Chapter 11: Classification

Questions and Exercises

1. Explain what you understand by each of the following terms –
   i) Object (or pattern) class.
   ii) Feature (or pattern) vector.
   iii) Feature space.
   iv) Discriminant function.

   To illustrate your answer to each of these parts, you may use a simple example which may be real or hypothetical.

2. The table below summarises data on 2 features obtained from a given image containing 8 examples of 3 types of food.

CIRCULARITY			**NO.**	**LINE-FIT ERROR**		
Pine-nuts	lentils	Rice-grains		Pine-nuts	lentils	Rice-grains
0.925	0.955	0.966	1	2.53	2.48	2.84
0.918	0.95	0.957	2	2.62	2.55	2.93
0.924	0.94	0.963	3	2.75	2.53	2.97
0.915	0.955	0.957	4	2.86	2.7	3.10
0.916	0.947	0.944	5	2.97	2.74	3.14
0.92	0.936	0.955	6	3.05	2.75	3.28
0.914	0.944	0.944	7	3.08	2.9	3.28
0.919	0.932	0.945	8	3.19	2.94	3.35

   Calculate the equations defining the decision boundaries for a linear, minimum-mean-distance classifier and sketch the boundaries in a 2-D feature space.

3. In the context of image processing and classification, explain the following terms :–
    i) Feature vector.
    ii) Feature space
    iii) Discriminant function

4. Draw a block diagram or flowchart outlining the main sequential steps in the design of an automated classifier. Describe the purpose of each step and how, in general terms, we might attempt to refine the design and performance of a classifier system.

5. A study is carried out on the shape characteristics of male and female faces. A principal component analysis on a sample containing an equal number of 10 male and 10 female faces results in the following data being obtained.

MALE	***PC1***	***PC2***	***FEMALE***	***PC1***	***PC2***
1	8.9	9.5	1	-7.2	-19.4
2	3.8	6.7	2	3.9	-5.6
3	5.13	0.9	3	-7.1	-14.0
4	13.7	14.8	4	1.2	-8.8
5	2.8	1.8	5	-2.5	-12.7
6	4.9	3.9	6	-3.4	-14.0
7	-3.2	-3.4	7	-1.17	-12.2
8	-5.3	-7.5	8	4.8	-7.7
9	3.4	3.2	9	10.4	1.5
10	12.2	16.2	10	9.1	-1.5

Calculate a linear discriminant function which achieves the classification into two classes.

Two faces of unknown gender having PC1=2.5 and PC2=-2.5 and PC1=7.5 and PC2=5 are input to the classifier. To what classes are the faces assigned ?

6. What is the main difference between *supervised* and *unsupervised* classification procedures ?

7. Describe the main idea and steps in the k-means classification algorithm.

8. A sample of two different species of palm leaves is taken and measurements made on two characteristics of their shape - the length and form-factor. The results are given in the table below.

PALM A	LENGTH	FORMFACTOR	**PALM B**	LENGTH	FORMFACTOR
	3.47	5.72		7.48	6.9901
	4.73	6.0		6.99	6.2189
	7.28	4.88		6.72	6.2617
	4.94	6.33		8.27	7.2134
	6.84	8.46		8.86	5.72
	3.96	5.78		6.48	5.87
	5.01	4.98		7.10	4.73
	3.83	6.37		6.19	4.34
	5.88	5.08		7.68	5.30
	4.62	3.34		4.64	6.28

By first calculating mean or prototype feature vectors for the two classes, derive a linear discriminant function for a minimum distance classifier.
How many misclassifications occur ? Do you consider length and form-factor to be a good choice for the formation of the feature space ? State your reasons.

9. An analysis of 5 examples each from 2 image classes produces the 2-D feature vectors given in the table below -

Class 1	***F1***	***F2***	***Class 2***	***F1***	***F2***
1	8.9	9.5	1	-7.2	-19.4
2	3.8	6.7	2	3.9	-5.6
3	5.13	0.9	3	-7.1	-14.0
4	13.7	14.8	4	1.2	-8.8
5	2.8	1.8	5	-2.5	-12.7

10. Plot the data and make an informed decision as to the correct classification of the point (F1=0,F2=2.5). Explain the basis of your decision.