SMAI

Assignment_1 Report

Name : **Akhilesh Soni**

Roll No : **201530001**

Q1)

For Breast Cancer Dataset:

Vanilla Perceptron:

for 10 Epochs accuracy is 94.853%

for 15 Epochs accuracy is 95.000%

for 20 Epochs accuracy is 96.176%

for 25 Epochs accuracy is 94.853%

for 30 Epochs accuracy is 94.118%

for 35 Epochs accuracy is 96.912%

for 40 Epochs accuracy is 95.294%

for 45 Epochs accuracy is 96.765%

for 50 Epochs accuracy is 93.676%

Voted perceptron:

for 10 Epochs accuracy is 96.618%

for 15 Epochs accuracy is 96.471%

for 20 Epochs accuracy is 96.912%

for 25 Epochs accuracy is 97.059%

for 30 Epochs accuracy is 97.059%

for 35 Epochs accuracy is 97.206%

for 40 Epochs accuracy is 97.206%

for 45 Epochs accuracy is 97.206%

for 50 Epochs accuracy is 97.353%

For Ionosphere Dataset:

Vanilla Perceptron:

for 10 Epochs accuracy is 80.571%

for 15 Epochs accuracy is 82.286%

for 20 Epochs accuracy is 83.714%

for 25 Epochs accuracy is 74.857%

for 30 Epochs accuracy is 84.286%

for 35 Epochs accuracy is 86.000%

for 40 Epochs accuracy is 81.714%

for 45 Epochs accuracy is 84.571%

for 50 Epochs accuracy is 84.000%

Voted perceptron:

for 10 Epochs accuracy is 82.571%

for 15 Epochs accuracy is 85.429%

for 20 Epochs accuracy is 84.000%

for 25 Epochs accuracy is 84.000%

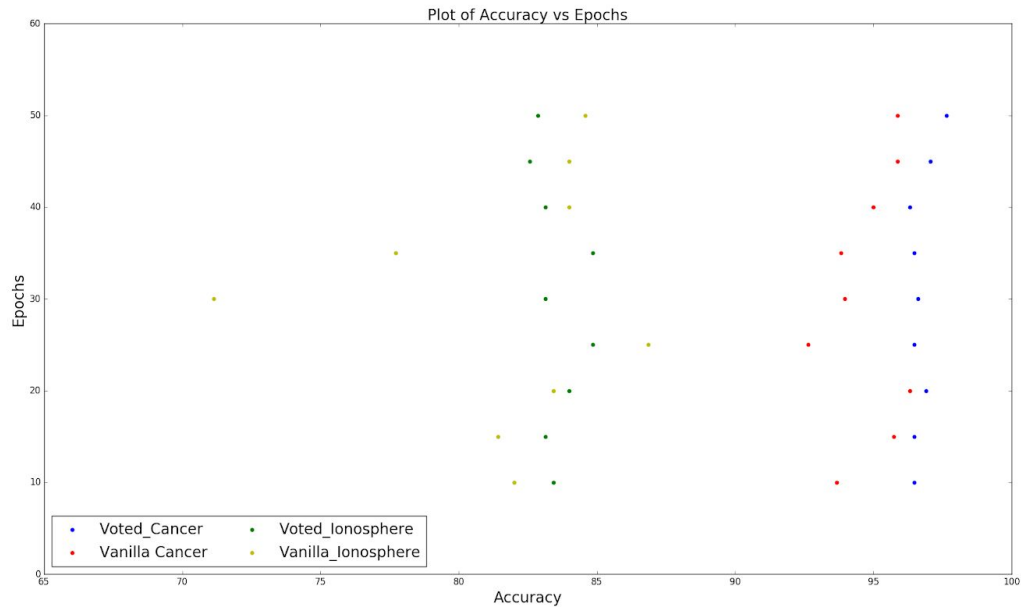for 30 Epochs accuracy is 85.714%

for 35 Epochs accuracy is 85.429%

for 40 Epochs accuracy is 85.143%

for 45 Epochs accuracy is 83.429%

for 50 Epochs accuracy is 83.143%

Accuracy vs Epochs:

Comment on Performance :

From the results obtained we can conclude that Voted perceptron is better than vanilla perceptron as we can see that the mean accuracy in both the data sets for vanilla is less than that of voted perceptron .

