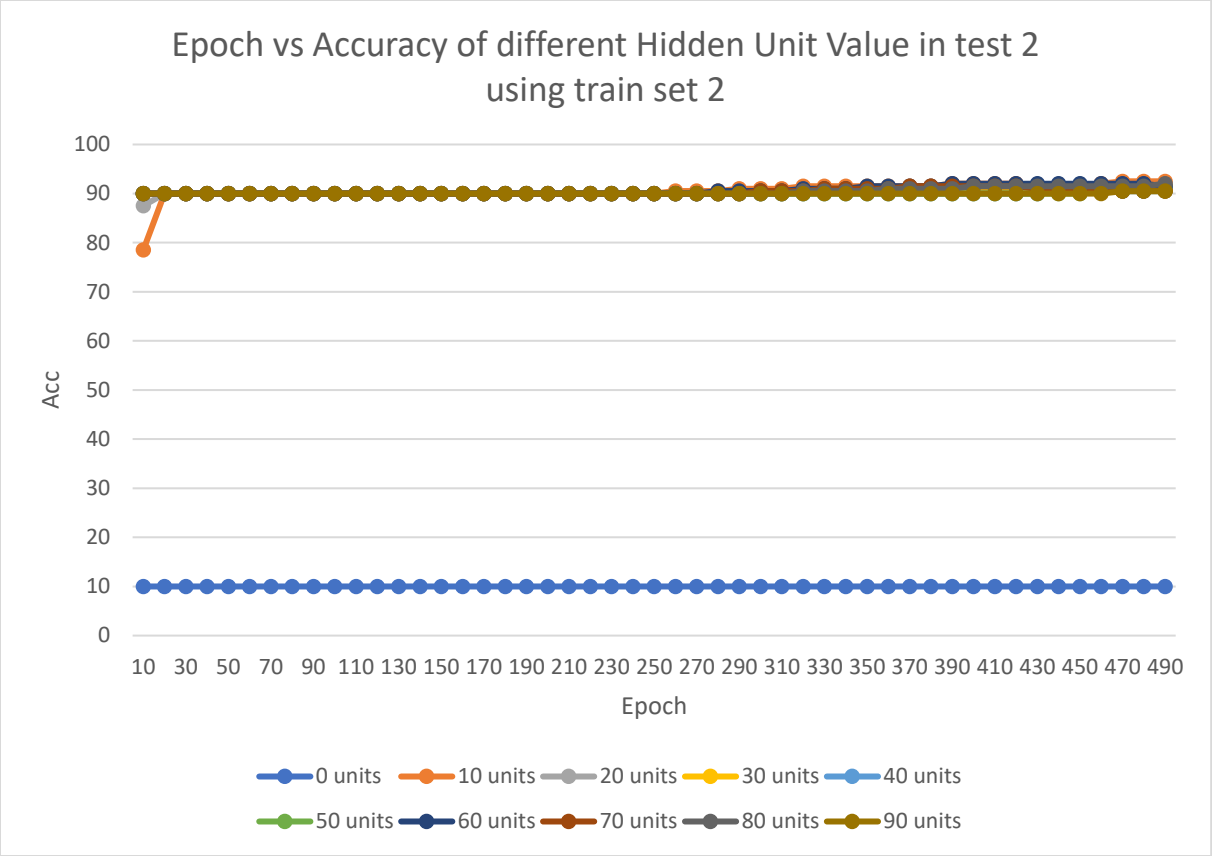
NOTE: all the graph data is next page

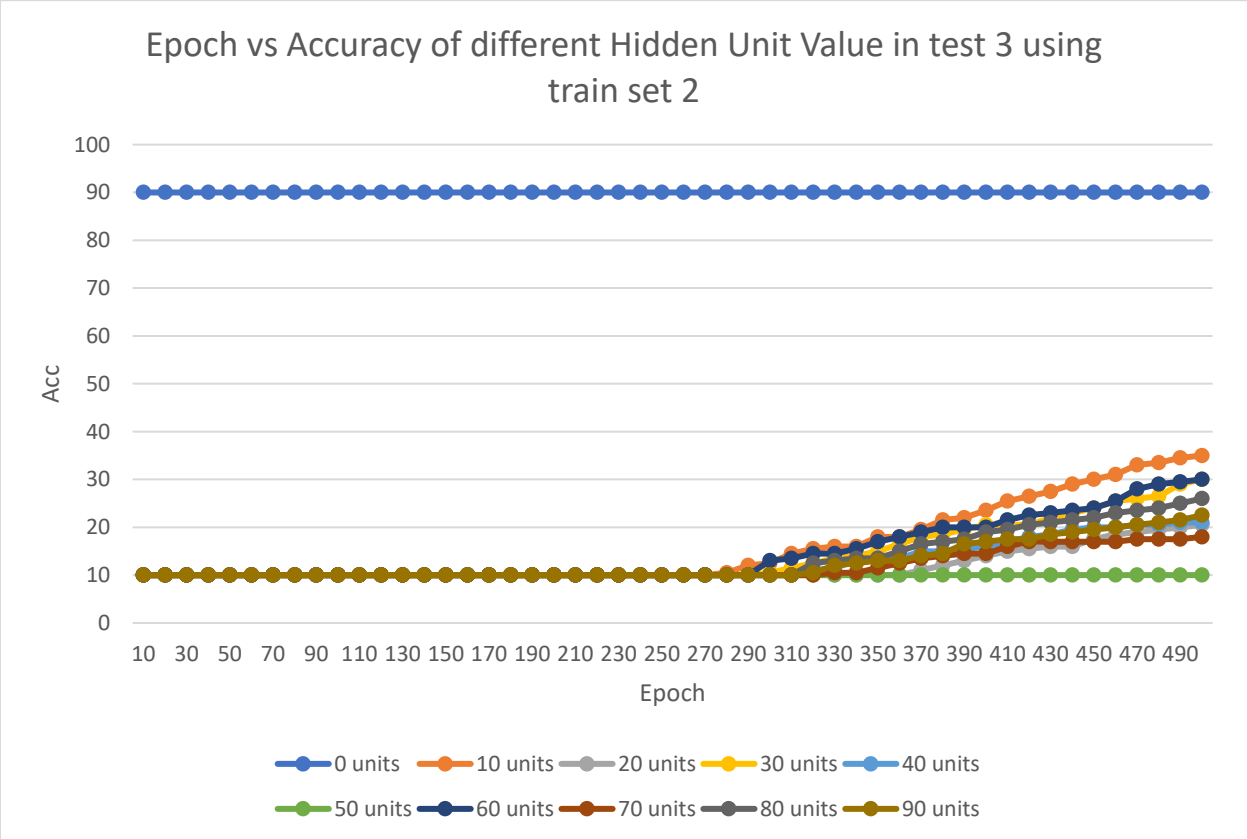

Epoch vs Accuracy of different Hidden Unit Value in test 1 using train set 1
Acc
90
85
80
75
70
65
60
55
50
45
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

