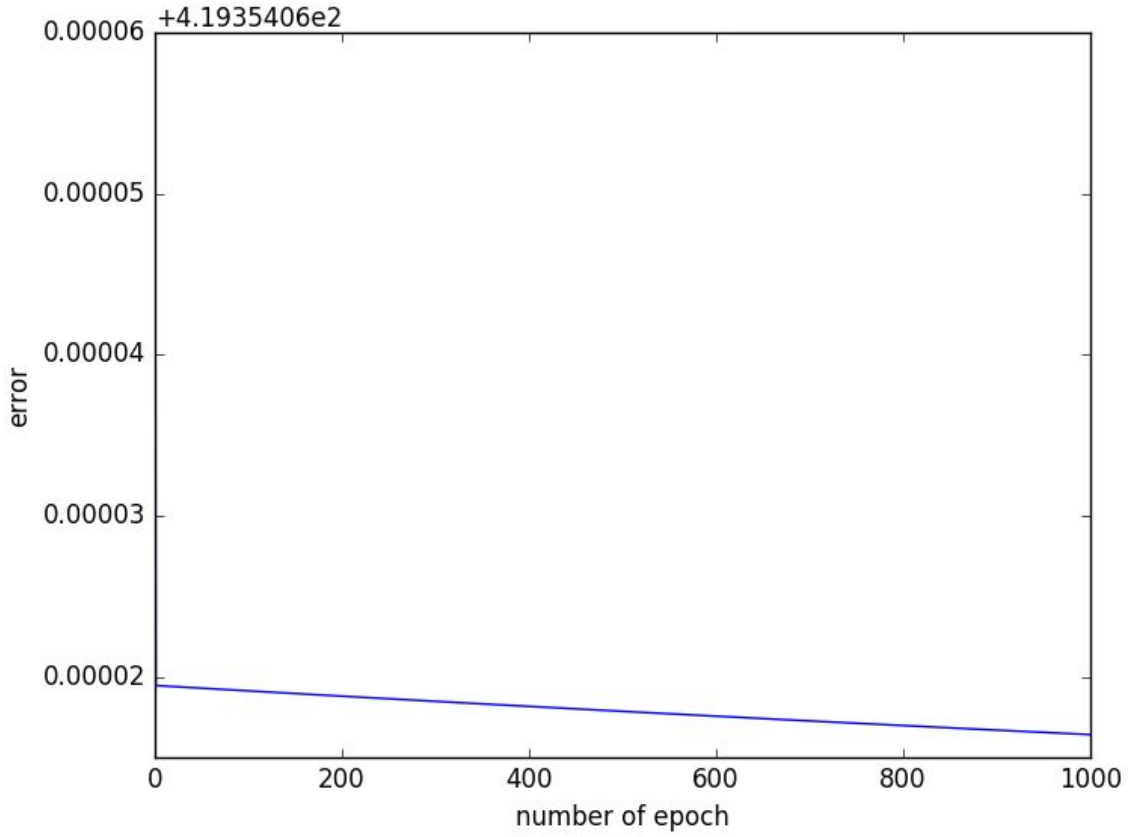
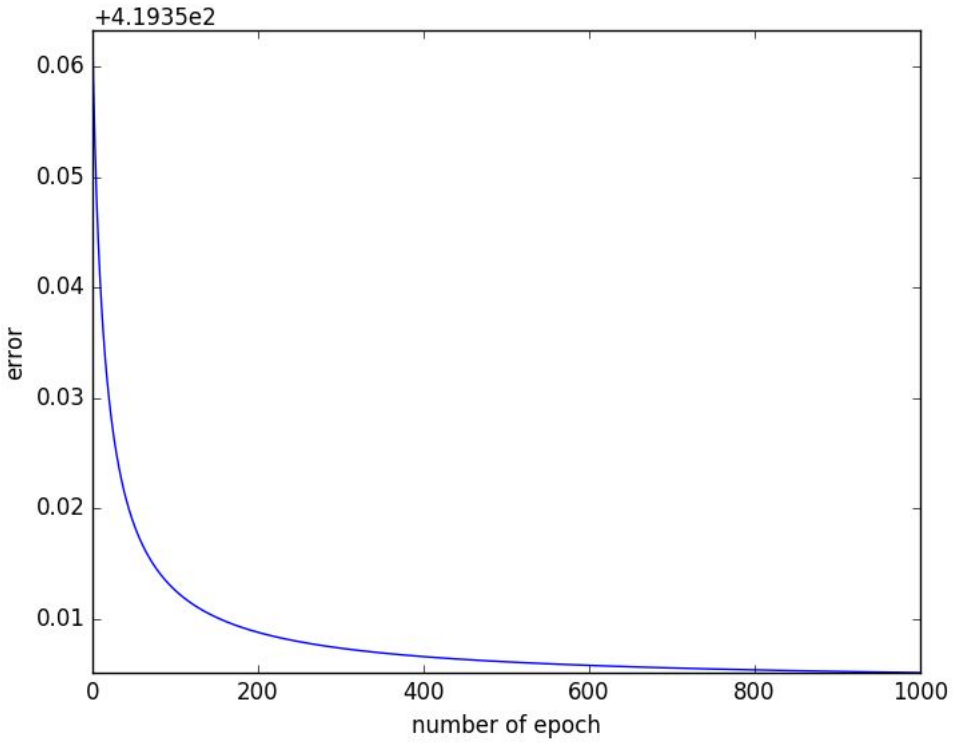
Report - Assignment 2: 160010031+160010011

1) Loss function graphs - Dataset 1 and 2 in order.

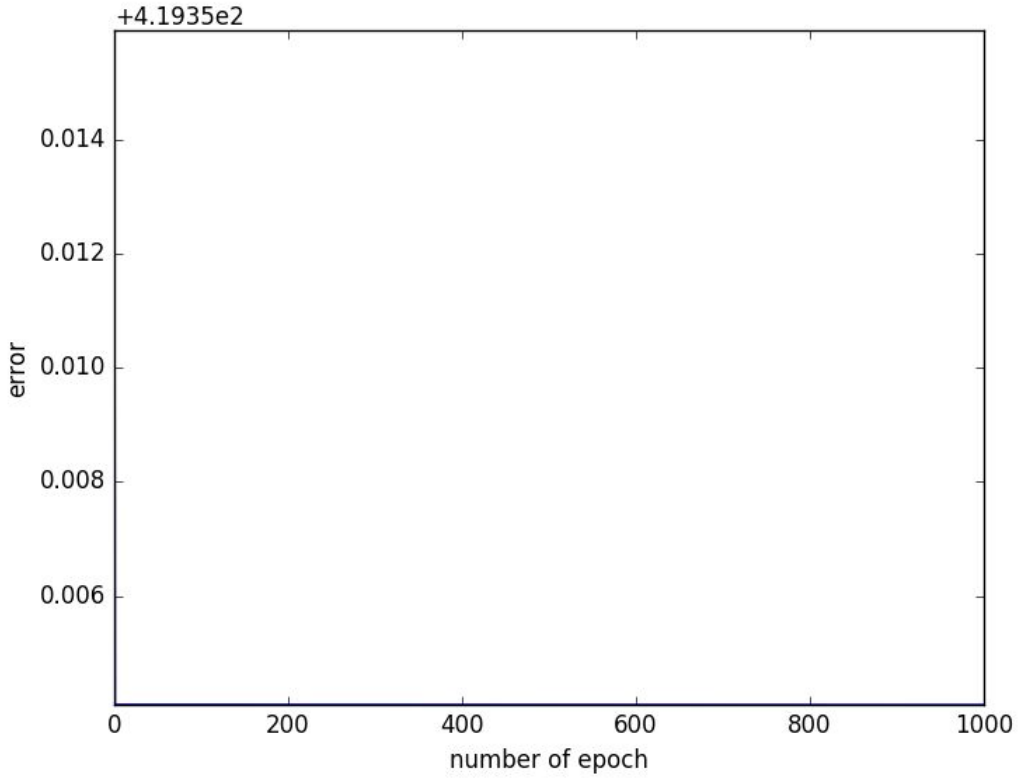
    a) Logistic Loss - No Regularizer (Only for dataset 1)