In voted perceptron we assign weights to different hyperplanes and take their contribution on the basis of how much data-points they are able to classify accurately whereas in case of vanilla perceptron we do not take weights of them and keep on updating them on the basis of classification.
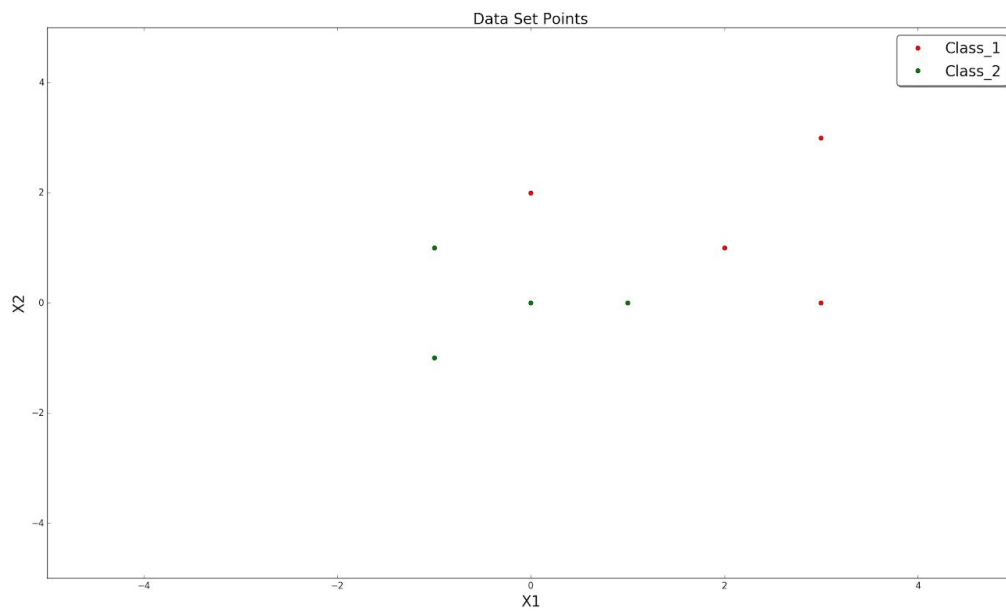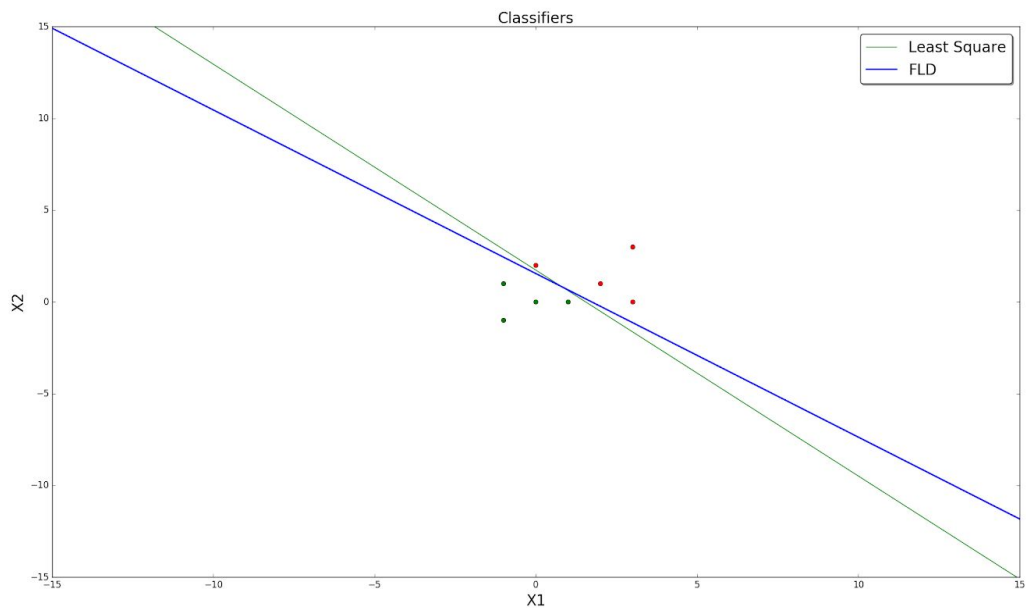
Q2)

Dataset 1:

Samples	Class_1	Class_2
1	(3,3)	(-1,1)
2	(3,0)	(0,0)
3	(2,1)	(-1,-1)
4	(0,2)	(1,0)

Dataset1 Graphical Visualisation :

Classifier Output 1:

For Least Square:

W for dataset 1 [[0.375], [0.33423913]]

B for dataset 1 [-0.57880435]

