Fisher's Linear Discriminant Analysis

- ## Overview

Fisher's LDA method is one of the popular methods used in binary and multi-classification problem. This involves reduction of parameters, so as to reduce the number of dimensions of a point. The basic idea is to find a vector W, that serves as a decision boundary between the classes (binary classification for the assignment problem). W vector is chosen so as to:

1. Maximise the difference of means of both classes
2. Minimise the sum of within class covariances of the classes.

- ## Implementation

1. After loading the data into an array, compute the mean(m1,m2) and the within class deviation from the mean (S1,S2), for both classes.
2. Compute the difference of means (m1-m2) and the sum of the covariance matrices (***Sw = S1+S2***).
3. The vector W will be proportional to the product of inverse of Sw and (m1-m2).
   $W = Sw^{-1} * (m1 - m2)$
4. Project the dataset from n dimensions to 1 dimension by taking projection from vector W. By doing this we obtain a new set of transformed points which are linearly separable.
5. Draw the normal curves for these projected points (separately for both classes), and find the point of intersection.
6. Project the point of intersection back onto the line to get the threshold value in 1D. All points that are projected onto this line and lie to the left of this threshold will be classified as negative points and others as positive.
7. Project the points back in 3D to obtain the decision boundary plane.
8. Accuracy of the model is calculated by the formula,
   Accuracy = (True positives + True negatives) / Total Number of points
   , where , True positives is the number of positive points classified correctly and True negatives is the number of negative points classified correctly.

Note: We have assumed that all the positive points in the dataset follow a Normal distribution with mean m1 and Variance S1, while negative points follow a Normal distribution with mean m2 and Variance S2.

- **<u>Plots</u>**

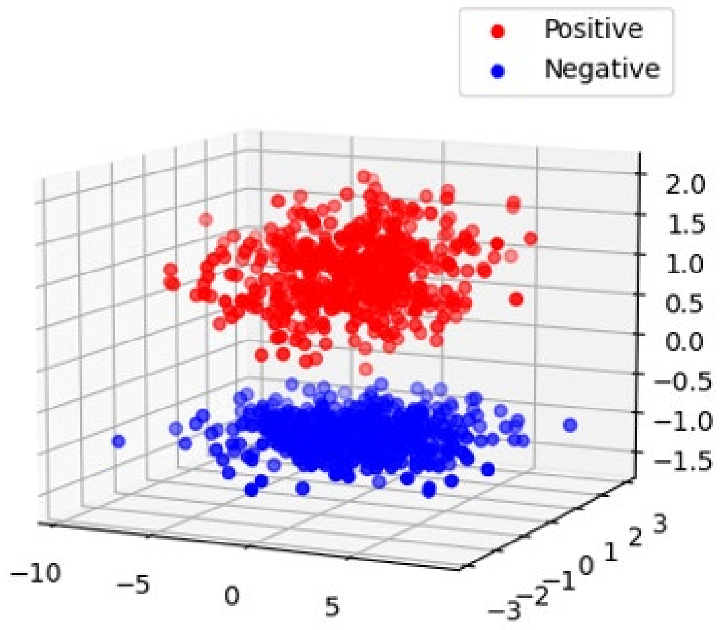

Figure 1: 3D Scatter Plot of the Points

Projecting the 3D points using W vector, onto 1D, and then constructing the normal distribution curves of both classes, we calculate the Point of Intersection of the curves (Threshold Point).