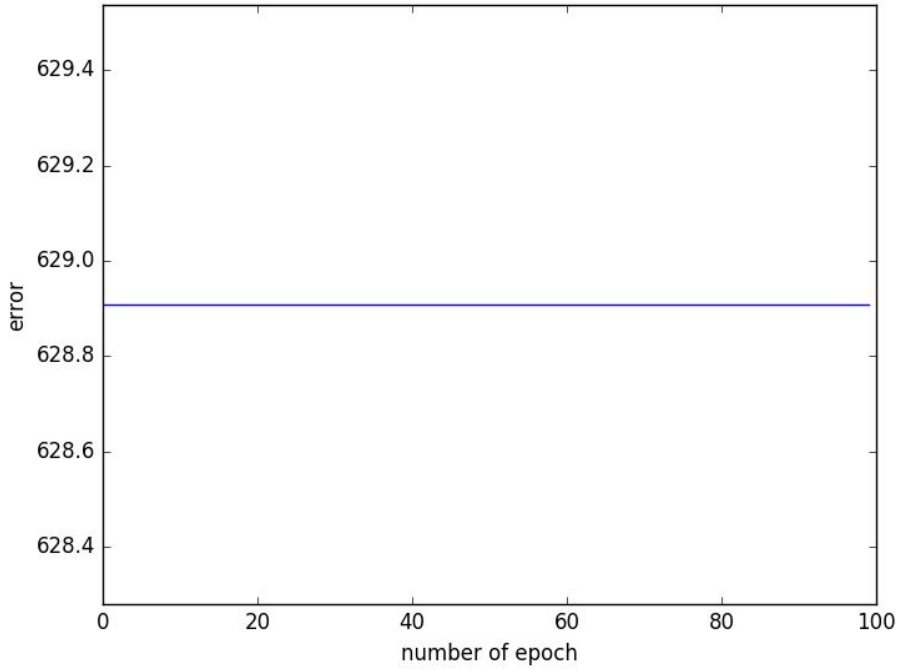
b) Logistic Loss - L2 regularizer (Only for dataset 1)

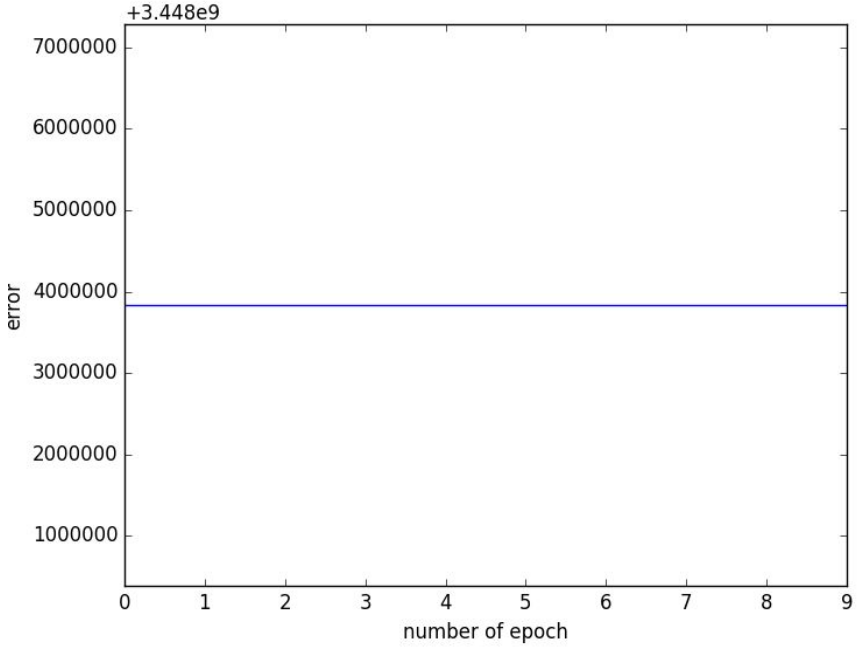


c) Logistic Loss - L4 regularizer (Only for dataset 1)