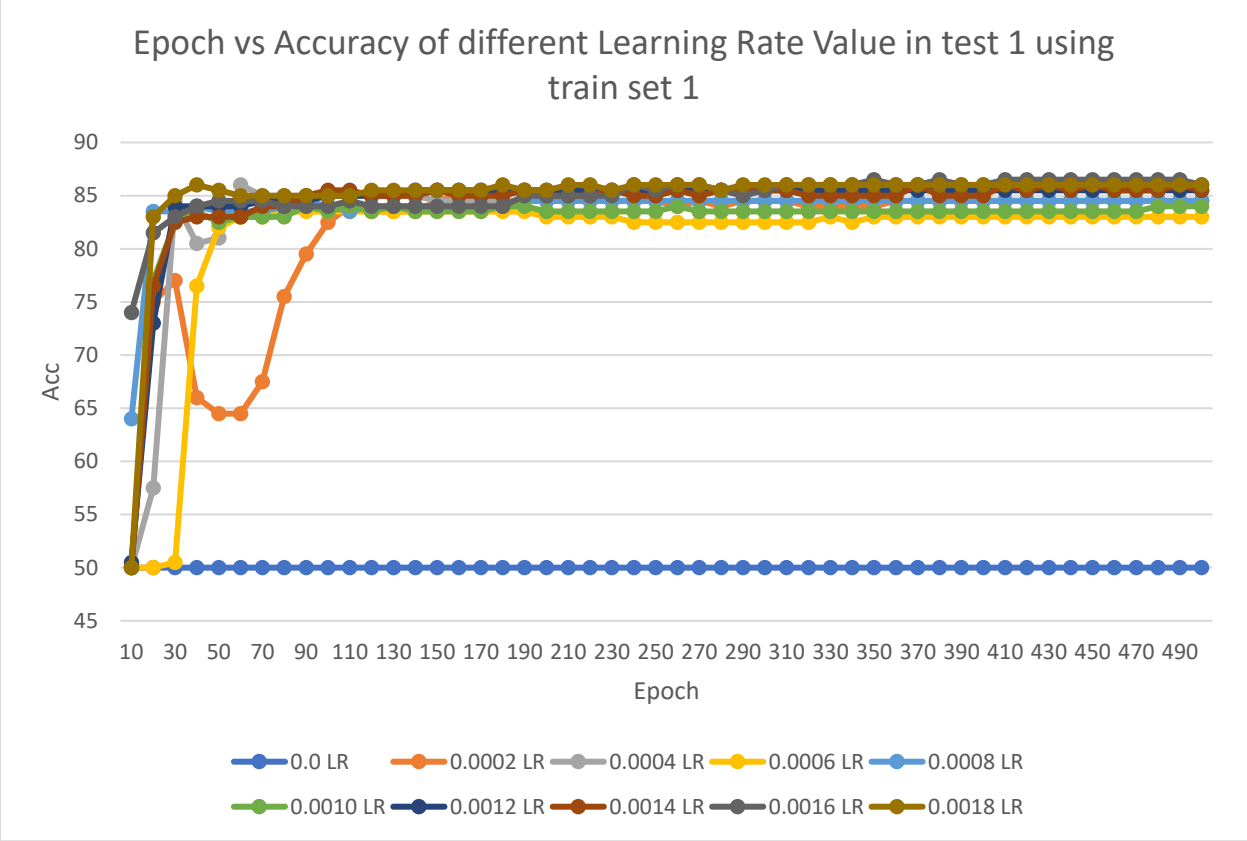

Epoch vs Accuracy of different Hidden Unit Value in test 2 using train set 1
Acc
100
90
80
70
60
50
40
30
20
10
0
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

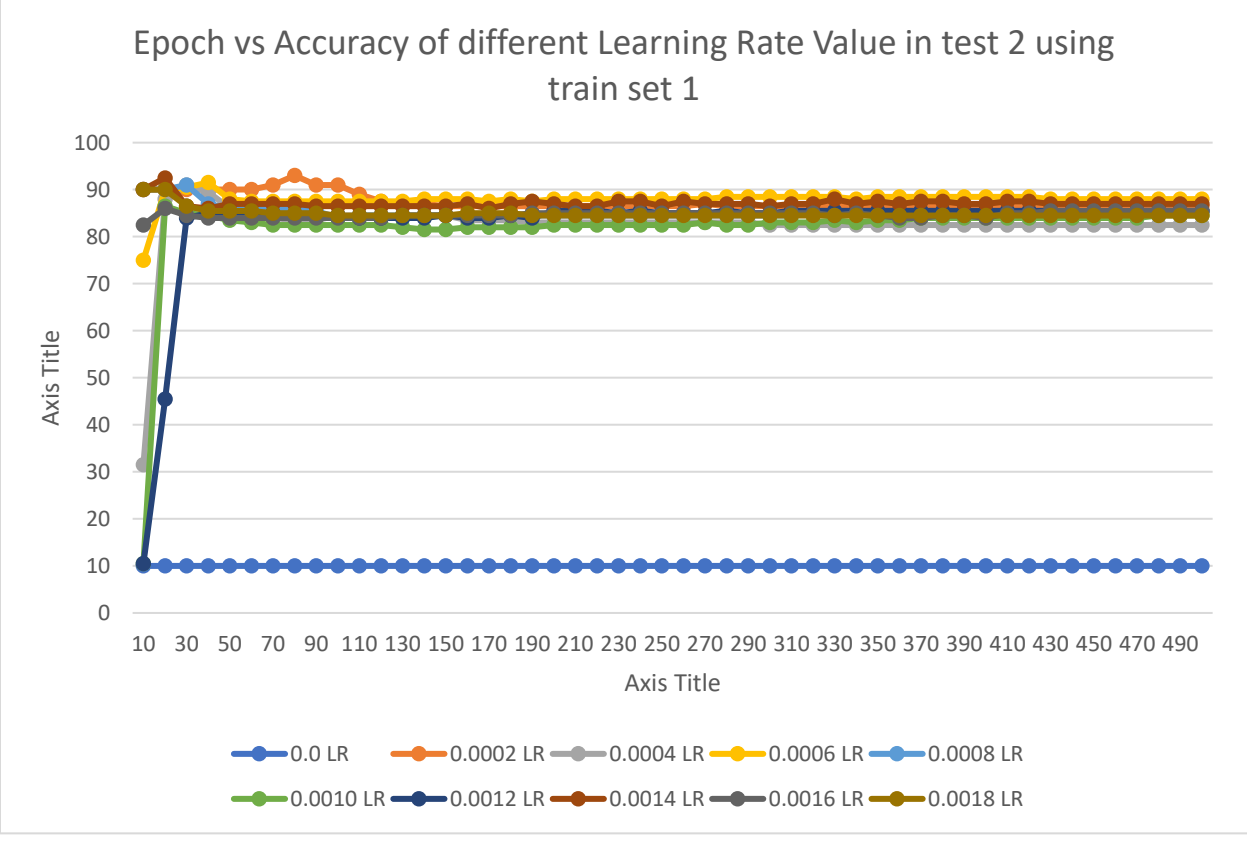

Epoch vs Accuracy of different Hidden Unit Value in test 3 using train set 1
Acc
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

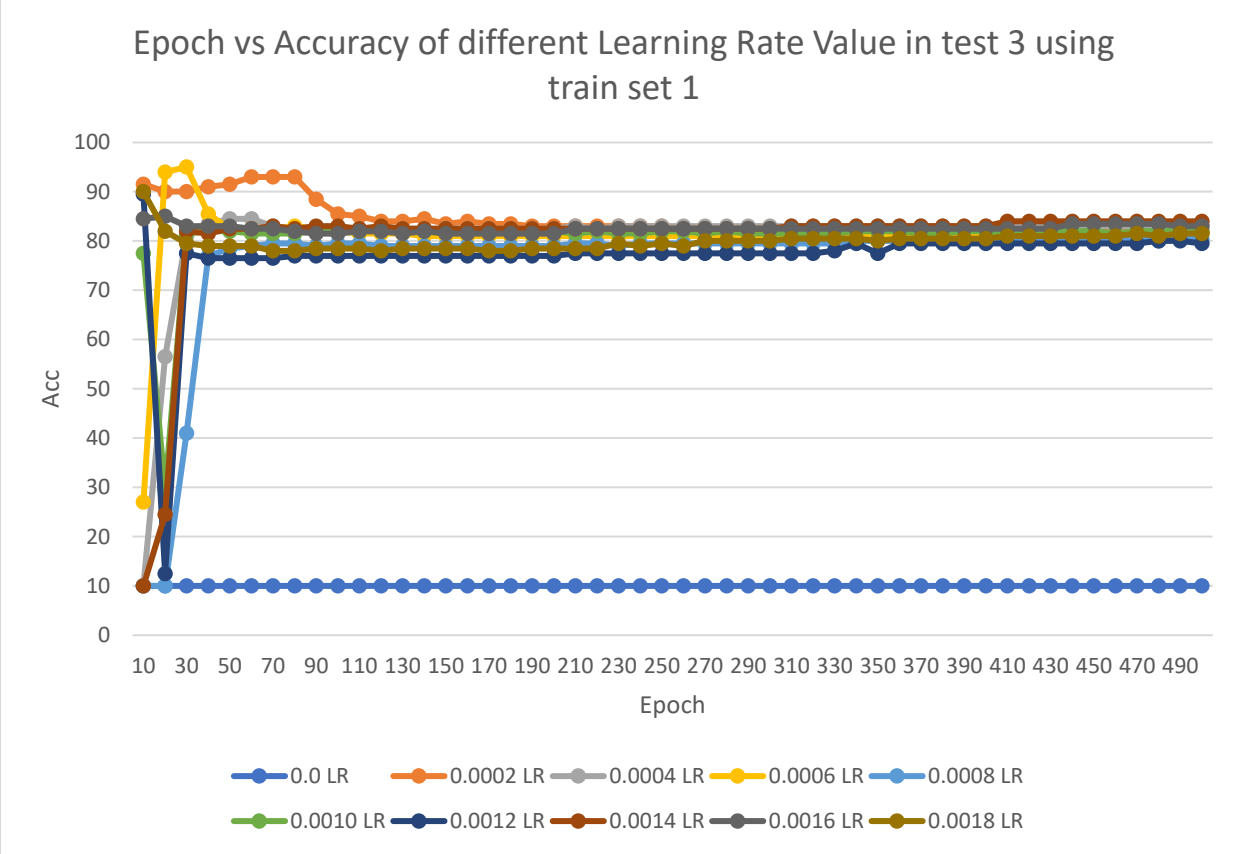

Epoch vs Accuracy of different Hidden Unit Value in test 1 using train set 2
Acc
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