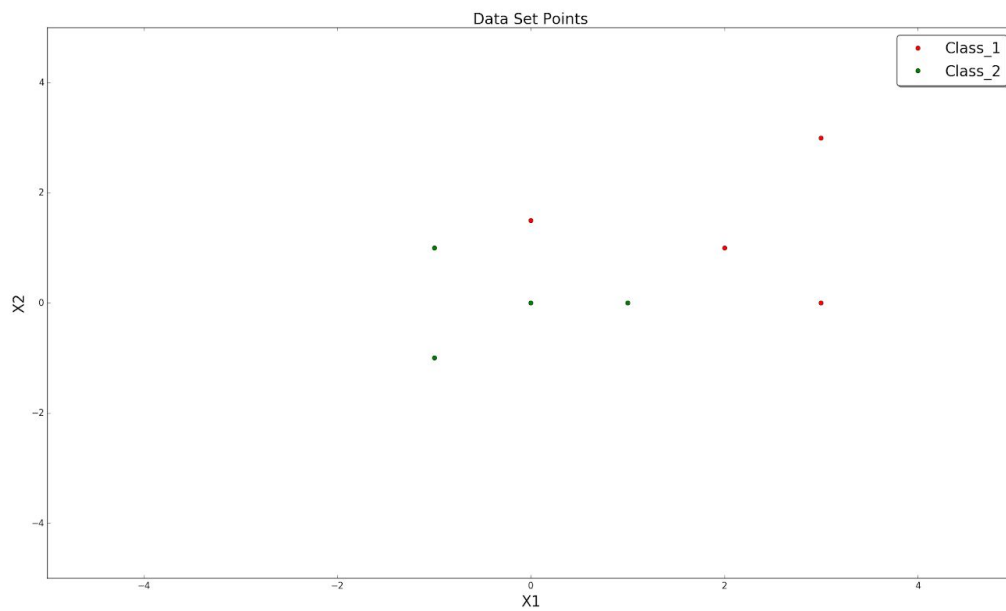**For Fisher:**

W for dataset 1 [[0.28630705], [0.25518672]]
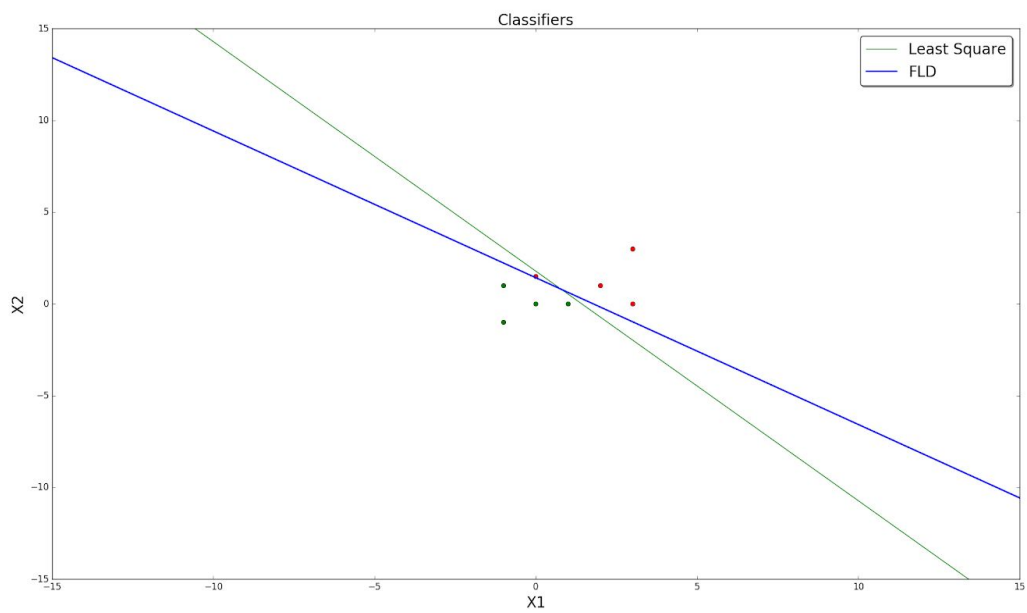
B for dataset 1 [0.44190871]]

Dataset 2:

Samples	Class_1	Class_2
1	(3,3)	(-1,1)
2	(3,0)	(0,0)
3	(2,1)	(-1,-1)
4	(0,1.5)	(1,0)

Dataset 2 Graphical Visualisation:

Data Set Points

Classifier Ouput_2 :

Classifiers

Least Square:

W for dataset2 : [[0.37779521], [0.30207925]]

B for dataset 2 : [-0.53825029]

For Fisher:

W for dataset2 : [[0.25714286], [0.20560748]

B for dataset2 :   [0.36635514]

Comment on Performance :

In 1st dataset both the classifiers are able to classify the two classes properly whereas in 2nd case Fisher's Linear discriminant is just able to classify the datasets and least square fails for 1 point in dataset2.

This is because even if the data is linearly separable the least square method may not be able to classify correctly whereas Fisher classifies correctly as we try to maximize between class variance and minimize within class variance.

Q3)

One vs Rest Perceptron

S.No.	Accuracy(%)	Discarded Pts	Epochs
1	90.3409090909	50	5
2	90.0568181818	50	5
3	94.3181818182	500	5
4	90.3409090909	250	5
5	86.0795454545	200	10

One vs Rest Perceptron

S.No.	Accuracy(%)	Discarded Pts	Epochs
1	90.3409090909	50	5
2	90.0568181818	50	5
3	94.3181818182	500	5
4	90.3409090909	250	5
5	86.0795454545	200	10

Cosine Similarity

S.No.	Accuracy(%)	Discarded Pts	Epochs
1	92.3302040509	50	5
2	93.0245671898	50	5
3	94.3123453345	500	5
4	96.2345356476	250	5
5	95.3254562423	200	10

The decline in accuracy may be because of overfitting. Accuracy for cosine similarity method keeps steadily increasing with increase in number of features selected because of increase in information gained.