Threshold = -0.3893

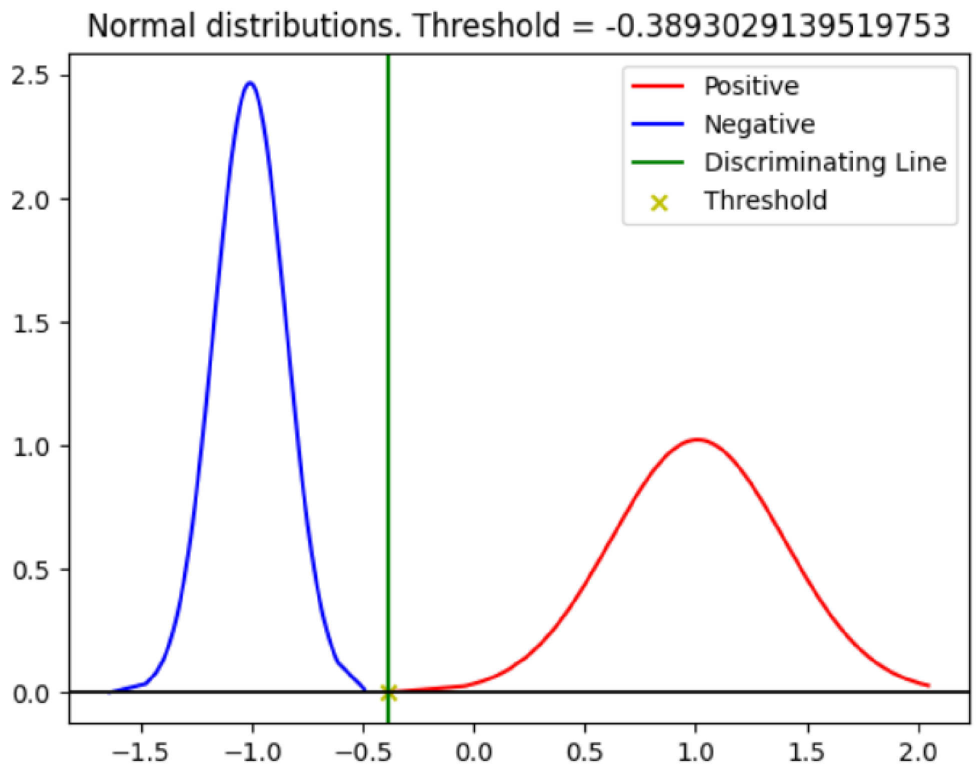

Figure 2: Normal Distributions with Discriminating Line

Passing through the Threshold and Perpendicular to the vector W (in 1D) is the Discriminating Line, which serves as the decision boundary in 1D.

Discriminating Line => x = -0.3893

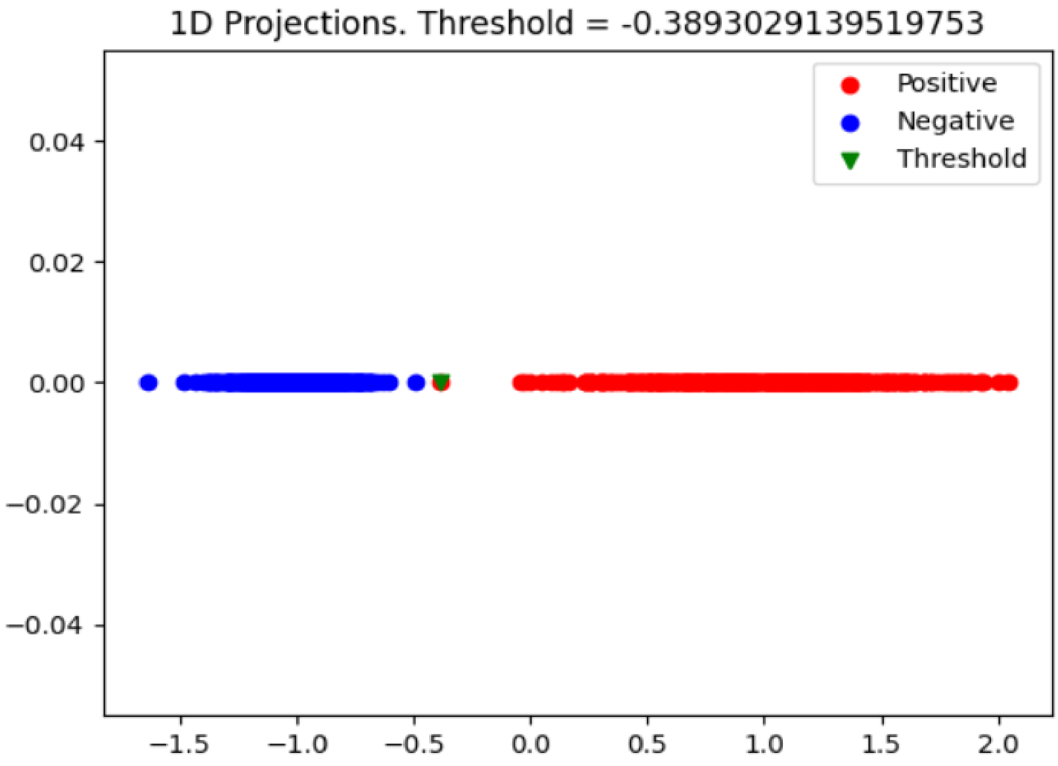

Figure 3: Projections onto 1D with Threshold

Reprojecting the 1D vector and decision boundary we obtain the Discriminating Plane, which serves as the decision boundary in 3D (Line projected to 3D is a plane).