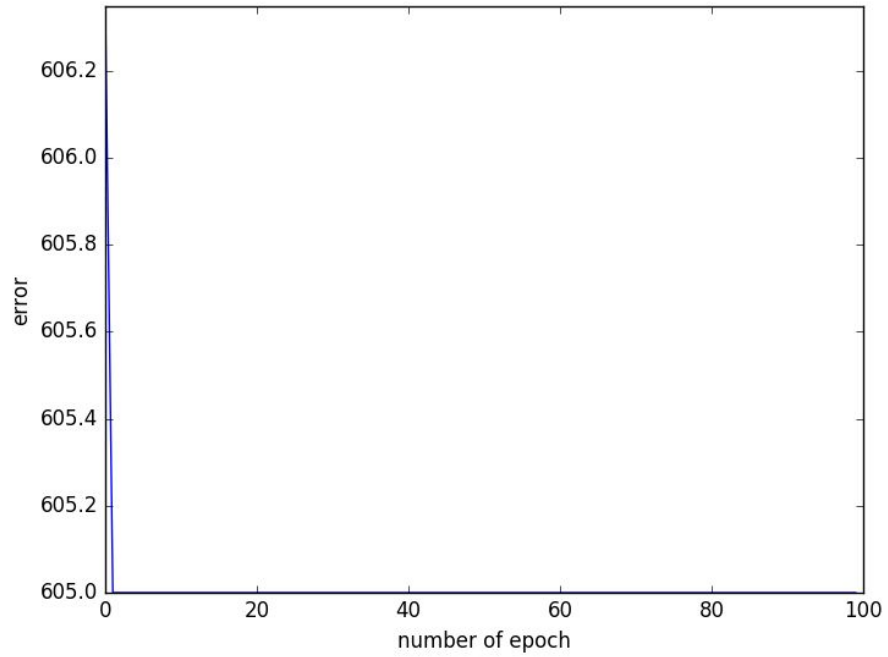


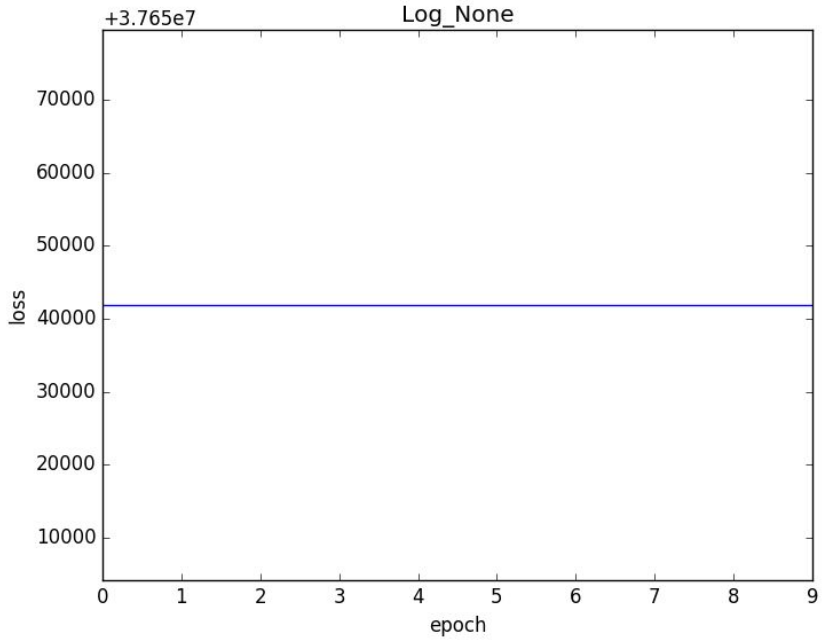
d) Sq. hinge Loss - No regularizer



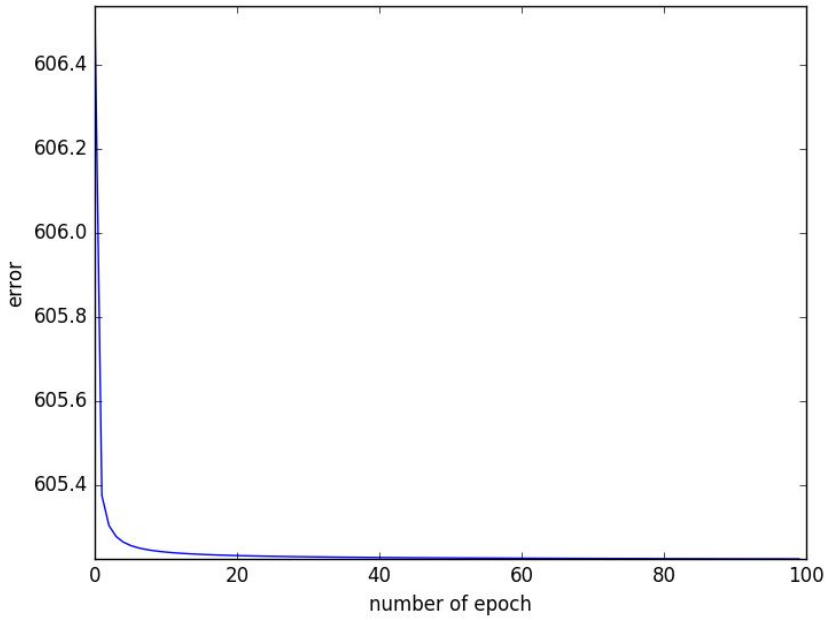


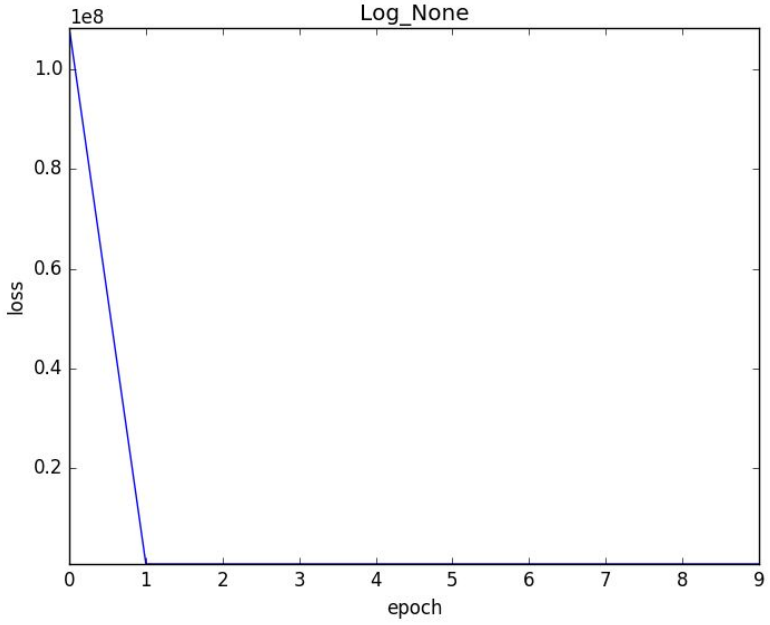
e) Sq. hinge Loss - L2 regularizer



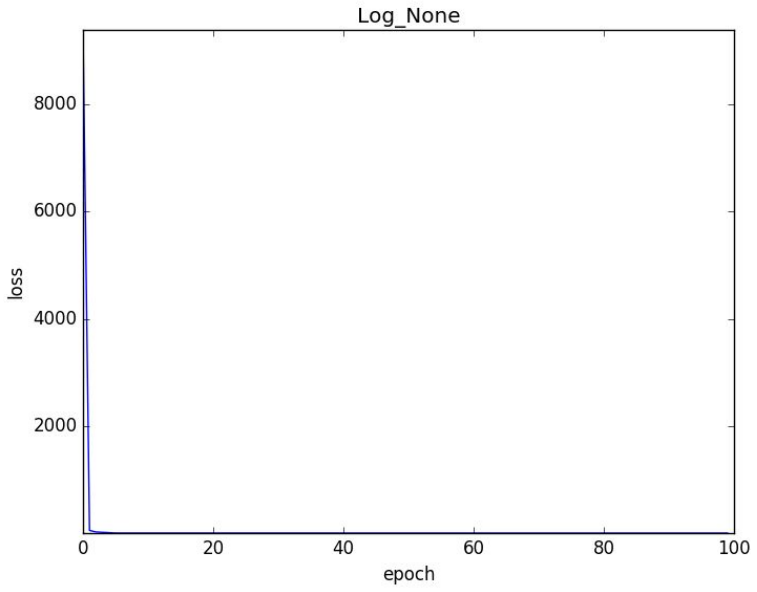


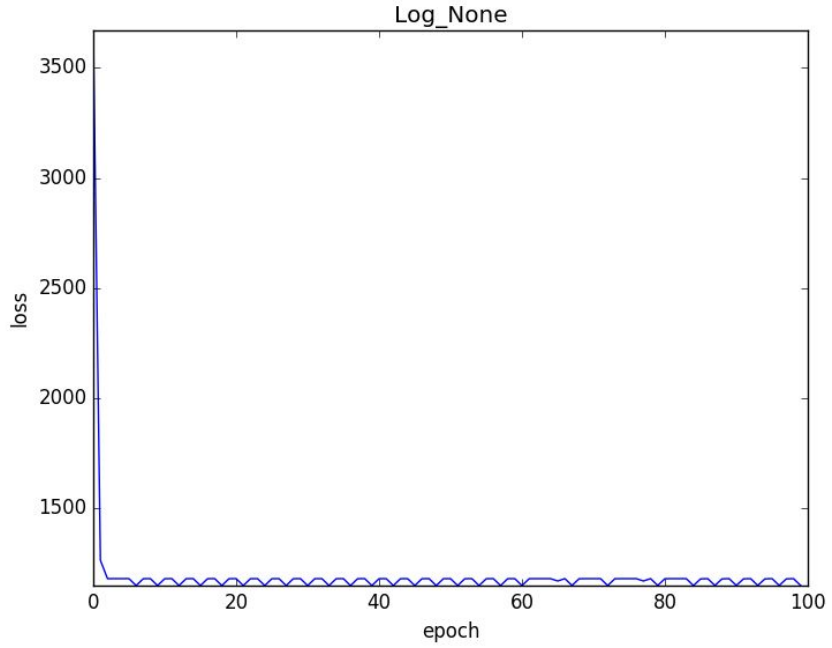
f) Sq. hinge Loss - L4 regularizer



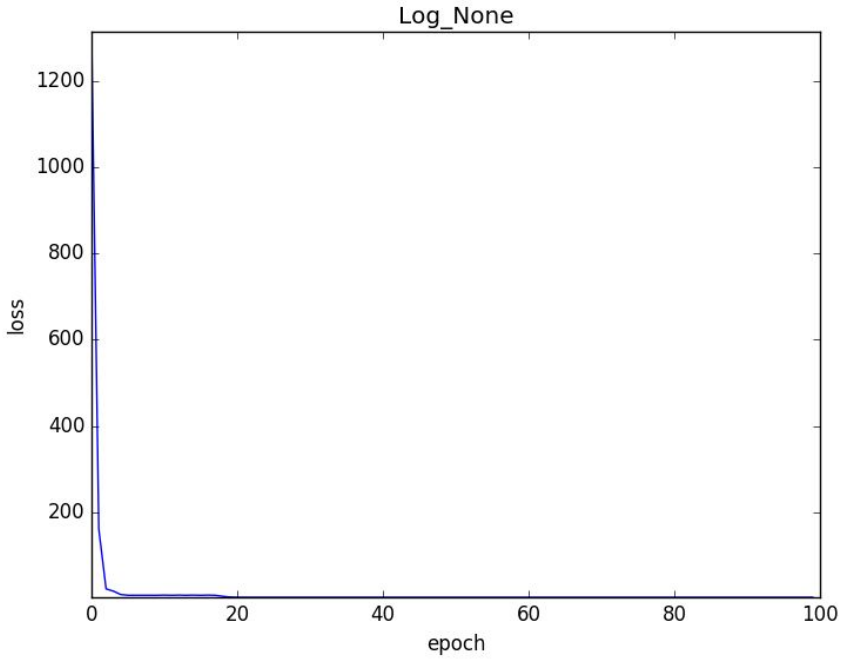


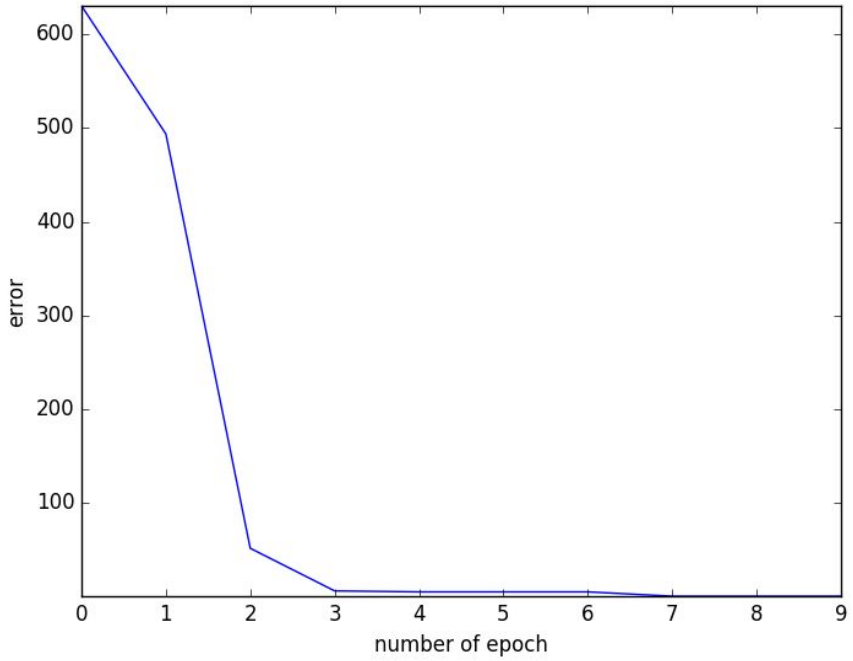
g) Perceptron Loss - No regularizer


Log_None
8000
6000
4000
2000
loss
0
20
40
60
80
100
epoch

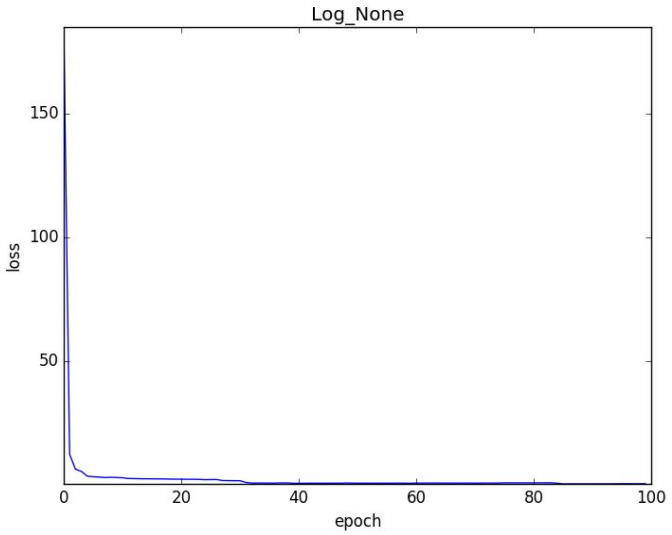

Log_None
3500
3000
2500
2000
1500
loss
0
20
40
60
80
100
epoch

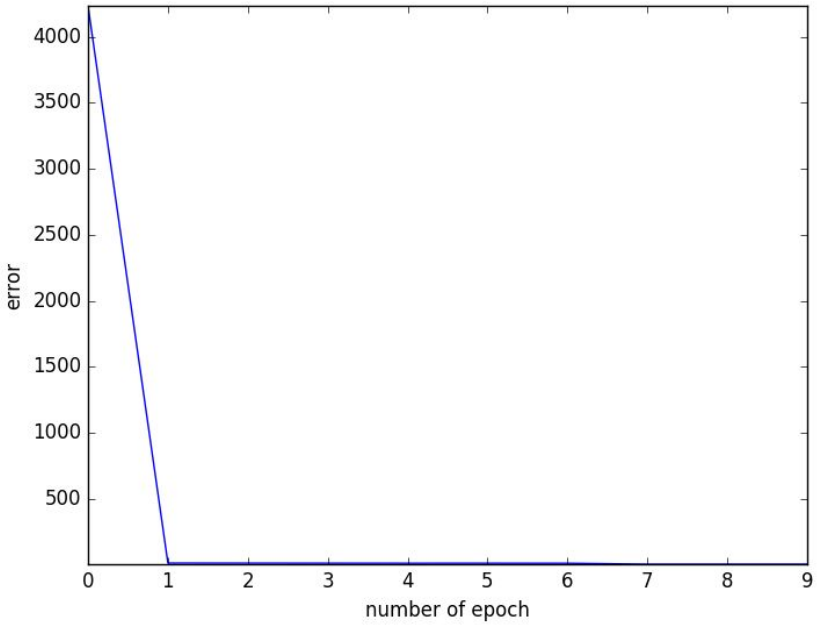
h) Perceptron Loss - L2 regularizer