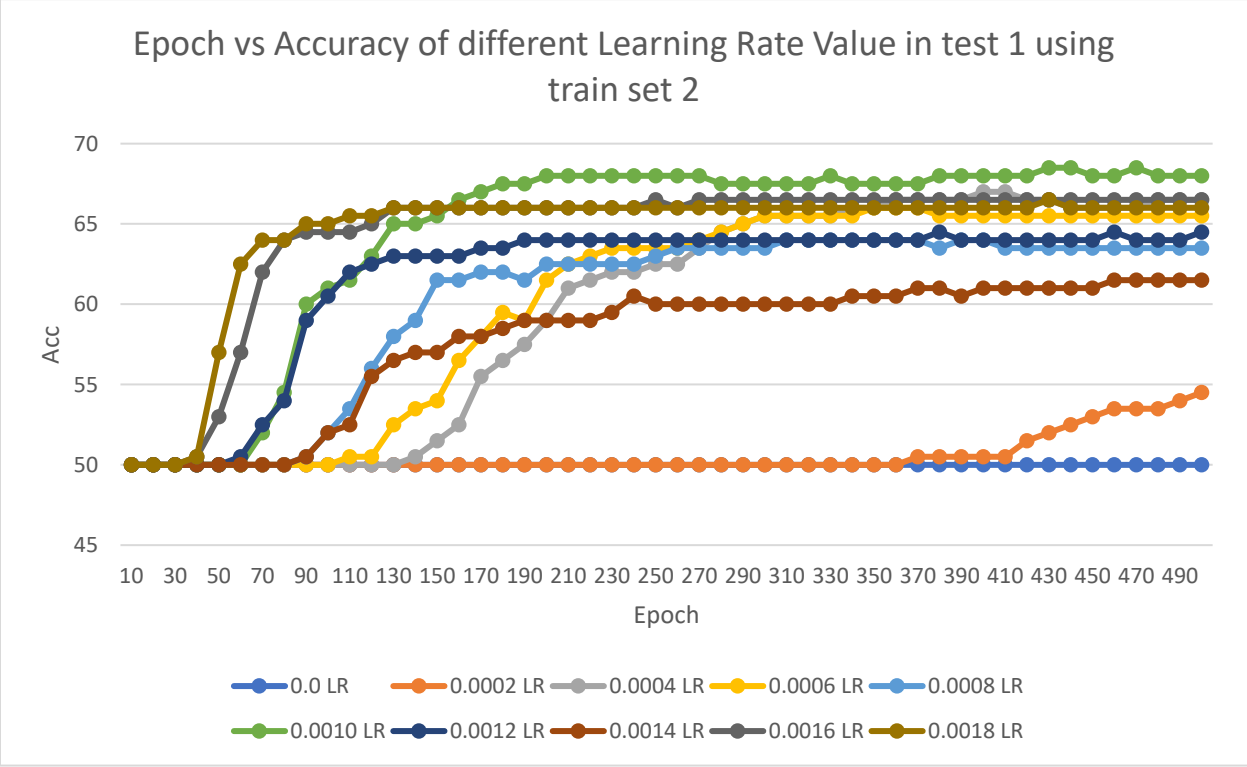

Epoch vs Accuracy of different Hidden Unit Value in test 2 using train set 2
Acc
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

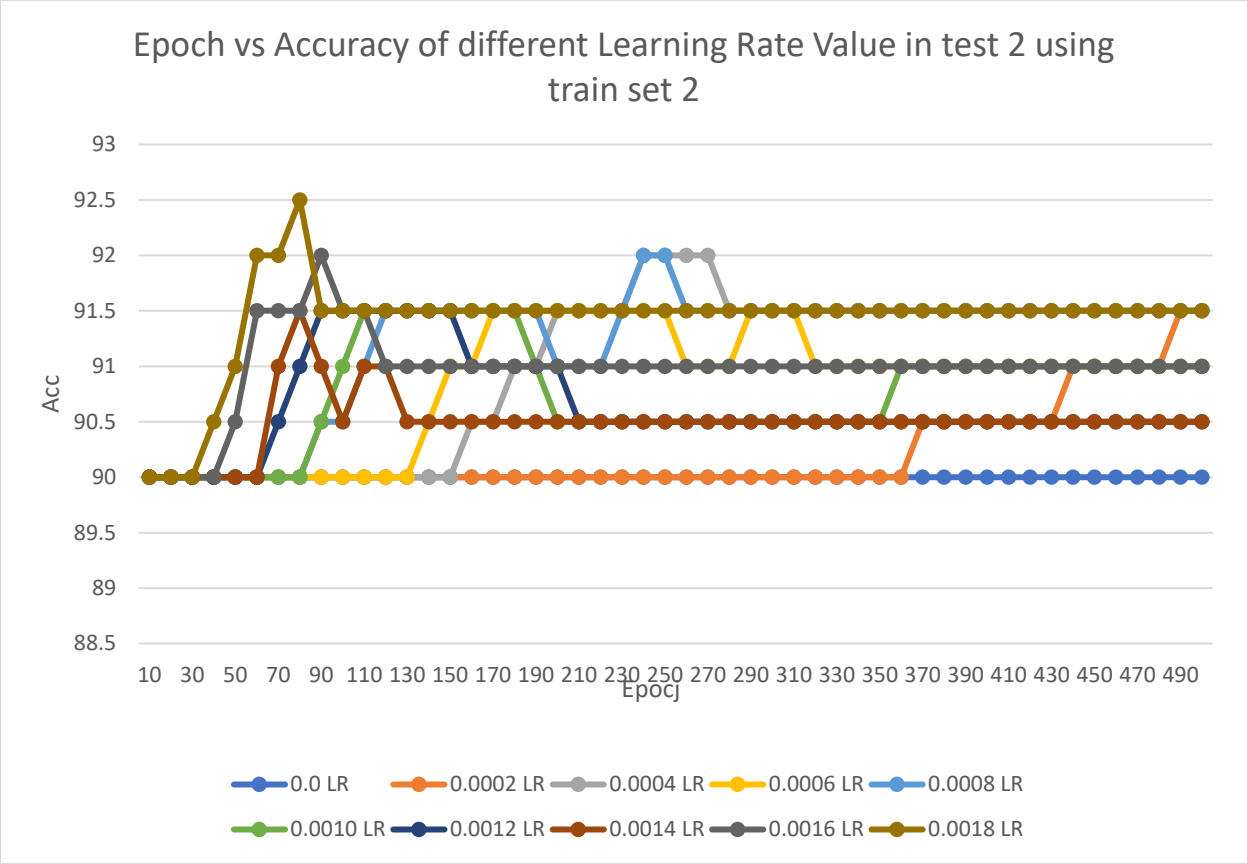

Epoch vs Accuracy of different Hidden Unit Value in test 3 using train set 2
Acc
Epoch
0 units
10 units
20 units
30 units
40 units
50 units
60 units
70 units
80 units
90 units

