STEPS:

1. Compute the within class and between class scatter matrices
2. Compute the eigenvectors and corresponding eigenvalues for the scatter matrices
3. Sort the eigenvalues and select the top ***k***
4. Create a new matrix containing eigenvectors that map to the ***k*** eigenvalues
5. Obtain the new features (i.e. LDA components) by taking the dot product of the data and the matrix from step 4

(a) Within Class Scatter Matrix

We calculate the *within class scatter matrix* using the following formula.

$$S_W = \sum_{i=1}^{c} S_i$$

where ***c*** is the total number of distinct classes and

$$S_i = \sum_{\boldsymbol{x} \in D_i}^{n} (\boldsymbol{x} - \boldsymbol{m}_i)\,(\boldsymbol{x} - \boldsymbol{m}_i)^T$$

$$\boldsymbol{m}_i = \frac{1}{n_i} \sum_{\boldsymbol{x} \in D_i}^{n} \boldsymbol{x}_k$$

where ***x*** is a sample (i.e. row) and ***n*** is the total number of samples with a given class.

For every class, we create a vector with the means of each feature.

(b) Between Class Scatter Matrix

Next, we calculate the *between class scatter matrix* using the following formula.

$$S_B = \sum_{i=1}^{c} N_i(\boldsymbol{m}_i - \boldsymbol{m})(\boldsymbol{m}_i - \boldsymbol{m})^T$$

where

$$\boldsymbol{m}_i = \frac{1}{n_i} \sum_{\boldsymbol{x} \in D_i}^{n} \boldsymbol{x}_k$$

$$m = \frac{1}{n} \sum_{i}^{n} x_i$$

(c) Then, we solve the generalized eigenvalue problem for

$$S_W^{-1} S_B$$

to obtain the linear discriminants.

(d) The eigenvectors with the highest eigenvalues carry the most information about the distribution of the data. Thus, we sort the eigenvalues from highest to lowest and select the first $\boldsymbol{k}$ eigenvectors. In order to ensure that the eigenvalue maps to the same eigenvector after sorting, we place them in a temporary array.