i) Perceptron Loss - L4 regularizer





2) Observations from the above plot -

a) L+No Regularizer - Linear Plot
b) L+L2 - Decreasing plot wrt epoch^(-2)
c) L+L4 - Decreasing plot wrt epoch^(-4)
d) Sq + None - Constant
e) Sq + L2 - Linear decrease then constant

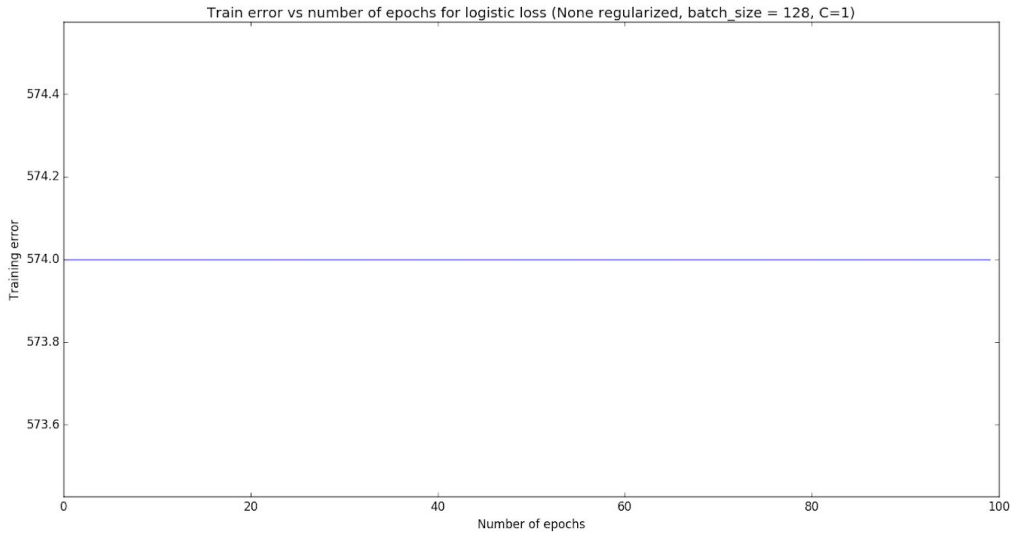
f) Sq + L4 - Exponential Decrease
g) Percp + None - Const
h) Percp + L2 - Linear decrease then constant
i) Percp + L4 - Exponential Decrease - C used = 2

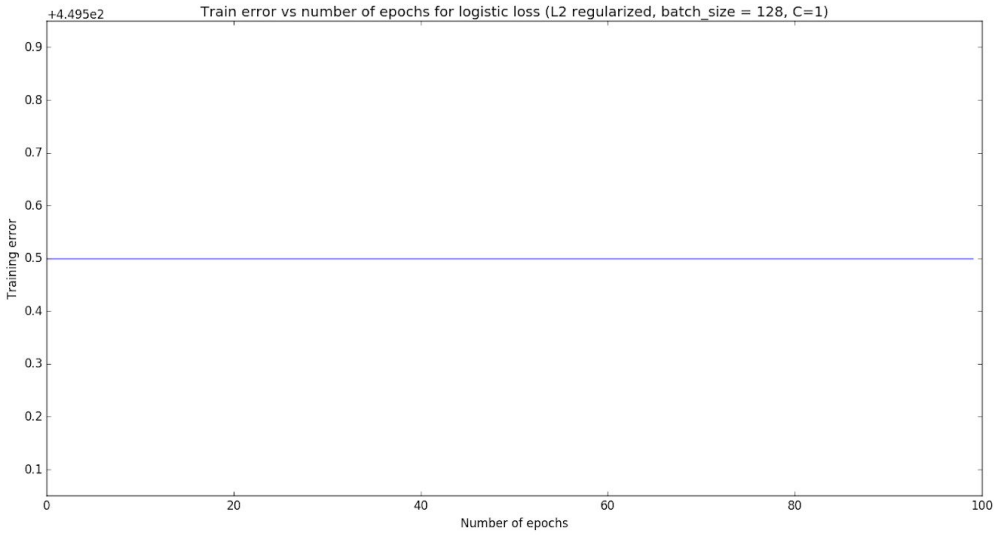
3) Preprocessing -
No data preprocessing.
It was observed that -
a) When C was too low, the loss reduction was fairly slow.
b) When C was large, then the loss would increase and oscillate.(Which is as expected)