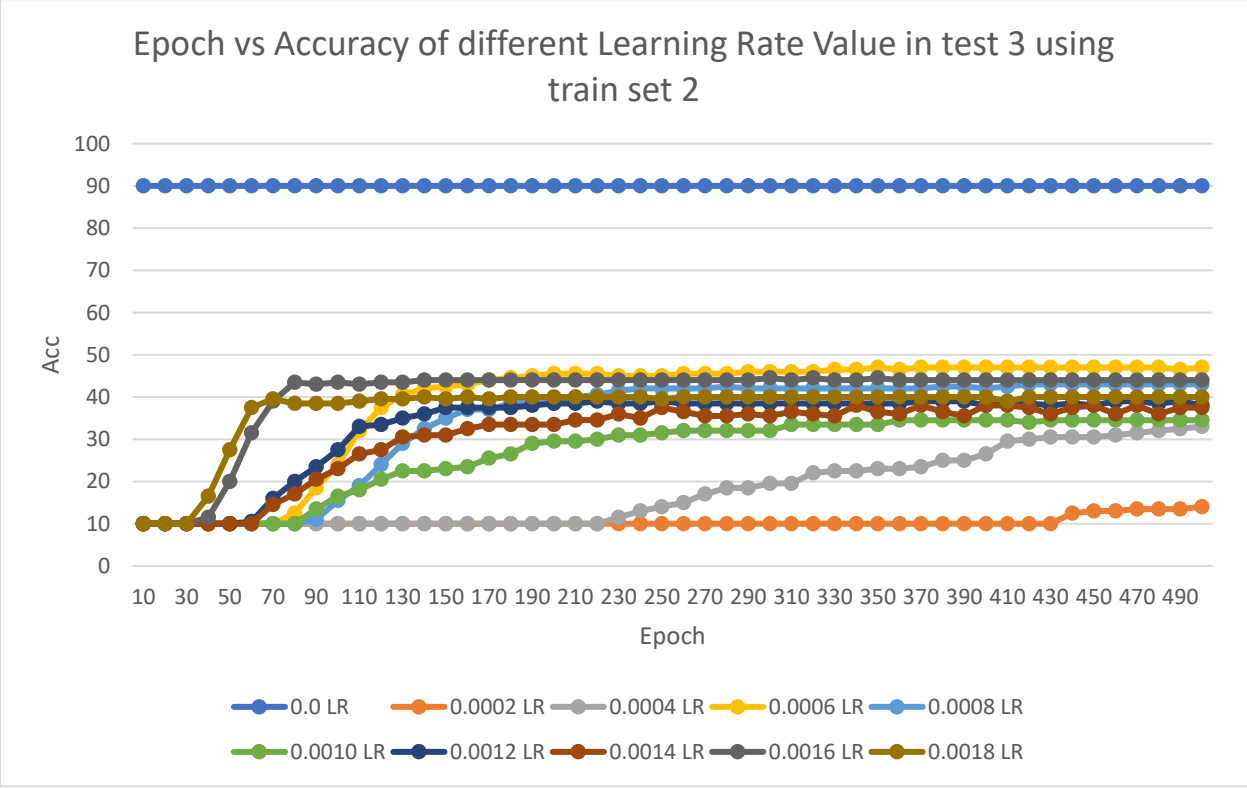

Epoch vs Accuracy of different Learning Rate Value in test 1 using train set 1
Acc
Epoch
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

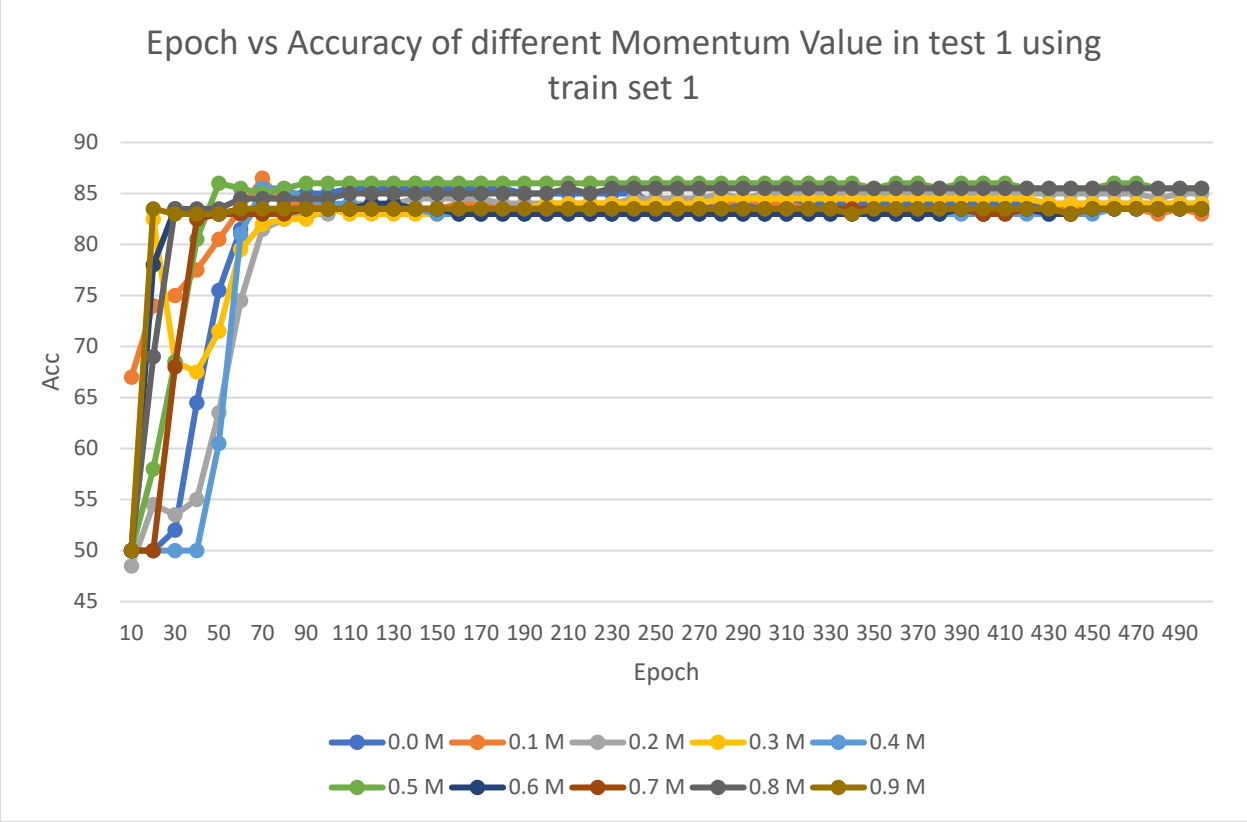

Epoch vs Accuracy of different Learning Rate Value in test 2 using train set 1
Axis Title
Axis Title
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

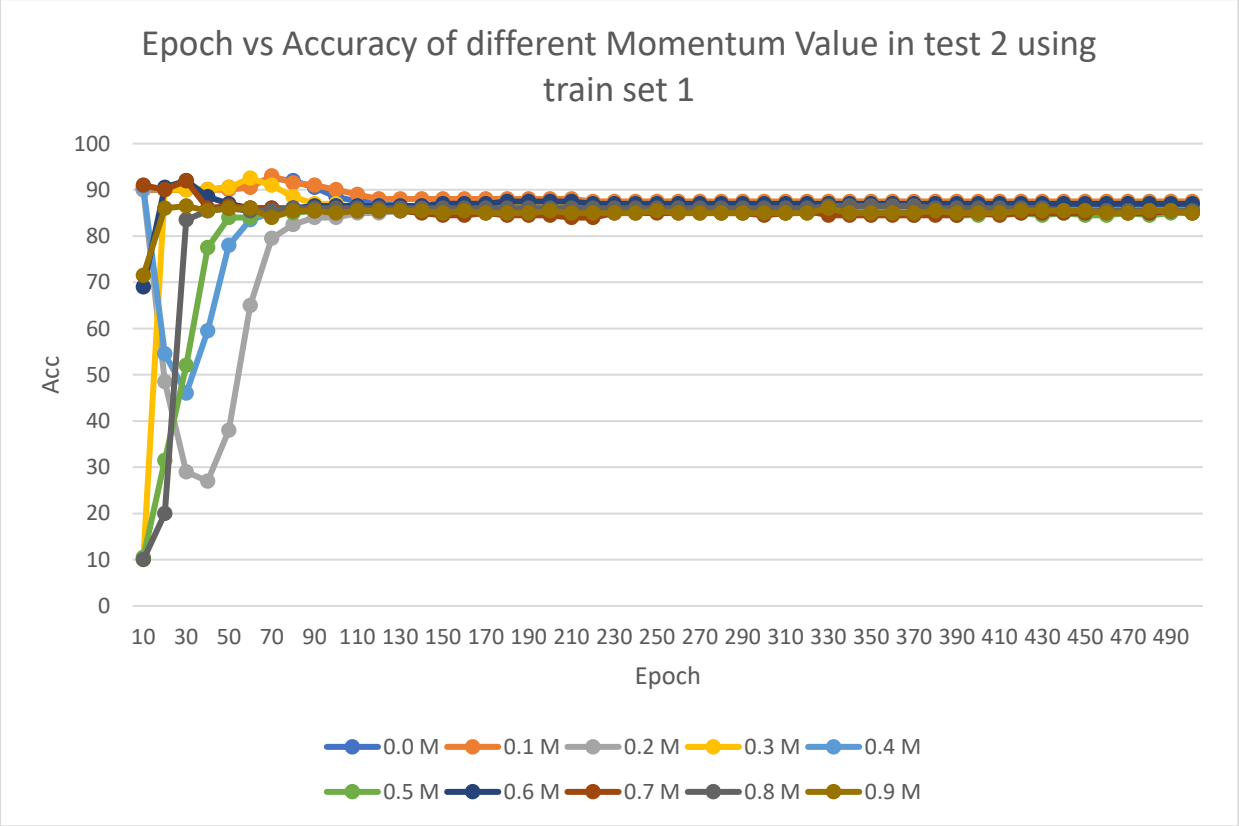

Epoch vs Accuracy of different Learning Rate Value in test 3 using train set 1
Acc
Epoch
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

