EE527: Machine Learning Laboratory

Assignment 8

Due Date: 28 March 2022

1. Application of the *Discriminant Functions* in Classification of Normal and Shouted Speech using MFCC features. These features are extracted from speech samples of a number of speakers uttering a few sentences normally or by shouting. The features are divided into train-test splits and are made available in two csv files. You are tasked to learn a discriminative model to classify normal and shouted speech. This example uses discriminative functions. The whole process is described as follows.

Consider the .csv file *"Train_file.csv"* containing *86060* instances of 61-dimensional arrays. The first *60* dimensions of the array contain the feature values for a particular instance and the last dimension contains its label. The label can be either *'0'* or '1'.

(a) Evaluate $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_1$, $\mathbf{C}_0$, $\mathbf{C}_1$ from instances in *"Train_file.csv".*

$\boldsymbol{\mu}_0$: Mean of all instances having label `0'

$\boldsymbol{\mu}_1$: Mean of all instances having label `1'

$\mathbf{C}_0$: Covariance matrix of all instances having label `0'

$\mathbf{C}_1$: Covariance matrix of all instances having label `1'.

(b) Construct parameters of the discriminant function with Gaussian assumption on instance distribution in classes.

$$g(x) = ln\left\{\frac{P(X1)}{P(X0)}\right\} - \frac{1}{2}\left({\mu_1}^T {C_1}^{-1}\mu_1 - {\mu_0}^T {C_0}^{-1}\mu_0\right) - \frac{1}{2} ln\left\{\frac{|C_1|}{|C_0|}\right\}$$
$$+ x^T\left({C_1}^{-1}\mu_1 - {C_0}^{-1}\mu_0\right) - \frac{1}{2}x^T({C_1}^{-1} - {C_0}^{-1})x$$

The decision rule for classification of an unseen instance x is given by its label $y(x)$ defined as

$$y(x) = \begin{cases} 1, g(x) \geq 0 \\ 0, g(x) < 0 \end{cases}$$

(c) Read *"Test_file.csv"* consisting of *21516* instances of 61 dimensional arrays. For each array, the first 60 dimensions contain the feature values for the test data and the last dimension contains its *actual label*. Predict the label of each data instance from the testing set using the decision rule mentioned above and compare the *predicted* and *actual* labels. Report the class-wise (ρ_0 and ρ_1) and overall (ρ) accuracy measures.

$$\rho_0 = \frac{No.of\ Correctly\ Classified\ Instances\ in\ Class\ "0"}{Total\ Number\ of\ Points\ in\ Class\ "0"}$$

$$\rho_1 = \frac{No.of\ Correctly\ Classified\ Instances\ in\ Class\ "1"}{Total\ Number\ of\ Points\ in\ Class\ "1"}$$

$$\rho = \frac{Total\ No.of\ Correctly\ Classified\ Instances}{Total\ Number\ of\ Test\ Instances}$$

2. Application of *Linear Discriminant Analysis (LDA)* in understanding the separability in the dataset of MFCC features used for classifying Shouted and Normal Speech. The two files *MFCC_N.npy* and *MFCC_S.npy* respectively store the features of normal and shouted speech.

a) Load the two datasets *MFCC_N.npy* and *MFCC_S.npy* into arrays ***N*** and ***S*** respectively. Use the python function *np.load(filename).*
b) Compute centroids *(*m_n *,* m_s *)* and covariance matrices *(*C_n*,* C_s*)* of both arrays.
c) Compute the optimal direction vector $\widehat{\omega}$ (unit vector) for LDA.
d) Project the vector data in arrays ***N*** and ***S*** to generate the respective array of scalars ***zN*** and ***zS*.**
e) Plot the normalized histograms of ***zN*** and ***zS*** in two different colors (red and blue).

3. Application of *K-Means Clustering* in image segmentation. Consider the *R-G-B* values of each pixel of the input image as *3*-dimensional feature vector. Initialize *K-Means* through data labels or cluster centroids. Perform *K-means* iterations till convergence and report the *K* cluster centroids. Revisit the Image and replace each pixel color (*R-G-B*) with the nearest cluster centroid (rounded) color values. Repeat this experiment with different values of *K* and visualize the results.