Fischer's Linear Discriminant Analysis

Group Members:

Sarthak Gaur - 2017A7PS0250H
Smit D Sheth - 2017A7PS1666H
Sourav Sanganeria - 2017A7PS1625H

Abstract:

Fischer's Linear Discriminant Analysis is a method used in statistics, pattern recognition, and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification.

Fischer's LDA tries to maximize the difference between the means of the two classes and minimize the within-class covariances in their transformed direction. With this objective, we get the best possible direction to which the data points need to be transformed.

After this, a point is decided in this direction which separates the two classes in the best way. This point can be the intersection point of the Gaussian curves for the two classes. When a new point comes, it is then tested against this point to classify it into one of the two classes.

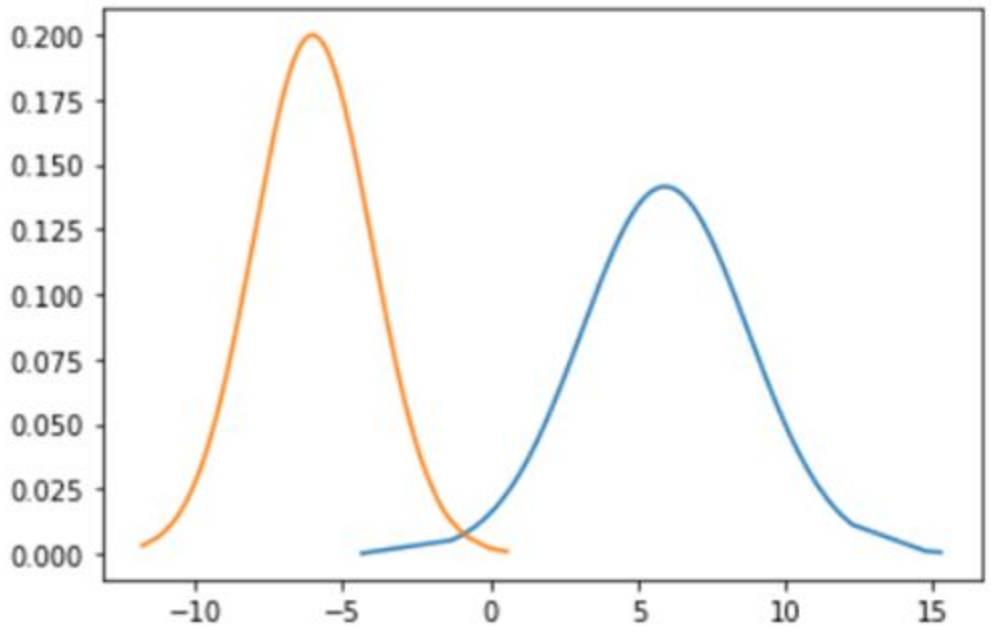
Performance:

This technique was applied in the datasets 1 and 2 and the following results were found out:

Dataset 1:

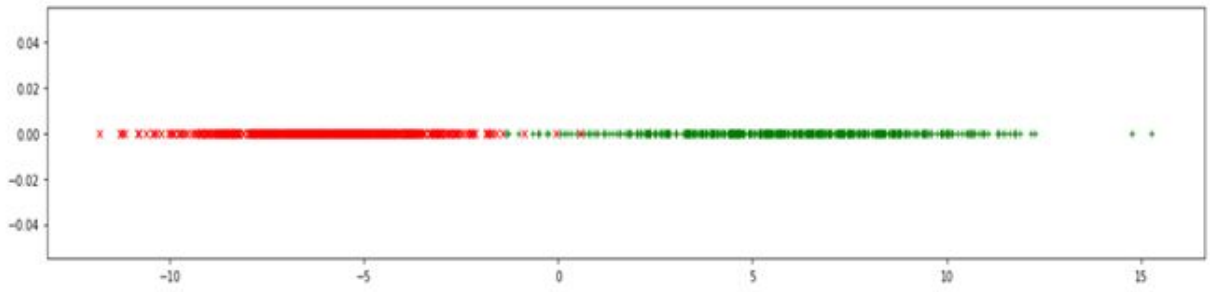
Transformation matrix(W): [0.19892987, 6.01643137]

Gaussian curves for the two classes:



Cut-point: -0.9154240171714463

Datapoints transformed in the direction found from Fischer's LDA:



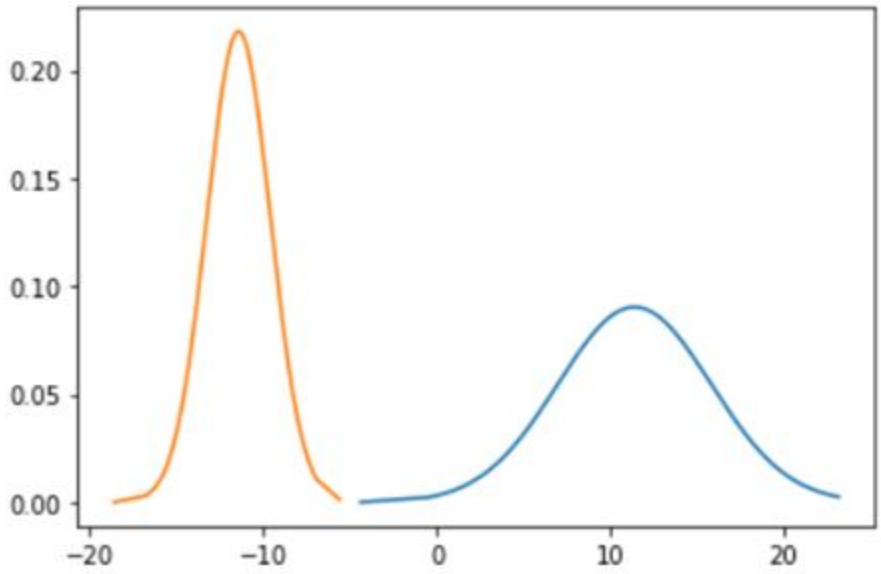
Accuracy: 99.3%

F score: 0.993

Dataset 2:

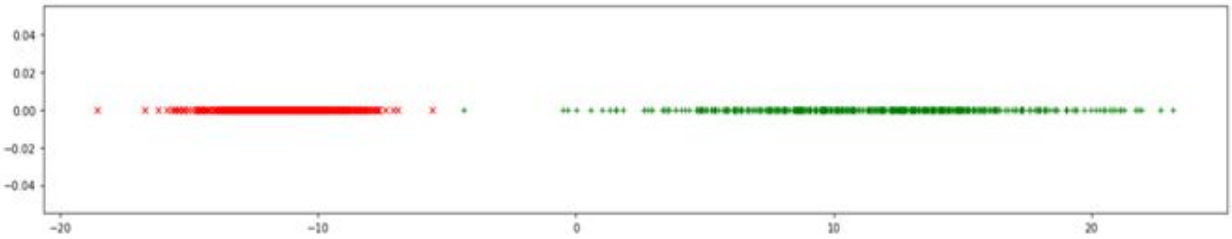
Transformation matrix(W): [-0.0742240, -0.2064484, 11.31793517]

Gaussian curves for the two classes:



Cut point: -4.4069315647339335

Datapoints transformed in the direction found from Fischer's LDA:



Accuracy: 100%

F score: 1.0

Conclusion:

The data points were classified into the two classes with an accuracy of 99.3% and 100% respectively. Their F scores are also close to 1. This proves that this classification technique works really well for both the datasets.