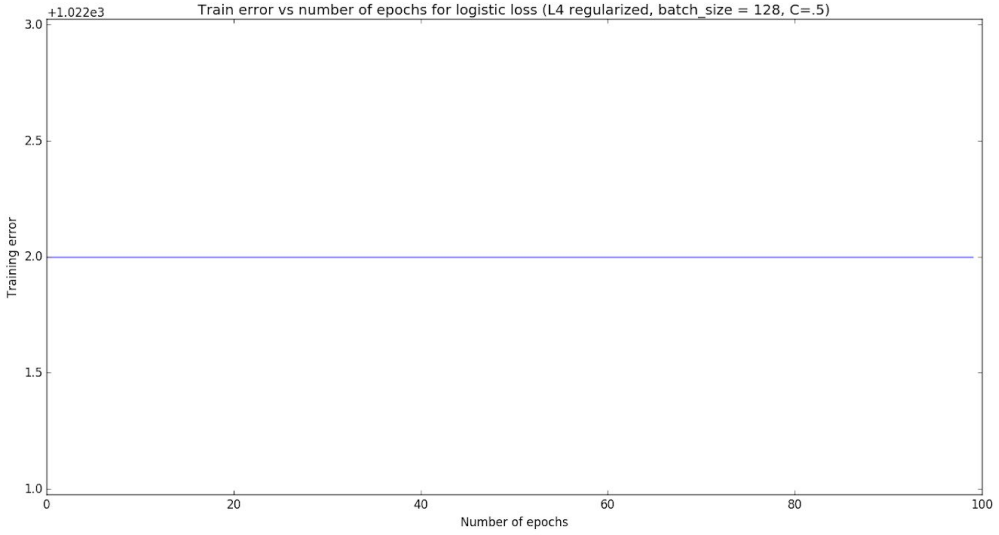
Training error graphs

Logistic Loss -

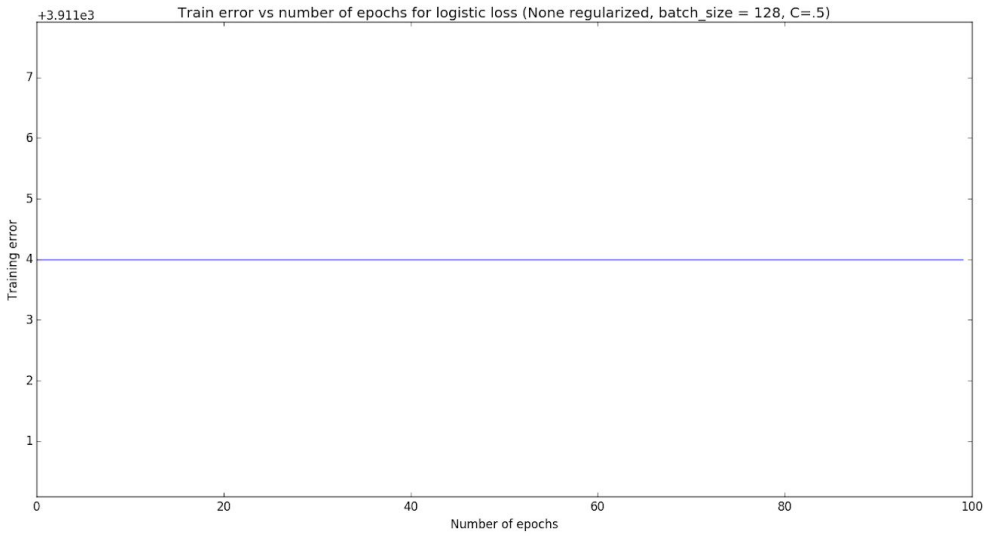
Dataset 1 -

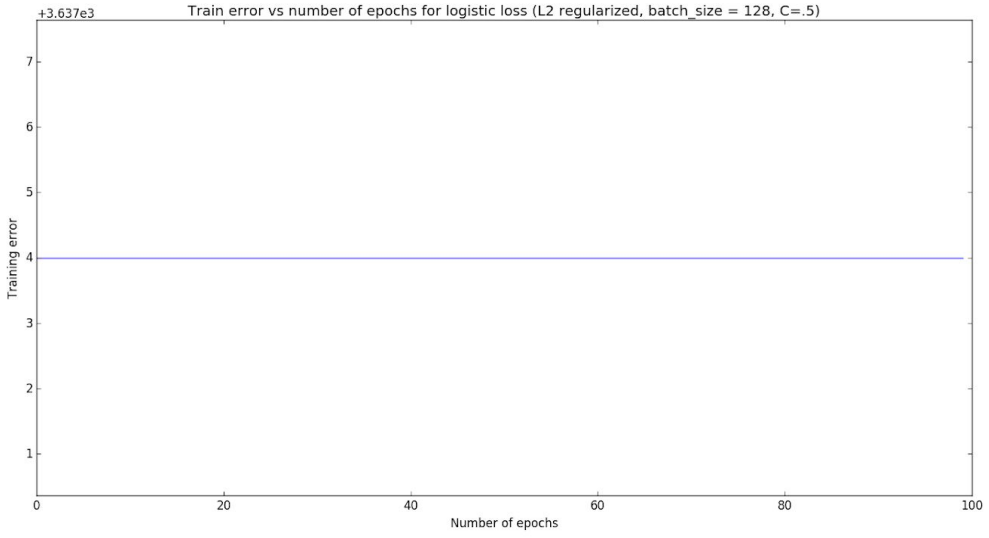


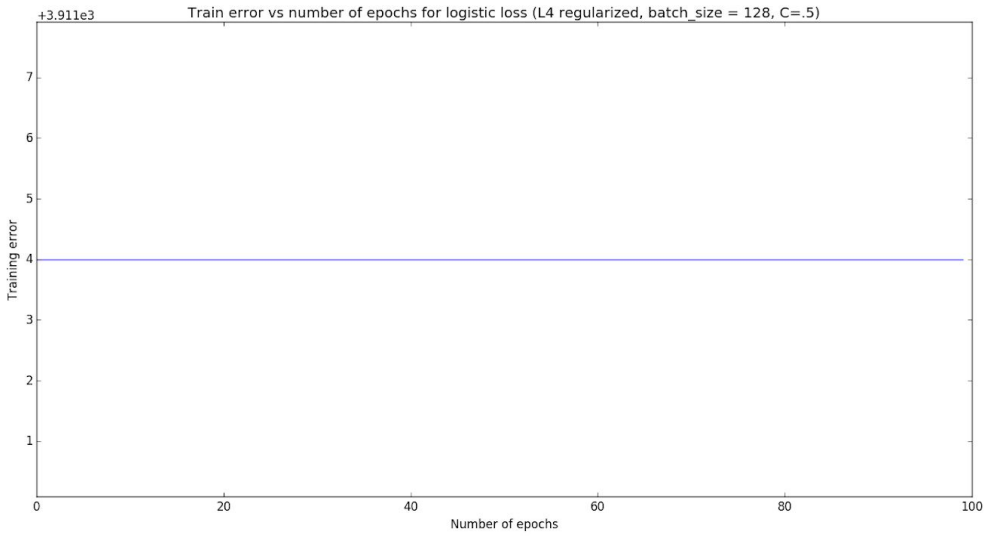

Train error vs number of epochs for logistic loss (L2 regularized, batch_size = 128, C=1)
+4.495e2
Training error
Number of epochs
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
20
40
60
80
100


Train error vs number of epochs for logistic loss (L4 regularized, batch_size = 128, C=.5)
+1.022e3
Training error
Number of epochs
3.0
2.5
2.0
1.5
1.0
0
20
40
60
80
100

Dataset 2 -

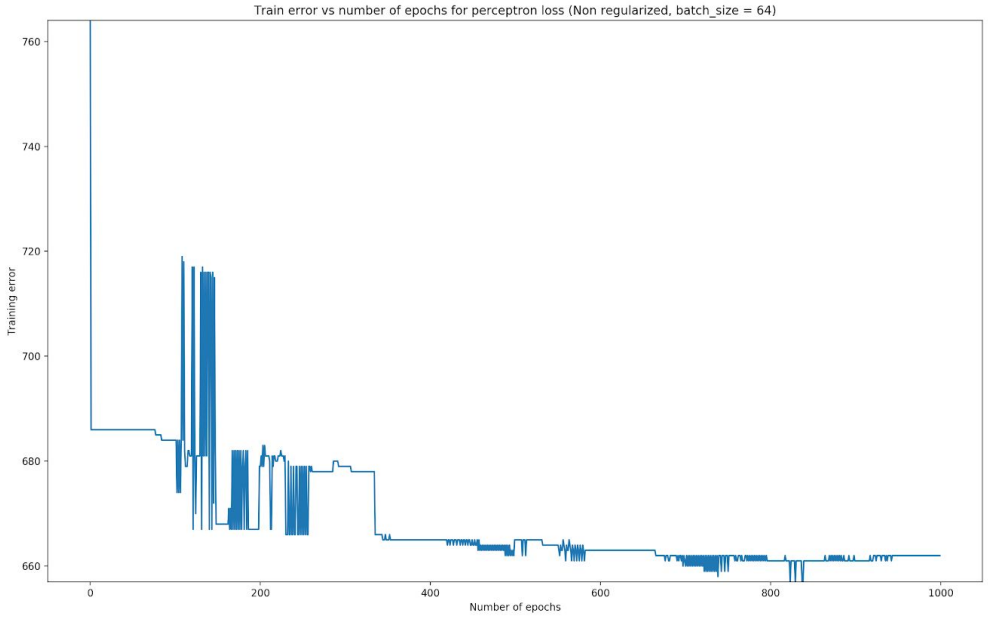

Train error vs number of epochs for logistic loss (None regularized, batch_size = 128, C=.5)
+3.911e3
Training error
Number of epochs
0
20
40
60
80
100
1
2
3
4
5
6
7

