

Indian Institute of Information Technology, Design and Manufacturing
(Chennai -600 127, India)
Pattern Recognition
Assignment - 1
Deadline : 11 Feb 2022

Note: Make use of MATLAB or Python for implementation.

Deliverables for this assignment:

Code file and output screenshots for all the questions.

1. Calculate the distance between the two normalized histograms H1 and H2 using each of the following methods:

- (a) KL Distance
- (b) Bhattacharyya Distance

H1 = [0.24, 0.2, 0.16, 0.12, 0.08, 0.04, 0.12, 0.04]

H2 = [0.22, 0.23, 0.16, 0.13, 0.11, 0.08, 0.05, 0.02]

2. Given $(hq - ht)^T = (0.5 \ 0.5 \ -0.5 \ -0.25 \ -0.25)$ and

$$A = \begin{pmatrix} 1 & 0.135 & 0.195 & 0.137 & 0.157 \\ 0.135 & 1 & 0.2 & 0.309 & 0.143 \\ 0.195 & 0.2 & 1 & 0.157 & 0.122 \\ 0.137 & 0.309 & 0.157 & 1 & 0.195 \\ 0.157 & 0.143 & 0.122 & 0.195 & 1 \end{pmatrix}$$

Find the quadratic form distance.

3. Compare two text files doc1.txt and doc2.txt using cosine distance.

doc1.txt

MATLAB is a program for solving engineering and mathematical problems. The basic MATLAB objects are vectors and matrices, so you must be familiar with these before making extensive use of this program.

doc2.txt

MATLAB works with essentially one kind of object, a rectangular numerical matrix. Here is some basic information on using MATLAB matrix commands.

4. Classify flower 1, 51, and 101 from the Iris Dataset (.csv file) attached along with the assignment document into one of the three classes as given in dataset specification:

Dataset Specifications:

Total number of samples = 150

Number of classes = 3 (Iris setosa, Iris virginica, and Iris versicolor)

The number of samples in each class = 50

Directions to classify:

- 1. Use features PetalLengthCm and PetalWidthCm only for classification.
 - 2. Consider flowers 1, 51 and 101 as test cases.
 - 3. Plot the distribution of rest 147 sample points along with their classes(differentiate classes with different colour). Consider PetalWidthCm along Y-axis and PetalLengthCm along X-axis.
 - 4. Capture the properties of the distribution and use suitable distance metrics to classify the flowers 1, 51 and 101 into one of the classes.
 - 5. Print their class and plot the points on the previous plot with a marker differentiating the three points.
5. Consider the following images. Obtain the histograms for each of the images. Using a suitable distance measure, find the distance between the query image and reference images.



Figure 1: Query Image



Figure 2: Reference Image 1

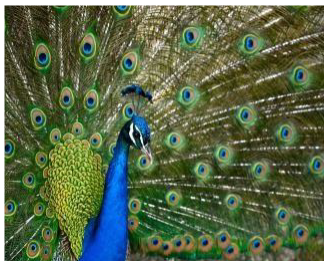


Figure 3: Reference Image 2