B.Sc. Engg. CSE 8th Semester 29 October 2018

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION **SUMMER SEMESTER, 2017-2018**

DURATION: 3 Hours **FULL MARKS: 150**

CSE 4835: Pattern Recognition

Programmable calculators are not allowed. Do not write anything on the question paper.

There are 8 (eight) questions. Answer any 6 (six) of them.

Figures in the right margin indicate marks.

---

1. a) Define the following terms: 4
   i. Class Model
   ii. Overfitting Problem
   iii. Decision Rule
   iv. Risk

   b) Consider a two-class problem with the following univariate Guassian class-conditional densities: $P(x\,|\,\omega_1) = N(0,1)$ and $P(x\,|\,\omega_2) = N(1,1)$.

   Assume $P(\omega_1) = 3/4$ and $P(\omega_2) = 1/4$.

   i. Calculate the threshold boundary value x^* which gives the probability of minimum error. 7

   ii. Is this classifier a linear machine? Justify your answer. 3

   iii. Assume $P(\omega_1) = P(\omega_2) = 1/2$. Find the minimum risk decision rule with loss functions as $\lambda_{11} = \lambda_{22} = 0, \lambda_{12} = 1, \lambda_{21} = 1/2$. Approximately show the effect on class regions. 8+3

2. a) Consider the three-category linear machine with discriminant functions :
   $g_i(x) = w_i^t x + w_{i0}, i = 1,2,3.$

   i. For the special case where x is two-dimensional and the threshold weights w_{i0} are zero, sketch the weight vectors with their tails at the origin, the three lines joining their heads, and the decision boundaries. 7

   ii. How does this sketch change when a constant vector c is added to each of the three weight vectors? 3

   b) Consider the following six data points: 15

   $$\omega_1 : \quad (1,5)^t, (2,9)^t, (-5,-3)^t$$
   $$\omega_2 : \quad (2,-3)^t, (-1,-4)^t, (0,2)^t$$

   Are they linearly separable? Derive the decision boundary equation to separate them. Show your detailed calculation up to three iterations. Use batch update with Perceptron criterion function. Assume any values if necessary.

3. a) Let random variable x have an exponential density: 10

   $$p(x\,|\,\theta) = \begin{cases} \theta e^{\theta x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and } \theta > 0.$$

Suppose that n samples $x_1, x_2, \ldots, x_n$ are drawn independently according to $p(x \mid \theta)$. Show that the maximum likelihood estimate of θ is given by:

$$\hat{\theta} = -\frac{1}{\frac{1}{n}\sum_{k=1}^{n} x_k}$$

5

b) Prove that the *minimum distance to class member* classifier is non-linear. 10

c) For the following sixteen samples in a one-dimensional problem:

$$D = \{0, 1, 3, 4.5, 5.5, 6.0, 6.5, 7.0, 7.2, 7.5, 8.0, 8.8, 9.2, 9.3, 11, 13\}$$

Give the values of the k-nearest neighbor estimate $p_n(x)$, for n=16 and $k_n = \sqrt{n}$, at x=*1.2*, x=*4.5*, x=*6*, x=*8*, and x=*10.3*.

4. a) What is the objective of Principal Component Analysis (PCA)? How does PCA address the issue of curse of dimensionality? 5

b) Multispectral data acquired from most remote sensors exhibit high interband correlations. Taking Landsat Multispectral Scanner (MSS) imagery as an example, band 1 (green) and band 2 (red) are highly correlated because of the relatively low reflectance of vegetation. Bands 3 and 4, the two infrared bands, are highly correlated because of the high reflectance of vegetation. Similarly, for the Thematic Mapper (TM) data, the first three visible bands (TMl, TM2, and TM3) are also highly correlated. Data processing with all of the spectral bands therefore involves a certain degree of redundancy. This increases the cost of data processing, especially when addressing change detection issues which involve images of more than one date. Devise a technique for detecting land-cover change with data reduction. 15

c) Suppose the covariance matrix **A** has trace value of 4, and two of its Eigen values are 1 and 0. Comment on the distribution pattern of the original samples in the 3D feature space from which that covariance matrix is computed. 5

5. a) Formulate the criterion function J of Linear Discriminant Analysis (LDA) for a two-class problem based on several factors which influence the design of that function. Extend your criterion function to support multiclass problem. 7+3

b) The classification of upper-limb movements based on surface electromyography (EMG) signals is an important issue in the control of assistive devices and rehabilitation systems. Contemporary research try to increase the efficiency of these multifunction EMG prostheses by increasing the number of movements recognized, which can directly increase the number of control commands. However, this leads to a need for increased information to be extracted from the EMG signals. There are two major ways used to increase the information derived from EMG recognition systems: obtaining information from different muscle positions and utilizing the information present in features of the signal. However, whilst increasing the number of EMG channels, EMG features yields a high dimensional feature vector. Design a pattern classification system for detecting five different movements with high level of efficiency in terms of both cost and accuracy. 15

6. A long time ago there was a village amidst hundreds of lakes. Two types of fish lived in the region, but only one type in each lake. These types of fish both looked exactly the same, smelled exactly the same when cooked, and had the exact same delicious taste - except one was poisonous and would kill any villager who ate it. The only other difference between the fish was their effect on the pH (acidity) of the lake they occupy. The pH for lakes occupied by the non-poisonous type of fish was distributed according to a Gaussian with unknown mean (μ) and variance (σ^2) and the pH for lakes occupied by the poisonous type was 25

distributed according to a different Gaussian with unknown mean (μ) and variance (σ^2). (Poisonous fish tended to cause slightly more acidic conditions). Naturally, the villagers turned to you for help.