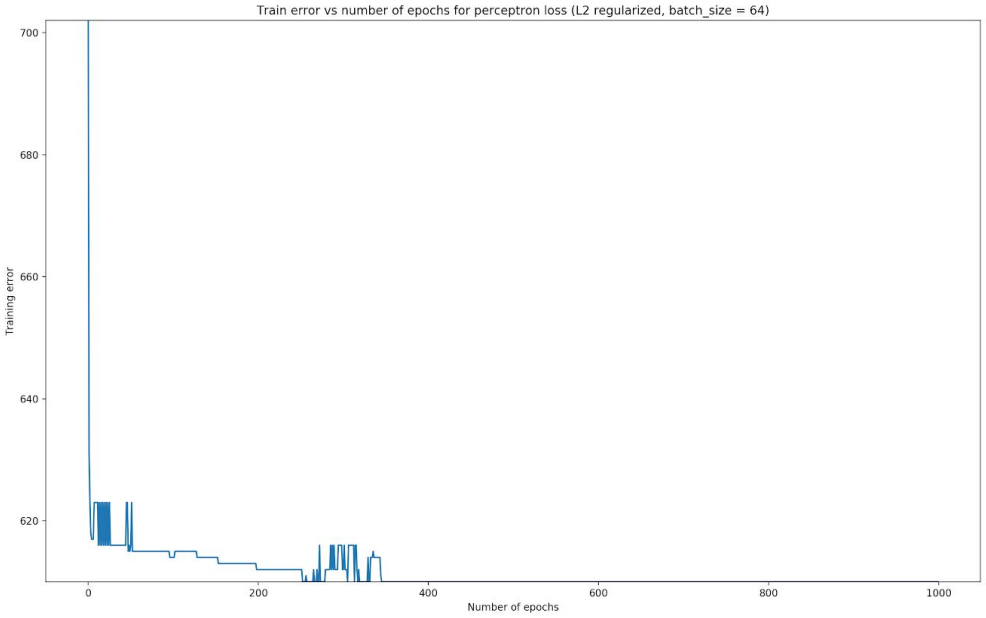

Train error vs number of epochs for logistic loss (L2 regularized, batch_size = 128, C=.5)
+3.637e3
Training error
Number of epochs
0
20
40
60
80
100
1
2
3
4
5
6
7