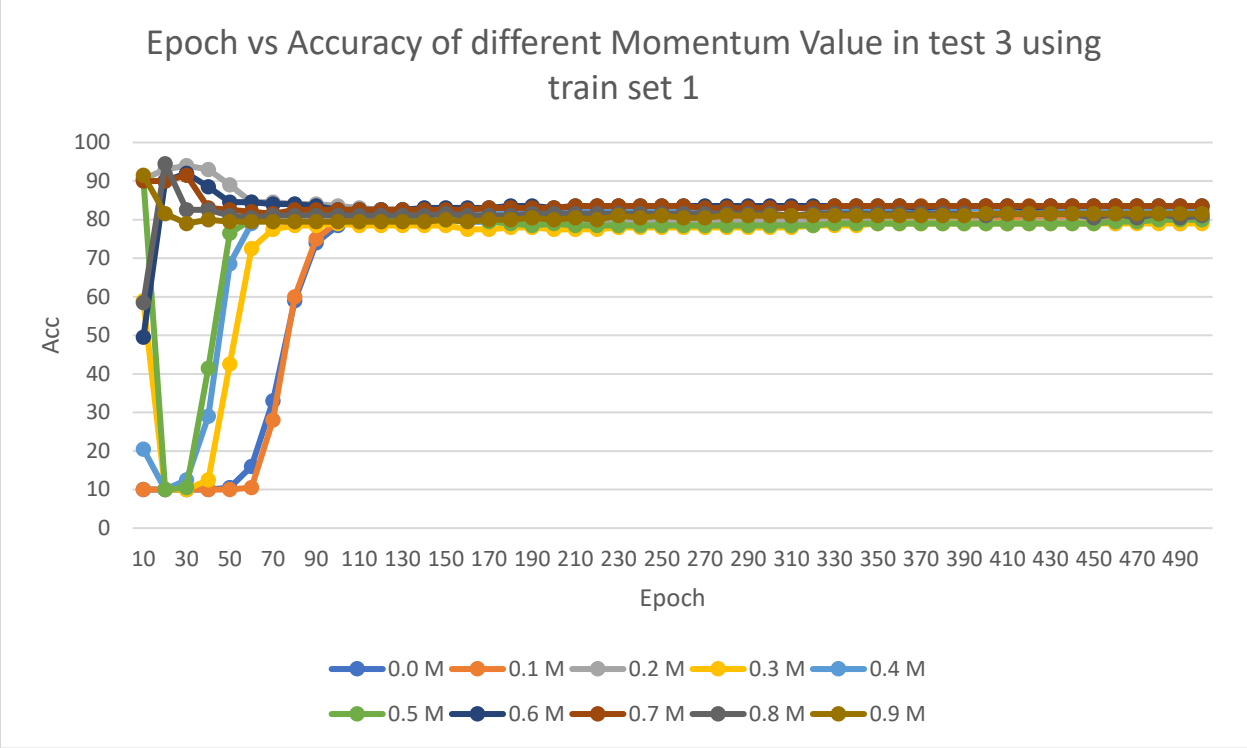

Epoch vs Accuracy of different Learning Rate Value in test 1 using train set 2
Acc
Epoch
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

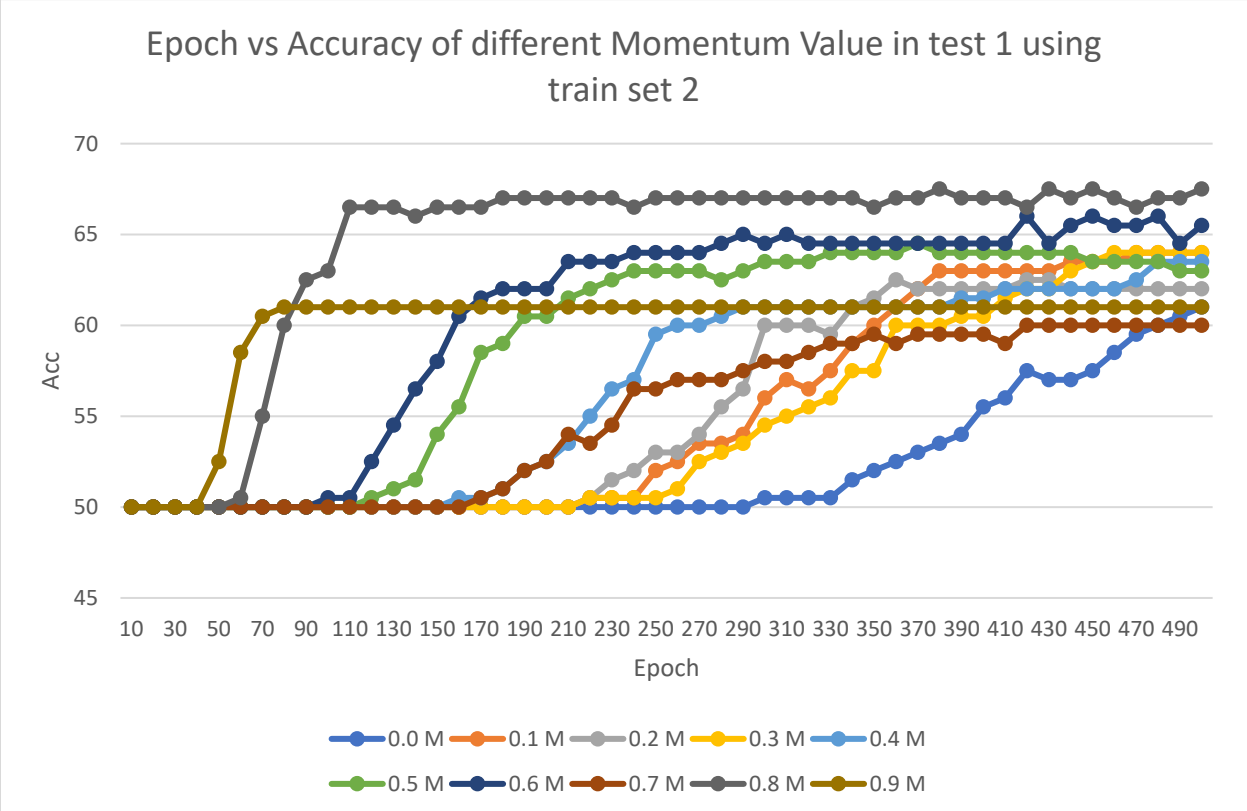

Epoch vs Accuracy of different Learning Rate Value in test 2 using train set 2
Acc
93
92.5
92
91.5
91
90.5
90
89.5
89
88.5
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epocj
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

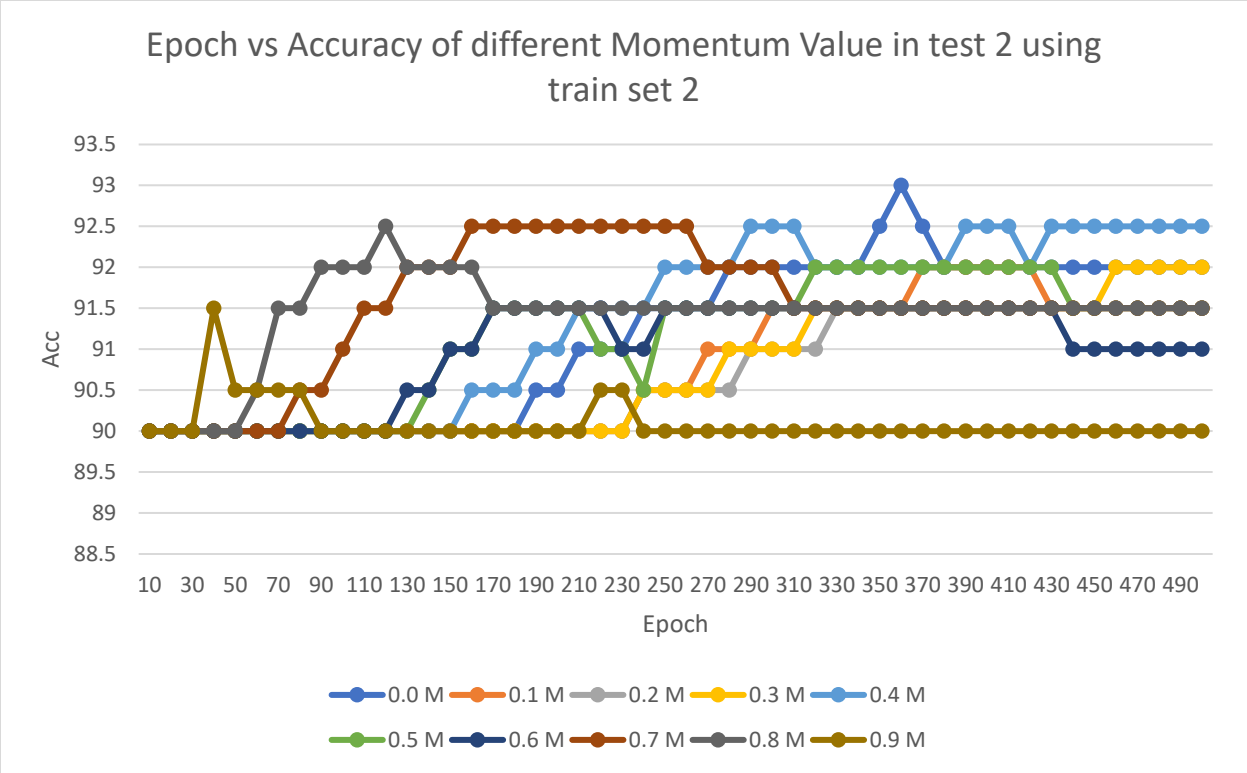

Epoch vs Accuracy of different Learning Rate Value in test 3 using train set 2
Acc
100
90
80
70
60
50
40
30
20
10
0
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0 LR
0.0002 LR
0.0004 LR
0.0006 LR
0.0008 LR
0.0010 LR
0.0012 LR
0.0014 LR
0.0016 LR
0.0018 LR