Design a complete solution based on the knowledge you have obtained from this course.

7. a) What equations are used for classification in a support vector machine (SVM) and how? 5
   b) For a two-class classification problem, we use an SVM classifier and obtain the separating hyperplane as shown in Figure 1. We have marked four instances of the training data. Identify the point which will have the most impact on the shape of the boundary on its removal. Explain your choice. 7

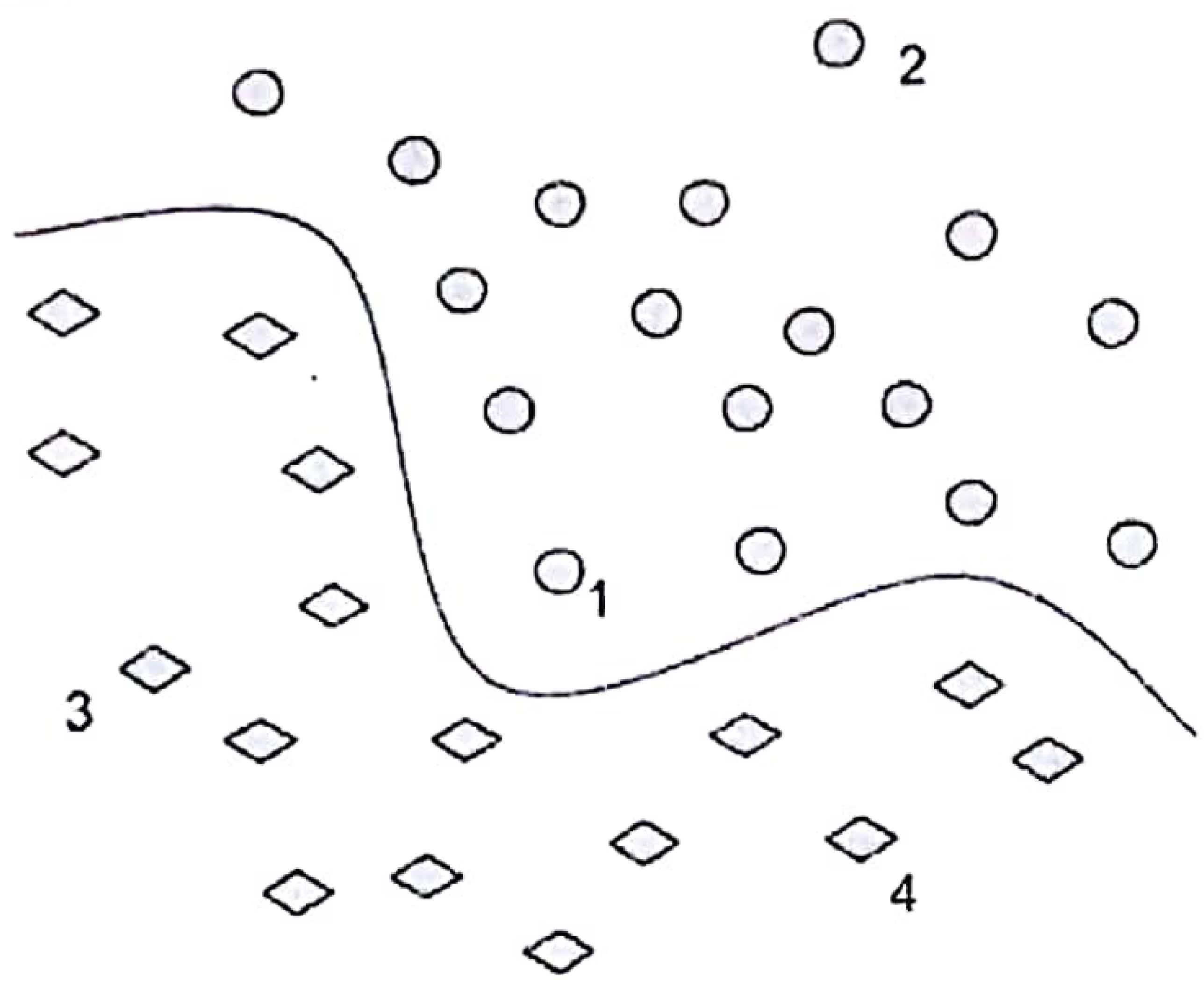

Figure 1. for 7(b)

   c) How does the kernel trick work in SVM? 5
   d) Consider the following optimization problem: 8

$$\min f(x) = x^2 + 1$$

subject to

$$g(x) = (x-2)(x-4) \leq 0$$

Find the extreme values using Lagrange Multipliers.

8. a) Draw a simple Neural Network to represent a generalized discriminant function $g_i(x) = w_i'x + w_{i0}$. 3
   b) In a feed-forward neural network (NN), the weights w_{ij} of the edges to the hidden nodes are adjusted by the following term. 10

$$\frac{\partial E}{\partial w_{ij}} = O_i \delta_j$$

$$\delta_j = O_j(1-O_j)\sum_{k \in K} \delta_k w_{jk}$$

Taking into consideration the usual meaning of the notations used, how did the back-propagation algorithm devise this adjustment factor?
   c) Define TPR and FPR. Draw the ROC curve with five different thresholds on the following data scores given in Table 1. Show all required calculations. 2+10

Table 1. for 8(c)

Sample No.	True Class	Score
1	P	0.90
2	P	0.80
3	N	0.70
4	P	0.65
5	P	0.50
6	P	0.53
7	N	0.47
8	N	0.43
9	P	0.42
10	N	0.40
11	P	0.37
12	N	0.31
13	P	0.26
14	N	0.22
15	N	0.19
16	N	0.15
17	P	0.12
18	N	0.11
19	P	0.04
20	N	0.01