INDIAN INSTITUTE OF TECHNOLOGY TIRUPATI
भारतीय प्रौद्योगिकी संस्थान तिरुपति

EE 5024 Machine Learning For Image Processing (Jul-Nov 2018)

Programming Assignment 1

K – NN Classification:

Download the CIFAR 10 Dataset from the zip file and experiment with K = 1, 3, 5, 9 and 11 for each of the following tasks.

Task	DATASET	Feature Vector To Extract	Metric To Be Used
1	CIFAR 10	Histogram of Image	Eucledian Distance
2	CIFAR 10	Histogram of Image	NCC

NCC-Normalized cross correlation

Compare the above mentioned tasks in terms of Accuracy and observe the variation of accuracy with K. (plot Accuracy vs. K and infer)

CIFAR10 Dataset:

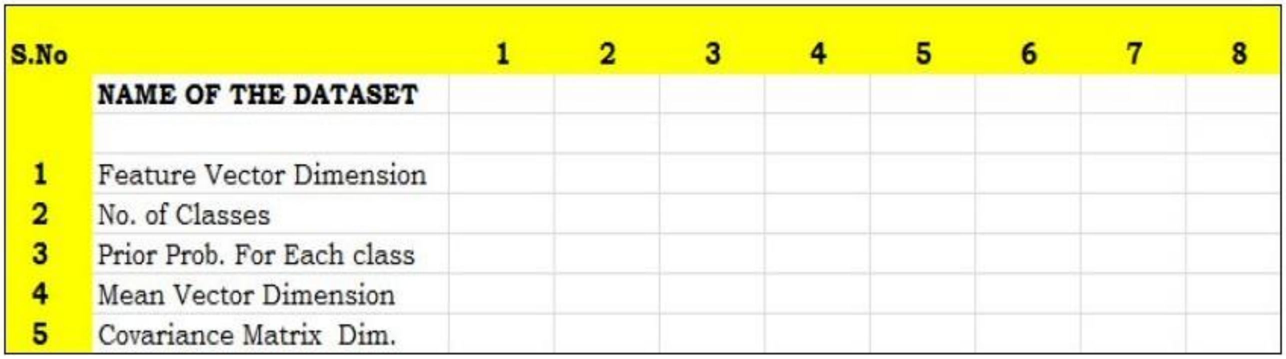
Here are the classes in the dataset, as well as 10 random images from each:

In both training as well as in test data each row stores a 32x32 color image. The first 1024 entries contain the red channel values, the next 1024 the green, and the final 1024 the blue. The image is stored in row-major order, so that the first 32 entries of the array are the red channel values of the first row of the image. When you load the training or testing data the corresponding labels will be automatically loaded.

Bayesian Classification:

Q1.

Go through the datasets uploaded in the zip file and complete the following table:



S.No		1	2	3	4	5	6	7	8
	NAME OF THE DATASET								
1	Feature Vector Dimension								
2	No. of Classes								
3	Prior Prob. For Each class								
4	Mean Vector Dimension								
5	Covariance Matrix Dim.								

Q.2

Select any one dataset. Then for each class select one feature and plot the 1D histogram i.e. p (x_k /w_i) for at least 3 classes where x_k is k^{th} feature of dataset and w_i represents i^{th} class . Now apply bayesian classification using the above likelihoods, you can experiment with different values of k. Repeat the same by selecting 2 features for at least 3 classes and plot the 2D histogram (you can use the inbuilt command for this).

Q.3

You need to perform Bayesian classification for the following dataset. Before starting divide the data of each class into 70% data as training and 30% for testing. Text File has 1500 data points. First 500 data points belong to class_1, next 500 to class_2 and last 500 to class_3. Perform Bayesian Classification for the following cases:

Case_1: Same Covariance Matrix for all the classes. Hint: Calculate Σ by considering all data points.