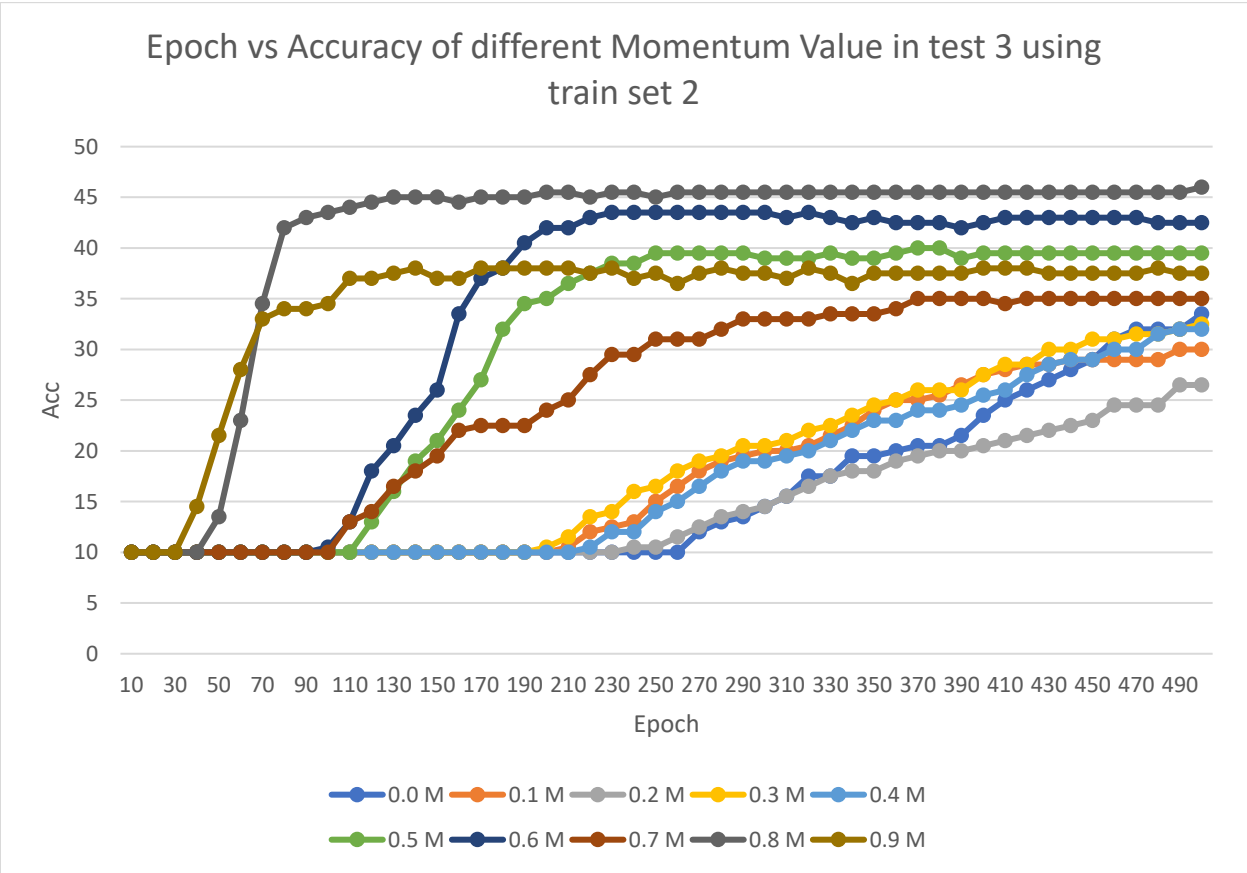

Epoch vs Accuracy of different Momentum Value in test 1 using train set 1
Acc
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

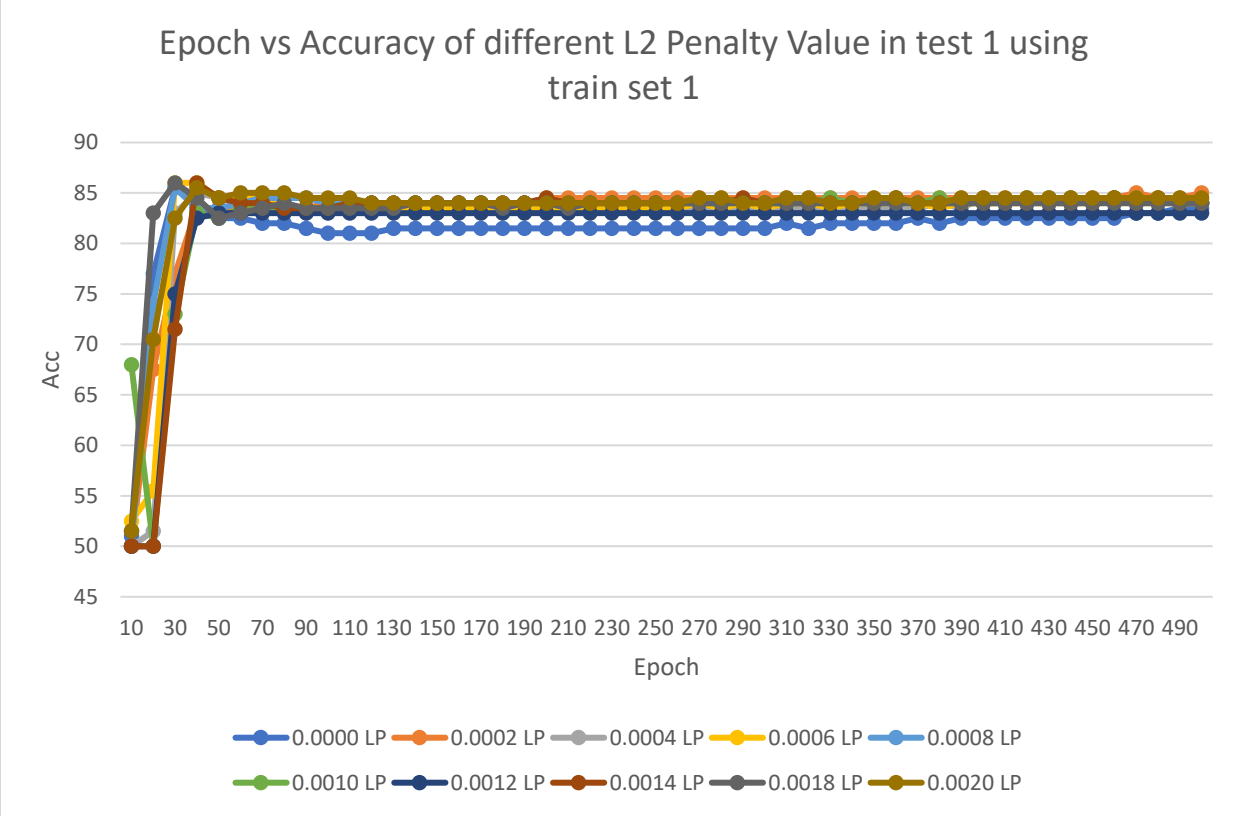

Epoch vs Accuracy of different Momentum Value in test 2 using train set 1
Acc
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

