BITS F464 Machine Learning

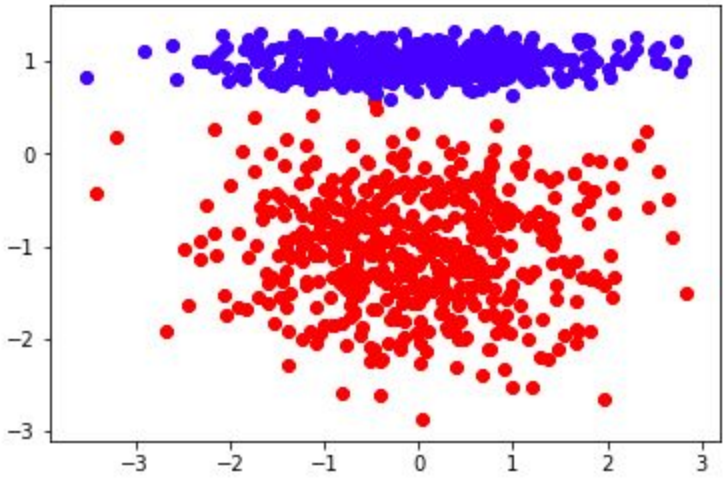
Assignment-1 Report

By:
Ayush Kumar Tiwary- 2016A2PS0567H
Mukesh Basira- 2016A7PS0116H
Srijan Soni- 2016A4PS0328H
Vivek Pratap Deo- 2016A7PS0056H

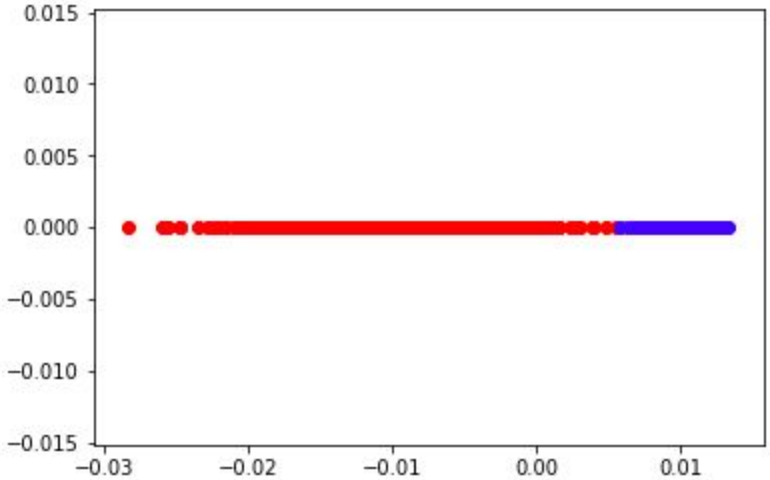
1. Fisher Linear Discriminant Analysis

Data 1:

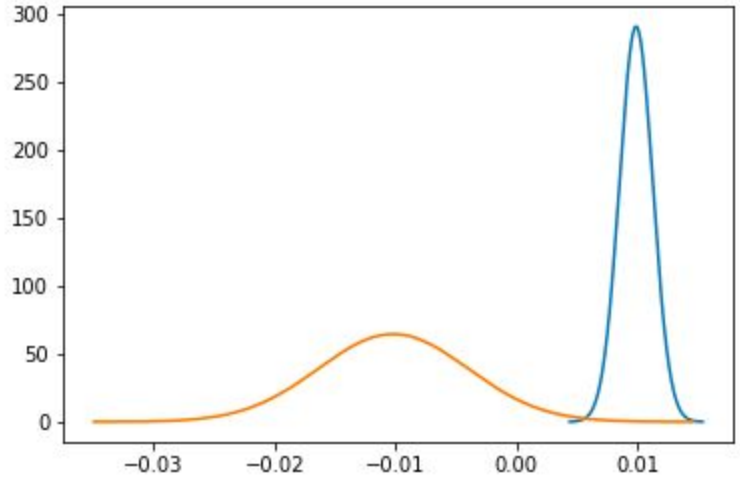
- Scatter Plot



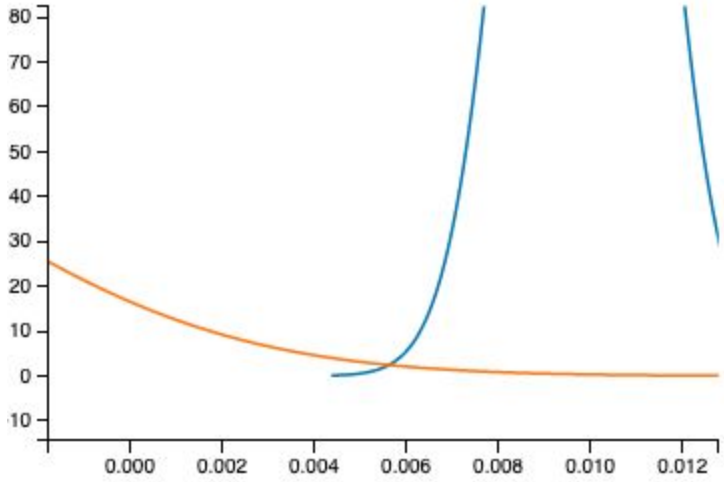
- Points Plotted



- Normal Plot

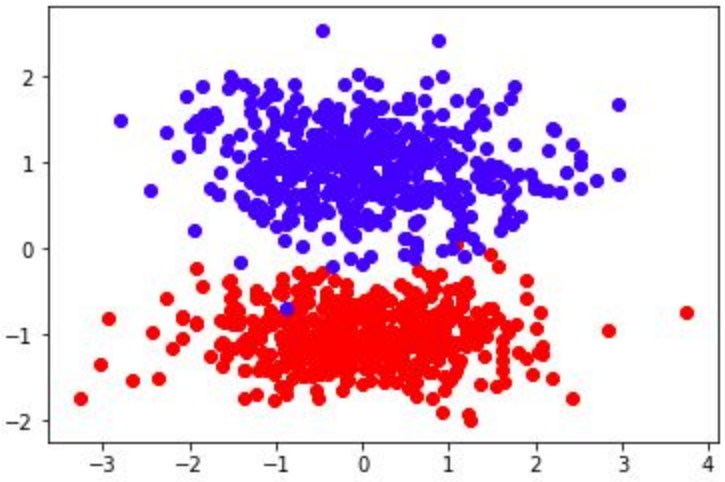


- Threshold plot

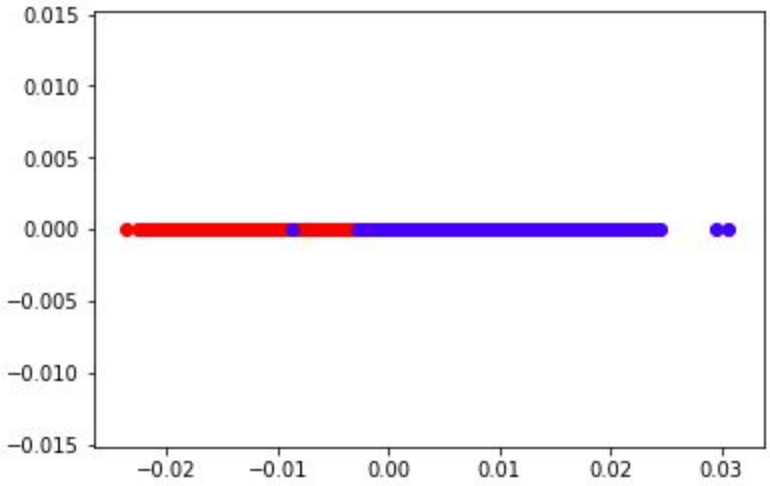


Data 2:

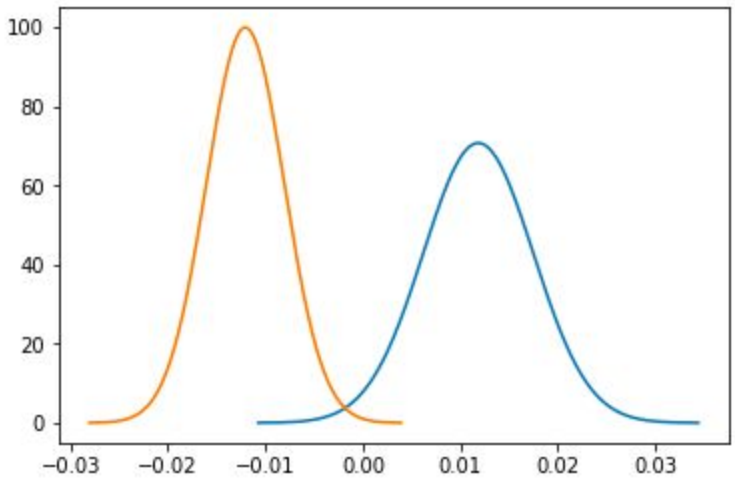
- Scatter Plot



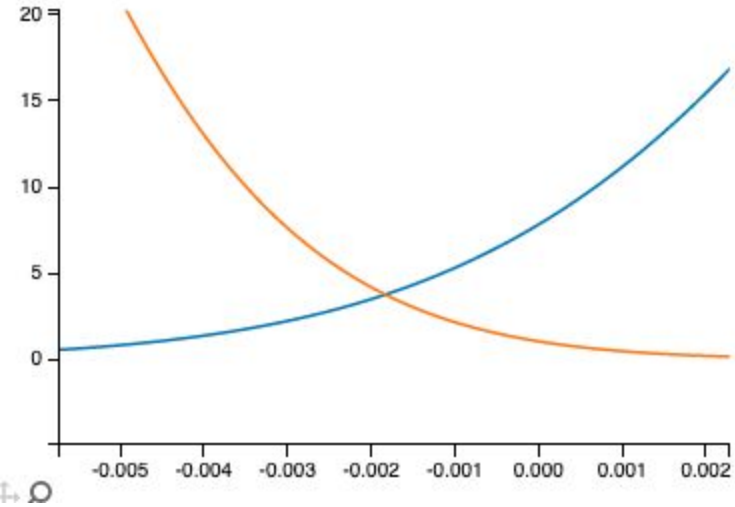
- Points Plotted



- Normal Plot

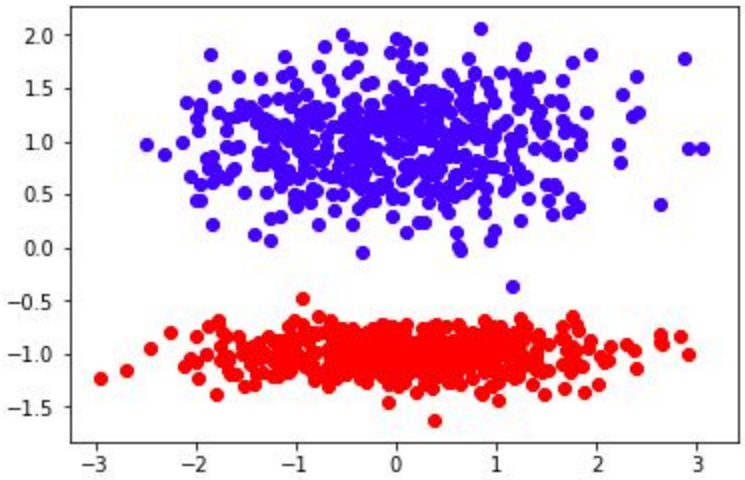


- Threshold plot

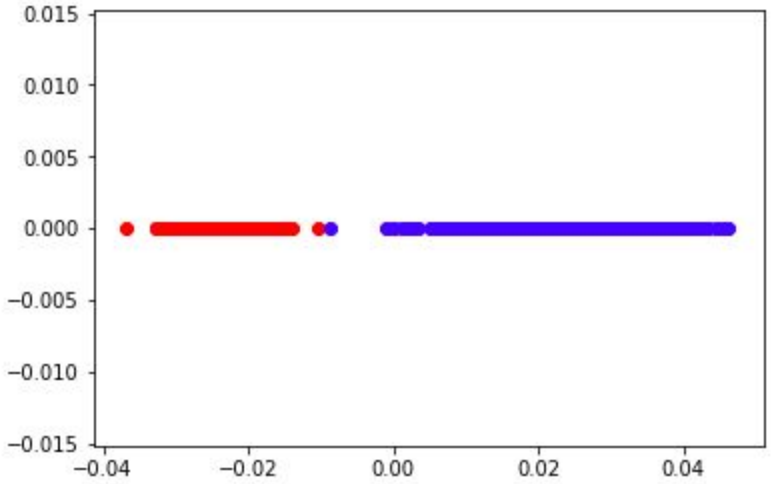


Data 3:

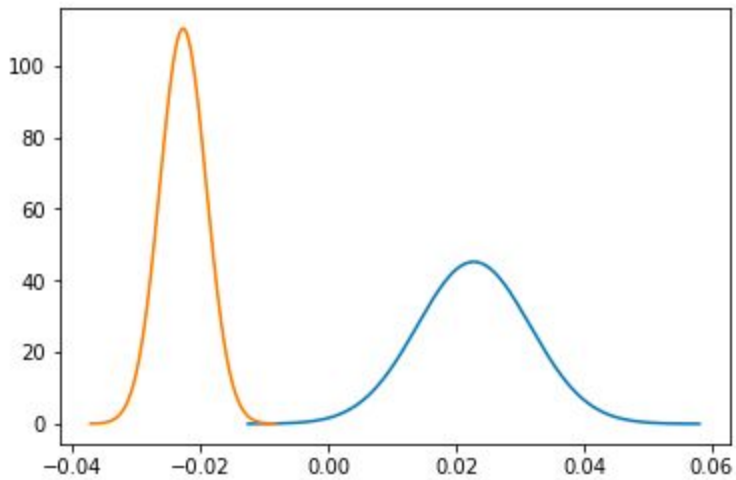
- Scatter Plot



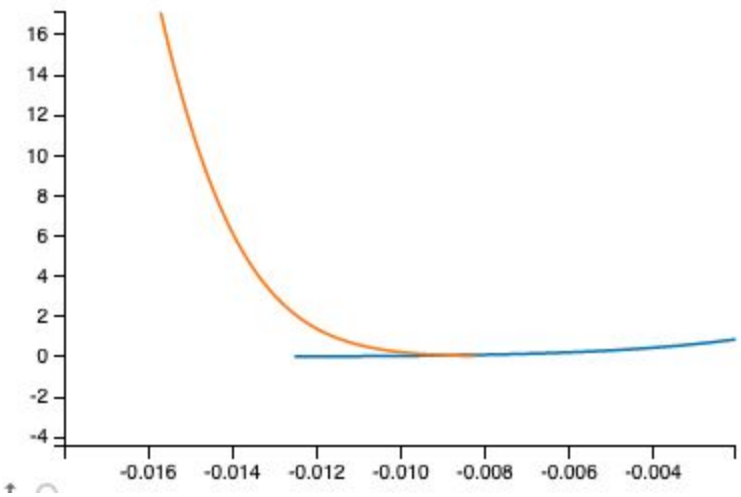
- Points Plotted



- Normal Plot



- Threshold plot



2. Perceptron

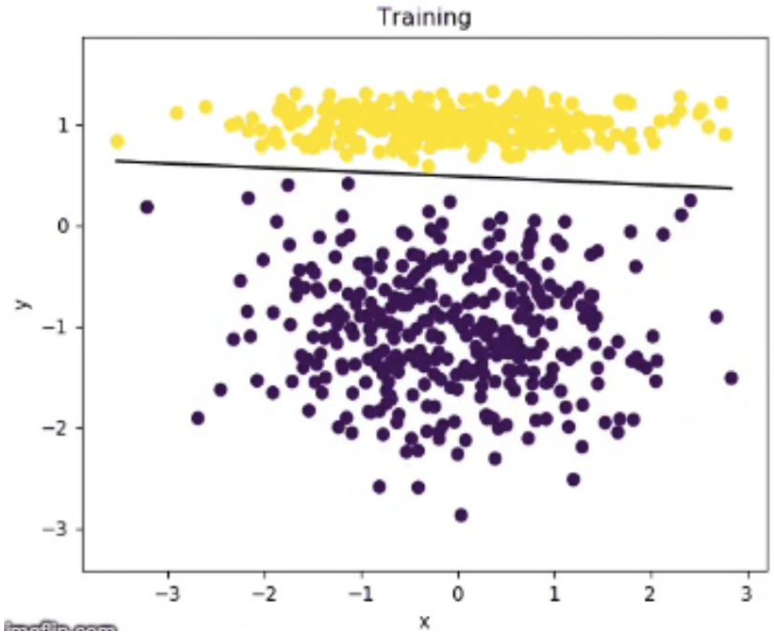
Perceptron Convergence Theorem

The perceptron convergence theorem basically states that the perceptron learning algorithm converges in finite number of steps, given a linearly separable dataset. More precisely, if for each data point x, $\| \mathbf{x} \| < R$, where R is certain constant number, $\gamma = (\theta^*)^T \mathbf{xc}$ where **xc** is the data point that is the closest to the linear separate hyperplane. It should be noted that mathematically $\gamma / (\| \theta^* \|^2)$ is the distance d of the closest datapoint to the