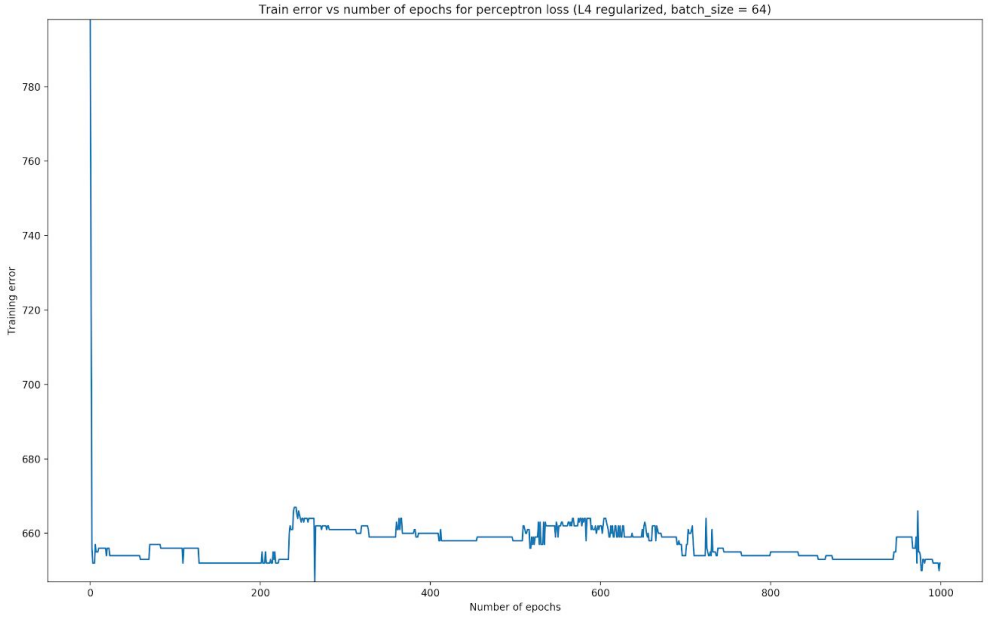


Perceptron loss

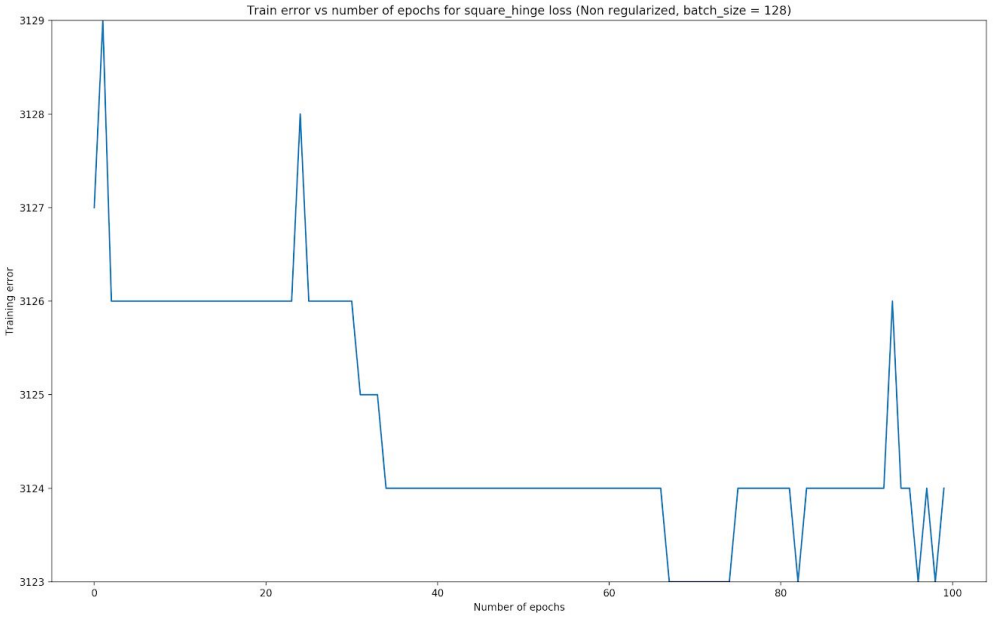
For dataset 1

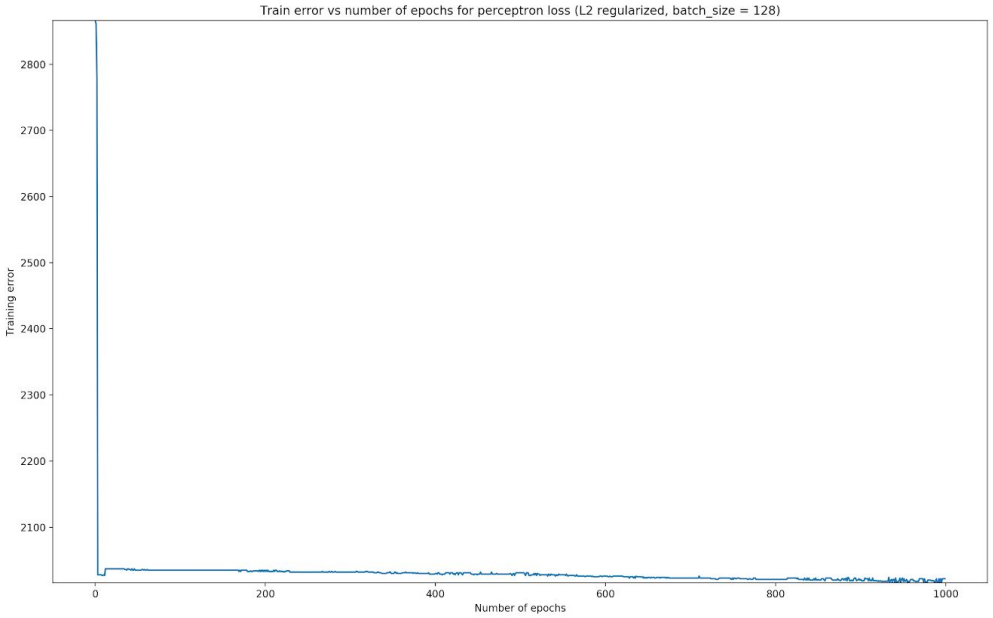


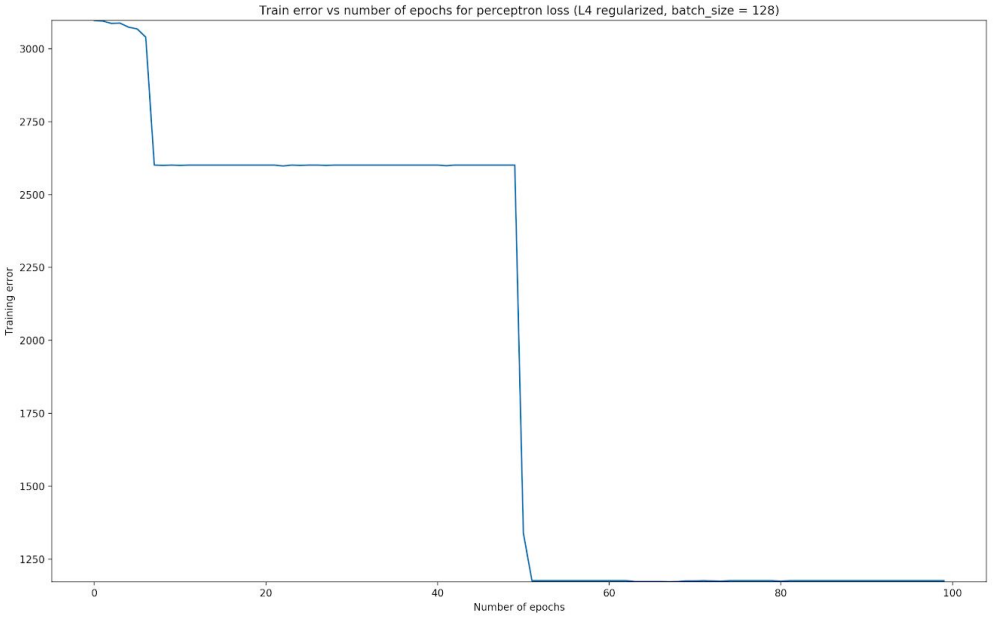

Train error vs number of epochs for perceptron loss (L2 regularized, batch_size = 64)
Training error
Number of epochs
700
680
660
640
620
0
200
400
600
800
1000


Train error vs number of epochs for perceptron loss (L4 regularized, batch_size = 64)
Training error
Number of epochs
780
760
740
720
700
680
660
0
200
400
600
800
1000

Dataset 2

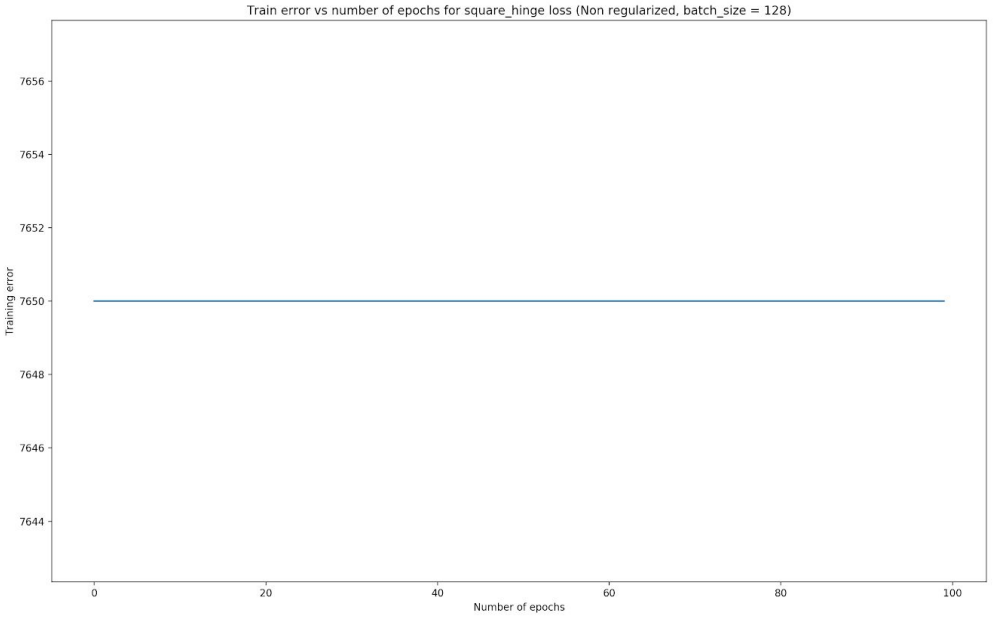

Train error vs number of epochs for square_hinge loss (Non regularized, batch_size = 128)
Training error
Number of epochs
3129
3128
3127
3126
3125
3124
3123
0
20
40
60
80
100


Train error vs number of epochs for perceptron loss (L2 regularized, batch_size = 128)
Training error
Number of epochs
2800
2700
2600
2500
2400
2300
2200
2100
0
200
400
600
800
1000