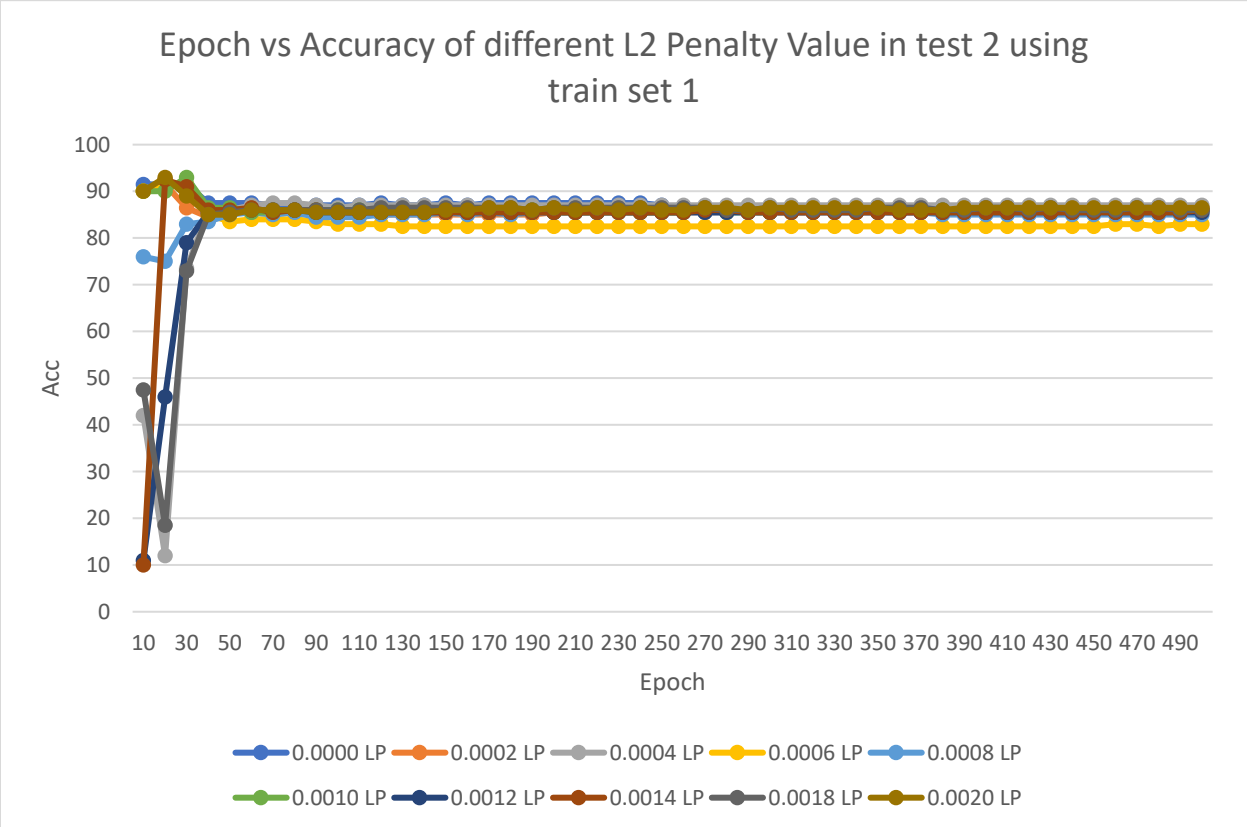

Epoch vs Accuracy of different Momentum Value in test 3 using train set 1
Acc
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

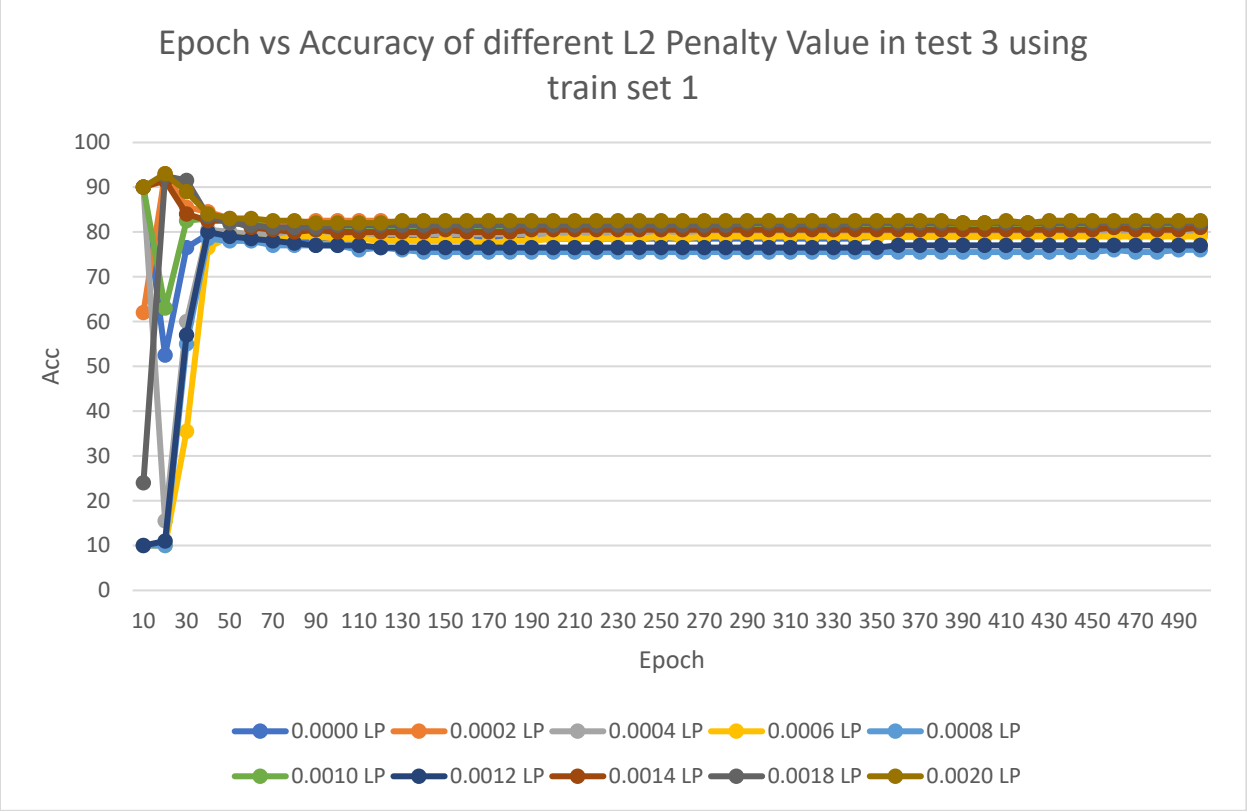

Epoch vs Accuracy of different Momentum Value in test 1 using train set 2
Acc
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

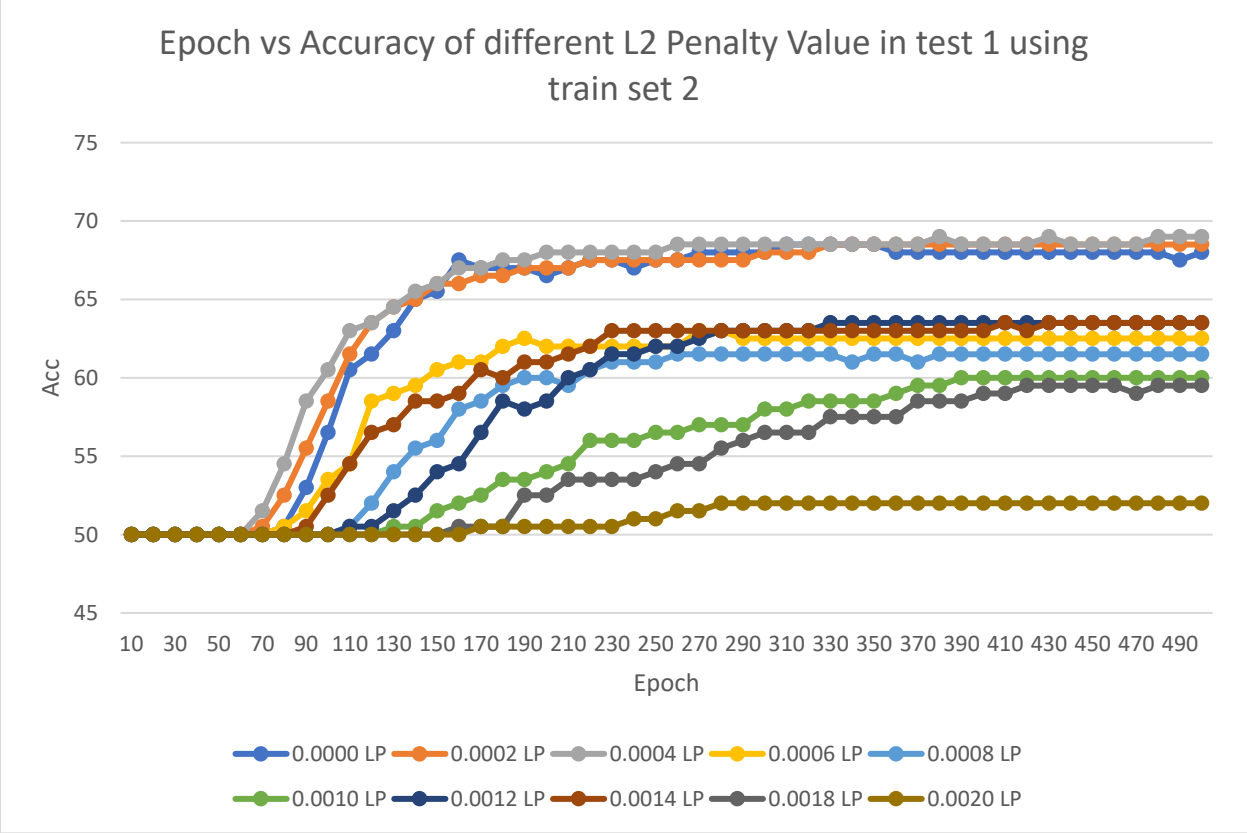

Epoch vs Accuracy of different Momentum Value in test 2 using train set 2
93.5
93
92.5
92
91.5
91
90.5
90
89.5
89
88.5
Acc
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

