**Machine Learning**

Assignment - 2

1. **Linear regression:**

Approach:

We have tried to achieve a hyper-plane that is closest to all the points in the space. The final plane is such that sum of distances of all points is the minimum.

* Of the given 308 elements in dataset, 200 random points are taken as training set and rest as testing set.
* First we train the model to find the hyperplane that fits the selected training data.
* After initializing the weight matrix, we perform gradient descent to reach at the optimum weight vector
  + In this, the loss function and gradient is computed in each iteration.
  + Then we alter the plane in each iteration slightly so that the loss reduces.
  + We stop when gradient becomes close to 0. (Practically less than 10^-4)
* Once done, we visualize the error function against total number of iterations.

Following are the results:

1.

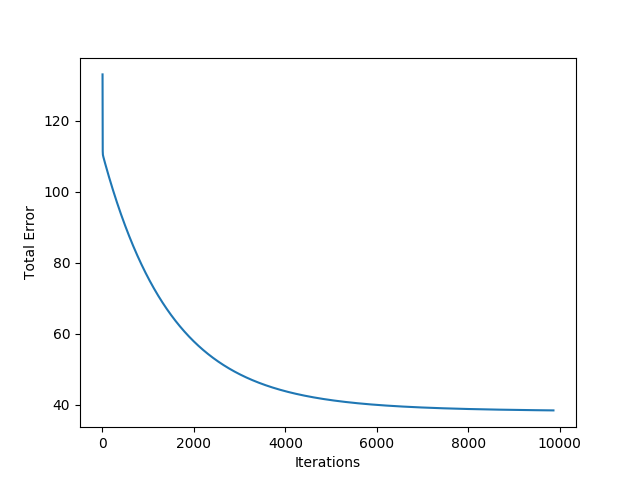
stopped at : 10^-2

training error: 38.40

testing error: 43.40

iterations: 9866

[w0,w1,…,w6] = [-0.547, -0.116, 5.591, -13.386, 3.400, 8.375, 109.049]



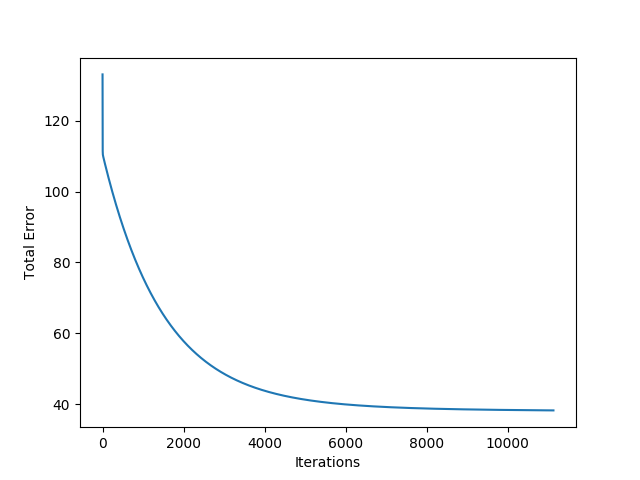
2.

Stopped at 10^-3

training error: 38.267

testing error: 43.06

iterations: 11130



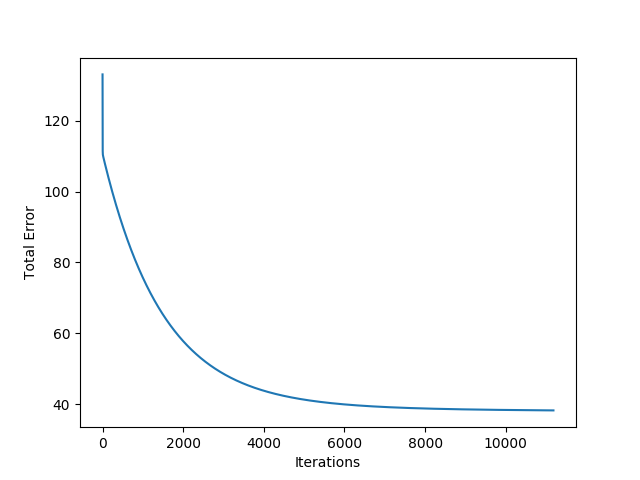
3.