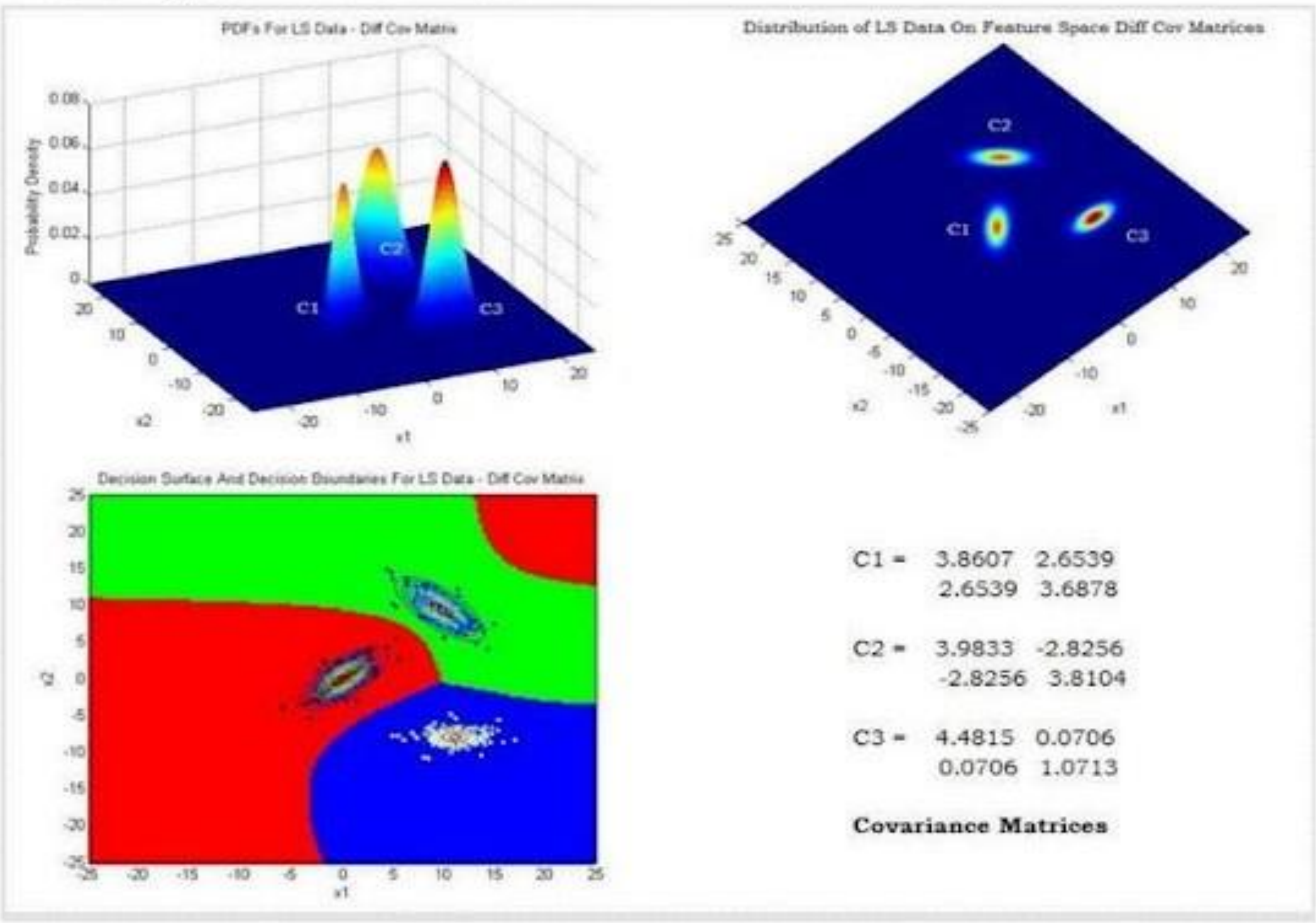
Case_2: Different Covariance Matrices. Hint: Calculate Σ1, Σ2 and Σ3 separately for each class.

Case_3: Diagonal Covariance Matrices. Hint: Make σ12= σ21=0 in covariance matrices generated in case_2.

Case_4: Apply whitening transform on covariance matrices obtained in case 2 and now use the transformed covariance matrices for applying Bayesian classification. (you may use inbuilt command for Eigen decomposition).

NOTE:

You have to plot eigen vectors for the covariance matrix and the Contours of equal probability on the feature space. You may use inbuilt functions like eig (Matlab) or numpy.linalg.eig (Python), quiver and contour.



$$C1 = \begin{matrix} 3.8607 & 2.6539 \\ 2.6539 & 3.6878 \end{matrix}$$

$$C2 = \begin{matrix} 3.9833 & -2.8256 \\ -2.8256 & 3.8104 \end{matrix}$$

$$C3 = \begin{matrix} 4.4815 & 0.0706 \\ 0.0706 & 1.0713 \end{matrix}$$