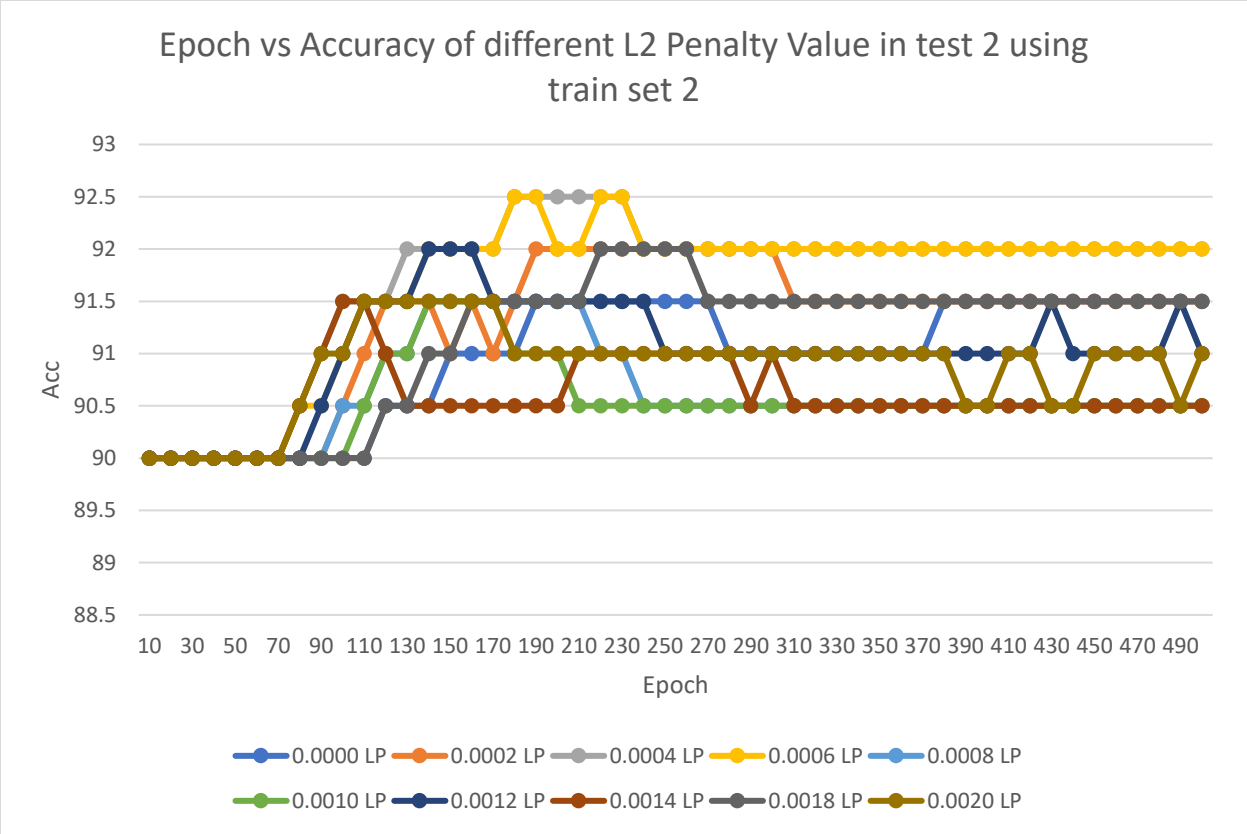

Epoch vs Accuracy of different Momentum Value in test 3 using train set 2
50
45
40
35
30
25
20
15
10
5
0
Acc
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0 M
0.1 M
0.2 M
0.3 M
0.4 M
0.5 M
0.6 M
0.7 M
0.8 M
0.9 M

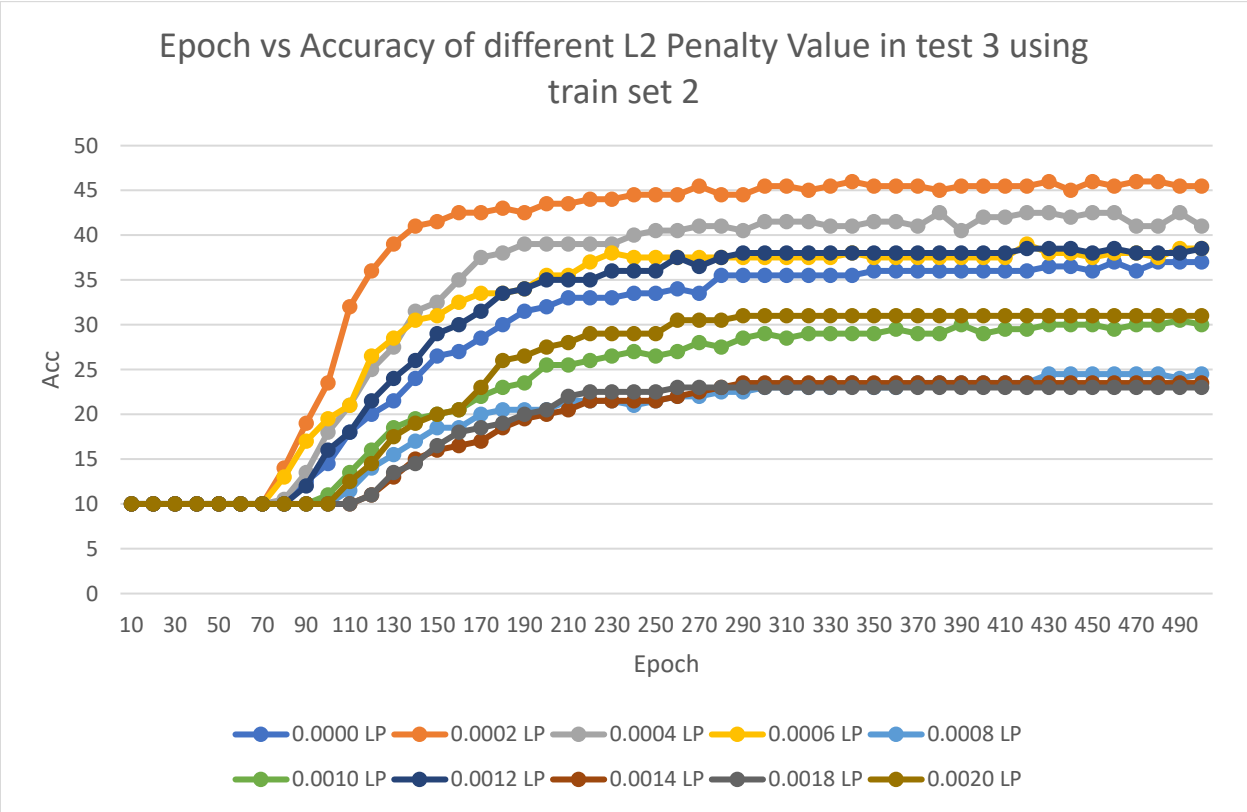

Epoch vs Accuracy of different L2 Penalty Value in test 1 using train set 1
Acc
90
85
80
75
70
65
60
55
50
45
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP


Epoch vs Accuracy of different L2 Penalty Value in test 2 using train set 1
Acc
100
90
80
70
60
50
40
30
20
10
0
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP


Epoch vs Accuracy of different L2 Penalty Value in test 3 using train set 1
Acc
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP


Epoch vs Accuracy of different L2 Penalty Value in test 1 using train set 2
Acc
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP


Epoch vs Accuracy of different L2 Penalty Value in test 2 using train set 2
93
92.5
92
91.5
91
90.5
90
89.5
89
88.5
Acc
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP


Epoch vs Accuracy of different L2 Penalty Value in test 3 using train set 2
50
45
40
35
30
25
20
15
10
5
0
Acc
10 30 50 70 90 110 130 150 170 190 210 230 250 270 290 310 330 350 370 390 410 430 450 470 490
Epoch
0.0000 LP
0.0002 LP
0.0004 LP
0.0006 LP
0.0008 LP
0.0010 LP
0.0012 LP
0.0014 LP
0.0018 LP
0.0020 LP