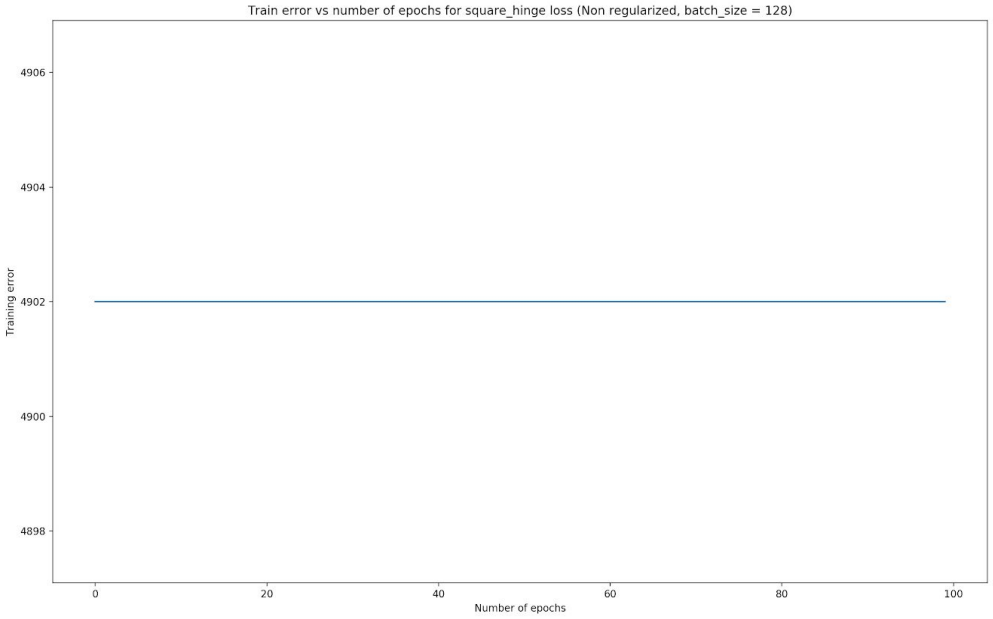


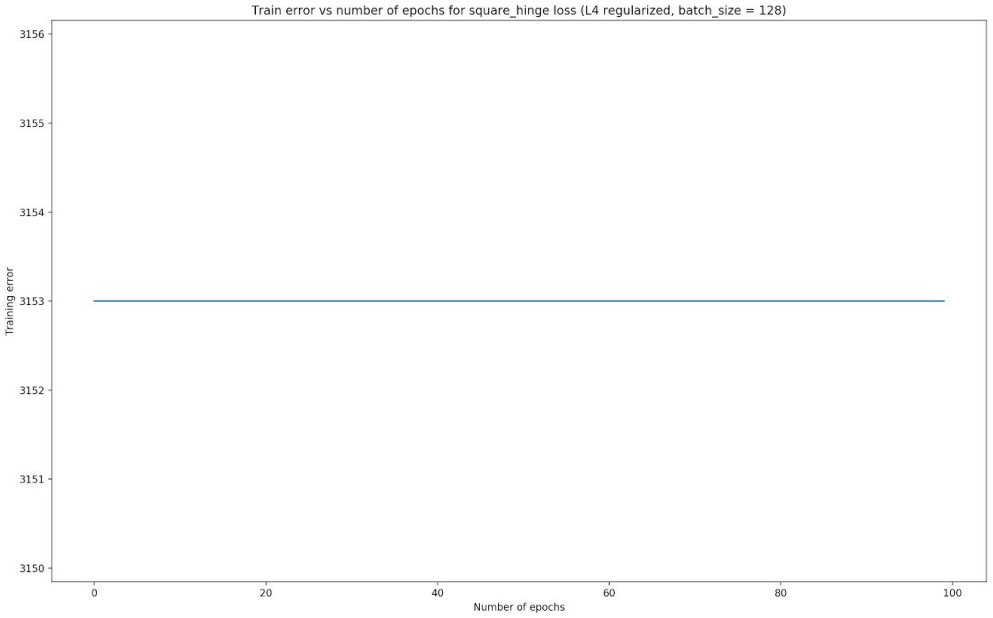
Square Hinge Loss

Dataset 2

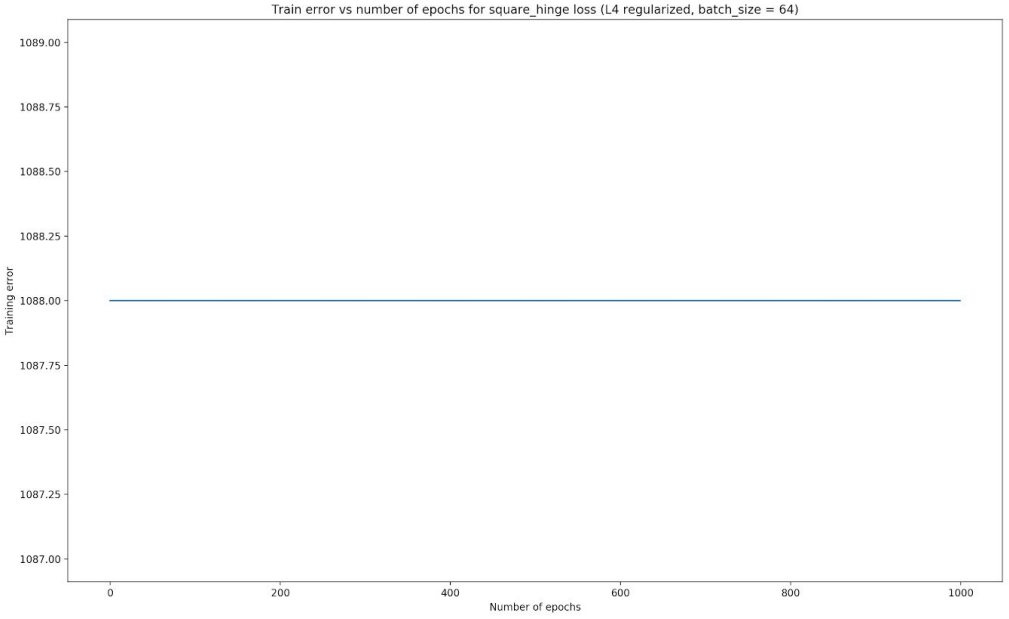


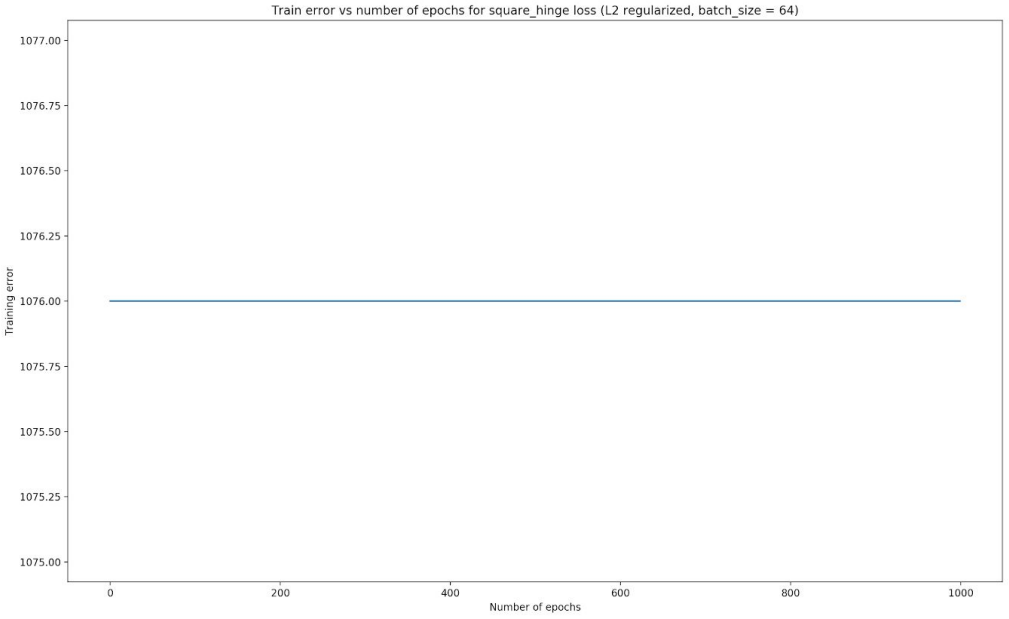
Below graph is for L2 regularized (Graph Title is incorrect)

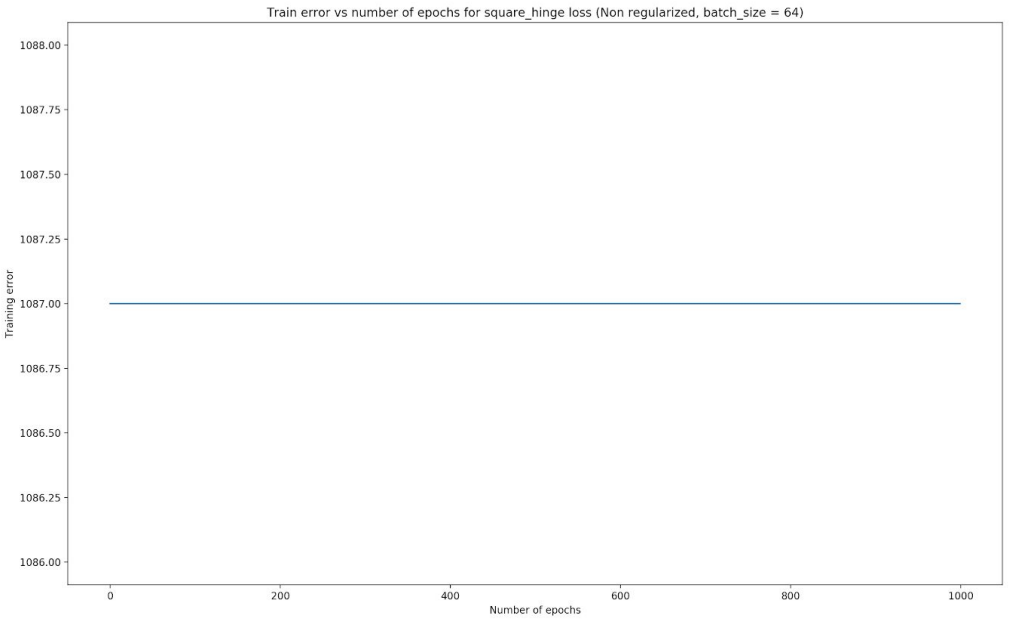



Train error vs number of epochs for square_hinge loss (L4 regularized, batch_size = 128)
Training error
Number of epochs
3156
3155
3154
3153
3152
3151
3150
0
20
40
60
80
100

Dataset 1


Train error vs number of epochs for square_hinge loss (L4 regularized, batch_size = 64)
Training error
Number of epochs
1089.00
1088.75
1088.50
1088.25
1088.00
1087.75
1087.50
1087.25
1087.00
0
200
400
600
800
1000


Train error vs number of epochs for square_hinge loss (L2 regularized, batch_size = 64)
Training error
Number of epochs
1077.00
1076.75
1076.50
1076.25
1076.00
1075.75
1075.50
1075.25
1075.00
0
200
400
600
800
1000


Train error vs number of epochs for square_hinge loss (Non regularized, batch_size = 64)
Training error
Number of epochs
1088.00
1087.75
1087.50
1087.25
1087.00
1086.75
1086.50
1086.25
1086.00
0
200
400
600
800
1000

Observations from training error vs number of epochs

The changes in error due to square hinge loss and logistic loss is minimal due to very low gradient

Perceptron loss shows expected behaviour and best Kaggle score was obtained with perceptron loss.

Parameters used for best kaggle score:

Dataset 1

Batch_size = 128
C = 2
Regularizer = L2 with lambda = 6
Perceptron loss

Dataset 2

Batch_szie = 512
C = 2
Regularizer = L4 with lambda = 5
Perceptron Loss