Machine Learning Assignment 1

Contributors-
Ameya Salankar - 2017A7PS0182H
Atharva Sune - 2017A7PS0183H
Ayush Laddha - 2017A8PS0717H

Fisher's LDA

Design decisions -

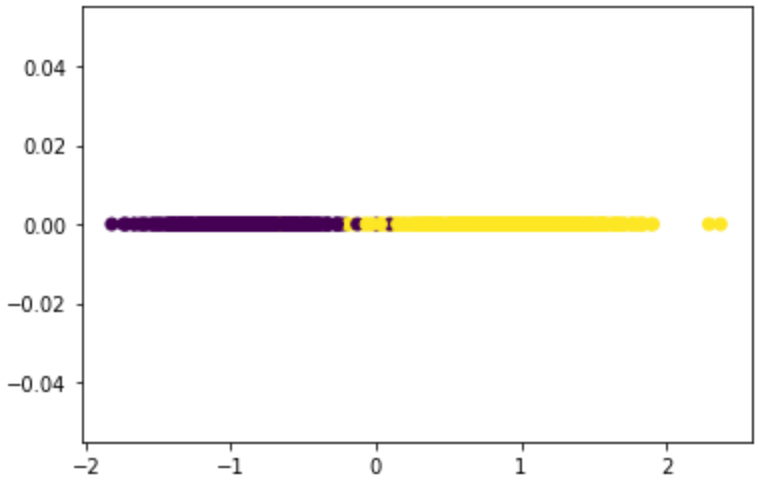
1. If w is the transformation vector and S_B is the between class variance and the S_W is the within class variance, we know that
$$(w^T S_B w) S_W w = (w^T S_W w) S_B w$$
Since the quantities inside the brackets are scalar and $S_B w$ is in the direction of $(\mu_2 - \mu_1)$, we can say that
$$w \, \alpha \, S_w^{-1}(\mu_2 - \mu_1)$$

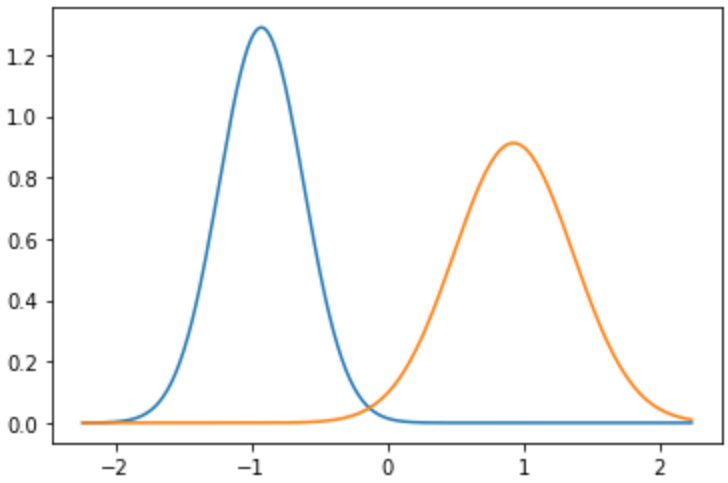
   Thus, the transformation vector was calculated as
$$w = S_W{}^{-1} \times (\mu_2 - \mu_1)$$
   The transformation vector was then made into a unit vector.
   This was done to get rid of the proportionality constant.

2. The intersection point of the two normal curves was obtained by solving the equations of the two normal curves. This yielded two roots. The root that was chosen was the one which was between the means of the normal distributions. This was done because on plotting the graphs, it was observed that the intersection point came between the two means.

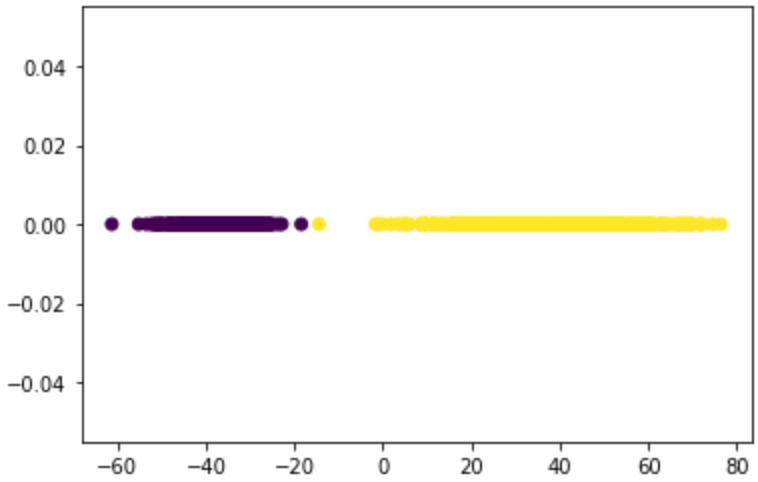
Results -

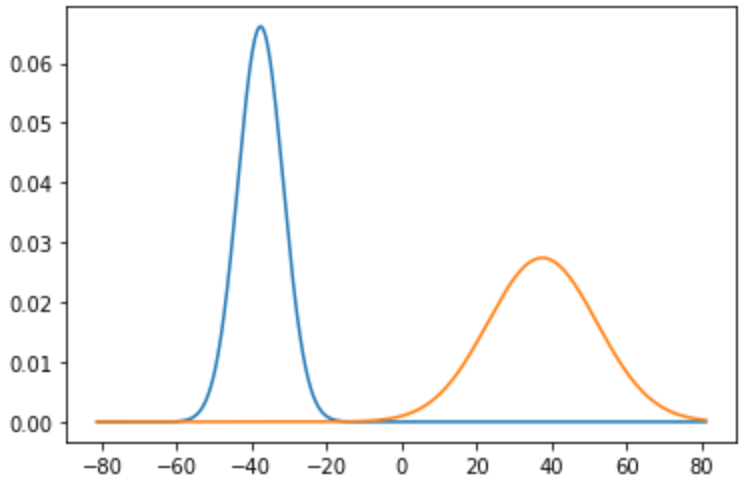
1. Applying the LDA on the first dataset (which was 2D) yielded
   Threshold point as **-0.13328610589766562**
   An accuracy of 99.3%
   An F-score of 0.993





2. Applying the LDA on the second dataset (which was 3D) yielded
   Threshold point as **-0.3707808853290484**
   An accuracy of 100%
   An F-Score of 1.0





Conclusion -

We can say that the Fisher's LDA does a good job on both the datasets, particularly the second one. This can be inferred from the high accuracy reported which were more than 99% for both the datasets.

Fisher's LDA performed good here because the data was linearly separable to a large extent. If this was not the case, the algorithm would have failed if the projection on the best line fails to result in a linearly separable situation.