stopped at 10^-4

training error: 38.26

testing error: 43.04

iterations: 11188



1. **Perceptron:**

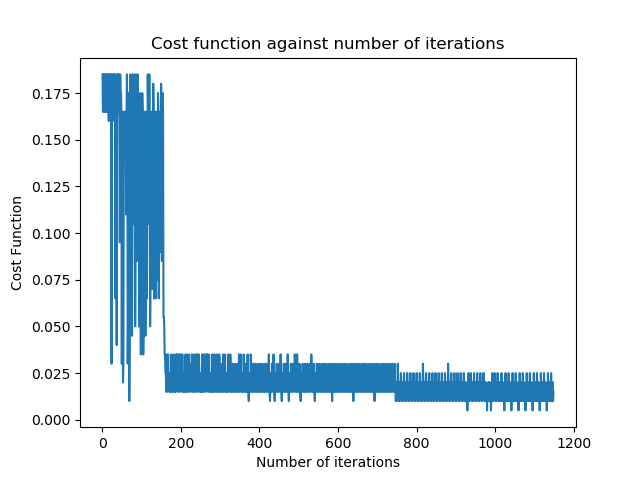
Approach :

We have tried to achieve a hyper-plane that best classifies the given data. To achieve the parameters for this hyper-plane, we did the following:

* We selected 70 random examples as training set and remaining 30 examples as test set
* After initializing the weight vector, we performed gradient descent to reach at the optimum weight vector
  + We continue calculating the gradient and updating the weight vector until the gradient becomes less than a specified theta(or threshold) value.
  + We have also plotted the mean squared error against the number of iterations to visualize the convergence of gradient descent.

With the final weight vector, we calculate the training and test set error by comparing the predicted and actual labels of data points.

Following are some common outputs:



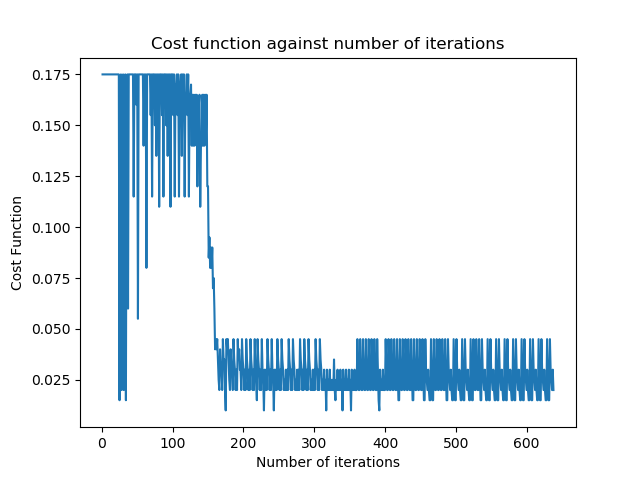
Training set error: 2.857142857142857 %

Training set accuracy: 97.14285714285714 %

Calculated weights: [-0.04590000000000018, -0.04993999999999965, -0.04525000000000039, 0.07098999999999986, 0.08341999999999856]

Test set error: 3.3333333333333335 %

Test set accuracy: 96.66666666666667 %



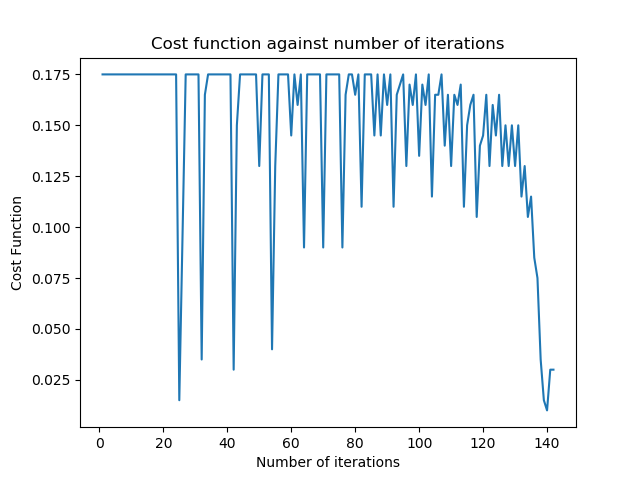
Training set error: 2.857142857142857 %

Training set accuracy: 97.14285714285714 %

Calculated weights: [-0.05260000000000005, -0.06747000000000068, -0.06030000000000073, 0.09769000000000058, 0.09543999999999919]

Test set error: 0.0 %

Test set accuracy: 100.0 %



Training set error: 0.0 %

Training set accuracy: 100.0 %

Calculated weights: [-0.03560000000000005, -0.07616000000000012, -0.05695000000000006, 0.10731999999999983, 0.08525999999999999]

Test set error: 6.666666666666667 %

Test set accuracy: 93.33333333333333 %