# **Depth Linear Discriminant Analysis**

This study introduces a new linear discriminant method, called Depth Linear Discriminant Analysis (D-LDA) to enhance the robustness of the original LDA. D-LDA systematically integrates the matrix depth concept into LDA, offering a systematic approach to address the challenges associated with scatter matrix estimation. As matrix depth measures how central or deep a particular matrix is within a distribution with respect to different directions, it is an efficient tool for computing a robust scatter matrix estimator that can handle outliers and complex data structures.

## **Linear Discriminant Analysis Methods**

Linear discriminant analysis is a Dimensionality Reduction (DR) technique that has been widely used in classification to find a projection of data points that maximize the ratio of between-class and within-class scatter. By maximizing the separability between data points, LDA can improve the accuracy of the classification process, effectively identifying and distinguishing different samples based on their features. Notably, LDA and its counterparts are not FS methods. In DR techniques, the objective is to find a set of features equivalent to the original features instead of finding the top set of features. LDA is a powerful tool for classification and dimensionality reduction, with several LDA variants and alternatives proposed; however, it has some performance limitations. In this review, we present existing works and categorize them based on the specific problems they aim to address.

One of the main problems with LDA is its sensitivity to outliers that are significantly different from the rest of the data. To address this problem, numerous studies have proposed methods, including robust measures into LDA, using the L1-norm of the projection matrix and its extension 2DLDA-T ℓ1 (Yang, Zheng, & Liu, 2023), integrating two-dimensional LDA and the T ℓ1 -norm, and using robust distance metrics such as the Mahalanobis distance, which are less sensitive to outliers. A recent example of an LDA-based approach is the Wasserstein Discriminant Analysis (WDA) (Flamary, Cuturi, Courty, & Rakotomamonjy, 2018; Kuhn, Esfahani, Nguyen, & Shafieezadeh-Abadeh, 2019), which is based on the Wasserstein distance. Similar to other discriminant algorithms, it aims to maximize the separation between classes while minimizing it within classes. Furthermore, WDA utilizes optimal transport theory to find an optimal transportation matrix that aligns with the original data.

## The Proposed Depth Linear Discriminant Analysis Method

To enhance the robustness of linear discriminant methods, D-LDA improves the accuracy and reliability of LDA by addressing overlapping classes, and data outliers, employing the concept of matrix depth for robust scatter matrix estimation. Matrix depth measures centrality or how far out matrix outliers are within a given distribution. The depth matrix can be used to develop a robust covariance and scatter matrix estimator for handling outliers and complex data structures. Matrix depth D(Σ, P) quantifies how well the estimator Σ captures the structure in the given data points by moving the estimator towards the deepest point to obtain the most accurate estimation of the scatter matrix that can enhance the performance of LDA. A lower matrix depth implies a greater centrality within the distribution.

## **Comparison D-LDA with other discrimination methods:**

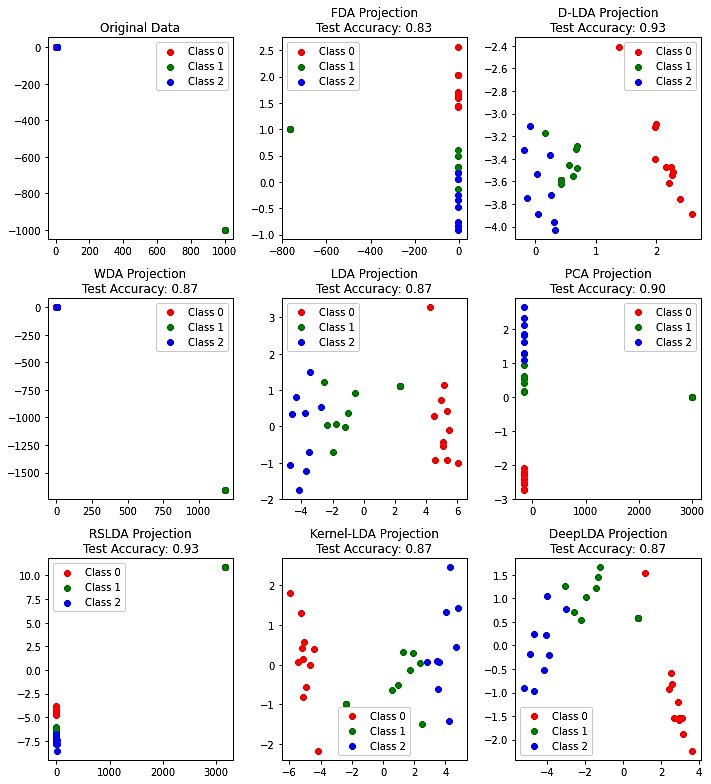
This section assesses the proposed depth linear discrimination analysis method, namely D-LDA, comparing D-LDA with several other DR methods, including FDA, Wasserstein Discriminant Analysis (WDA), LDA, PCA, Robust Sparse LDA (RSLDA), Kernel LDA (KLDA), and Deep LDA (DeepLDA). All these methods can reduce the feature space; therefore, our evaluation focused on assessing the impact of dimensionality reduction on the accuracy and robustness of the methods.

To illustrate the ability of D-LDA to linearly separate classes and handle outliers, Figures 1 and 2 present a comparison of D-LDA with existing methods using a simple dataset. These figures also show the accuracy of each method. **Fig. *1*** shows the results of applying the dimensionality reduction methods to normal data, whereas **Fig. *2*** shows the results after introducing outliers. The results show the superiority of the D-LDA method over the other methods in the presence and absence of outliers, except for the RSLDA method, which obtained results close to those of the D-LDA method. The results of KLDA and RSLDA were competitive and comparable to those of D-LDA.

A screenshot of a graph

Description automatically generated

**Fig. 1** Data transformation and accuracy comparisons without outliers.



**Fig. 2** Data transformation and accuracies compared in the presence of outliers.