Discriminating Plane => $W^T * X = -0.3893$

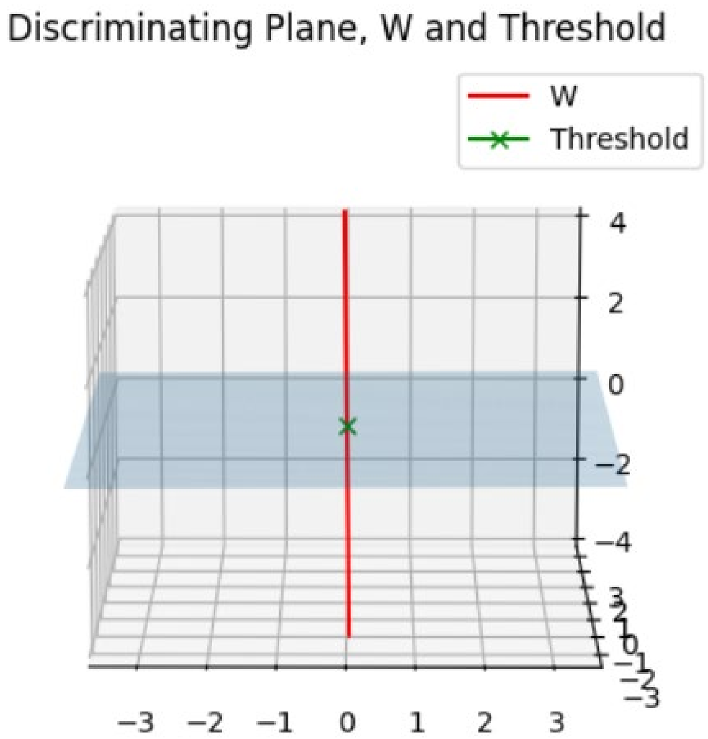

Figure 2: Discriminating Plane with W and Threshold

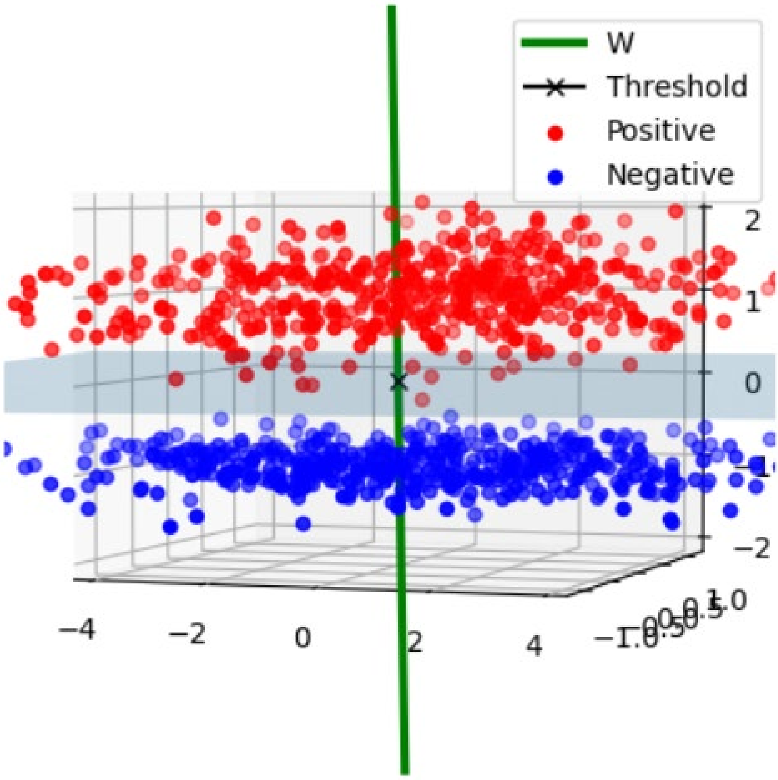

Figure 3: Discriminating Plane with original points and Threshold

- **Results:**

W -> [-0.00655686, -0.01823739 , 0.9998122]
Threshold -> -0.38930291576654724
Discriminating Line -> x = -0.38930291576654724
Discriminating Plane-> W ^{T} * X = -0.38930291576654724 , where X is [x, y, z]
Accuracy -> 100%

All points in the 3D space that lie below the Discriminating Plane will be classified as negative points while others as positive points.

Conclusion: Since the points are linearly separable, Fisher Discriminant Analysis gives 100% accuracy, however, had the points not been almost linearly separable then FDA would've given inaccurate results.