Definition 1 Linear discriminant analysis (LDA) *or discriminant function analysis is a generalization of Fisher's linear discriminant, a method used in statistics and other fields like Machine Learning, to find a linear combination of features that characterizes or separates two or more classes of objects or events*

Linear discriminant analysis is a dimensional reduction and supervised algorithm that maximized the component axis for class separation. LDA finds the component axes with a multiple class that maximizes the variance of the individual classes of the data, there by finding the axis that optimized the separation between multiple classes.
Linear Discriminant Analysis seeks to best separate (or discriminate) the samples in the training dataset by their class value. Specifically, the model seeks to find a linear combination of input variables that achieves the maximum separation for samples between classes (class centroids or means) and the minimum separation of samples within each class.

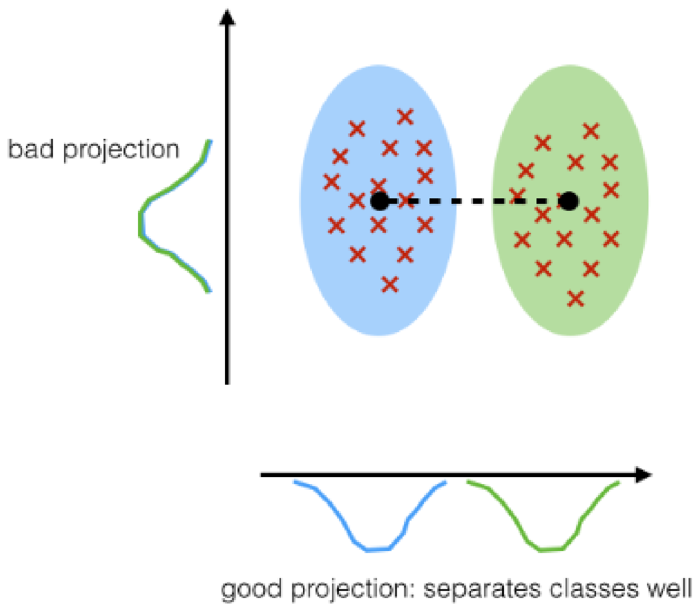

Figure 1:

Definition 2 curse of dimensionality *is the dramatically impact on the performance of machine learning algorithms fitted on data with many input features.*

When dealing with high dimensional data, it is often useful to reduce the dimensionality by projecting the data to a lower dimensional subspace which captures the "essence" of the data. This is called dimensionality reduction

Mathematical Representation

The goal is to find the eigenvalues and eigenvectors of the matrix the dot product of the inverse of the within-class scatter matrix and the between class scatter matrix.

Within-Class Scatter Matrix

$$S_W = \sum_c S_c \tag{1}$$

where,

$$S_c = \sum_{i \in c} (x_i - \bar{x}_c) \cdot (x_i - \bar{x}_c)^T$$

Between-Class Scatter matrix

$$S_B = \sum_c n_c \cdot (\bar{x}_c - \bar{x}) \cdot (\bar{x}_c - \bar{x})^T \tag{2}$$

where, n is the number of class labels

suppose M is the matrix of the LDA then we have,

$$M = S_W^{-1} \cdot S_B \tag{3}$$

Procedure:

- Finding the Eigen-values, and Eigen-vectors of M
- Sort the vector according to their Eigen-values in decreasing order
- Choose k Eigen-vectors and that will be the k-dimension
- Transform the original n dimensional data points onto k-dimensional(i.e projection with dot product)