linear separate hyperplane (it could be negative). The number of steps is bounded by $(R^2 \| \theta* \|^2)/\gamma^2$ or R^2/d^2.

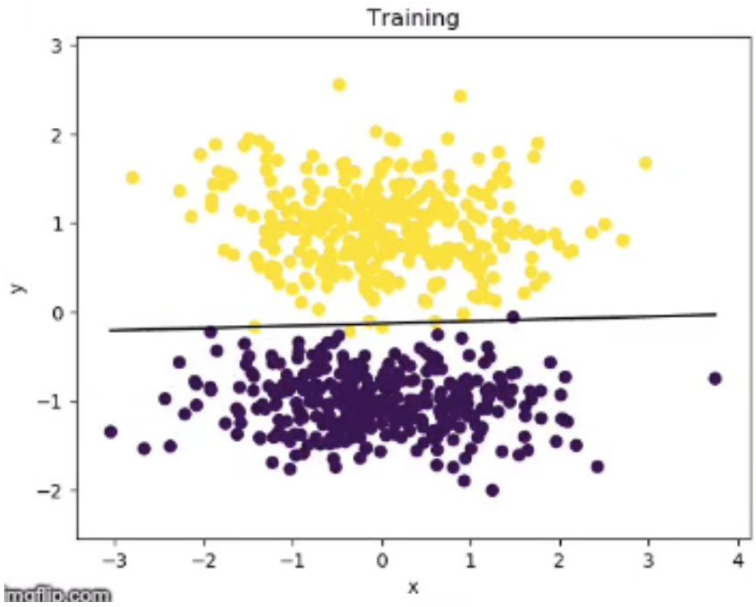
First Dataset is linearly separable. The perceptron converges to give 0 misclassified points in 100 epoches at 0.01 learning rate. Second Dataset is not linearly separable. The perceptron converges to give 5 misclassified points in 100 epoches at 0.01 learning rate. The algorithm will not give a perfect discriminant line. Third Dataset is linearly separable.It also converges to give 0 misclassified points in 100 epoches at 0.01 learning rate.

The gifs for the three datasets have been saved in Data folder.

Data 1 discriminant:



Data 2 discriminant:



Data 3 discriminant: