B.Sc. Engg. CSE 8th Semester

21 August 2017 (Afternoon)

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

MID SEMESTER EXAMINATION — SUMMER SEMESTER, 2016-2017

DURATION: 1 Hour 30 Minutes — FULL MARKS: 75

CSE 4835: Pattern Recognition

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

There are **4 (four)** questions. Answer any **3 (three)** of them.

Figures in the right margin indicate marks.

---

1. a) Briefly explain why generalization issue should be considered rather than overfitting in classification. **5**

   b) What characteristics should an ideal feature extractor hold during extracting a feature vector representing a sample object? Explain each of them with examples **7**

   c) For a 'minimum distance to class member' classifier, find the weight vector w and bias w_0 of its discriminant function $g(x)$. Show that this piecewise linear classifier can be viewed in terms of a linear machine for classifying subclasses of patterns. **8+5**

2. a) Show that the distance from the hyperplane $g(x)=w^t x + w_0 = 0$ to the point x_a is $|g(x_a)|/\|w\|$. Also prove that the projection x_p on the hyperplane is given by **5+5**

$$x_p = x_a - \frac{|g(x_a)|}{\|w\|^2} w$$

   b) There are generally three ways to devise multicategory classifiers employing linear discriminant functions. Describe each of such designs along with their pros and cons. Use necessary illustrations. **9**

   c) Define the Perceptron criterion function J for finding the augmented weight vector of a linear discriminant function $g(x)$ in a two-class problem. Fit this criterion function into the steps of a basic Gradient Descent technique for getting solution vector. **1+5**

3. a) In many pattern classification problems, one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. **10**

$$\text{Let } \lambda(\alpha_i \mid \omega_j) = \begin{cases} 0 & i = j \quad i,j = 1,2,\ldots,c \\ \lambda_r & i = c+1 \\ \lambda_s & \text{otherwise} \end{cases}$$

   where, λ_r is the loss incurred for choosing the $(c+1)^{th}$ action that is rejection, and λ_s is the loss incurred for making a substitution error. Show that the minimum risk is obtained if we decide ω_i if $P(\omega_i|x) \geq P(\omega_j|x)$ for all j and if $P(\omega_i|x) \geq 1 - \lambda_r/\lambda_s$, and reject otherwise.

   b) Devise the decision rule for the Bayes classifier with minimum risk. Explain the effects of increasing the loss function λ_{21}. **5**

   c) Suppose, in a Bayes classifier the likelihood probability follows a normal distribution and all classes have the same covariance matrix but with different prior probabilities. Devise the equation of the decision boundary for each pair classes. **10**

4. a) How are the estimates from Maximum Likelihood Estimation (MLE) different or similar to the estimates of Bayes Parameter Estimation? 5

b) Use MLE method for estimating the unknown parameter $\theta = \{\mu, \sigma^2\}$ of a univariate Gaussian distribution. 8

c) In nonparametric density estimation techniques which conditions are required for the estimate $p_n(x)$ to converge to true $p(x)$? What do they assure? 3

d) Describe the general principal of Bayes Parameter Estimation technique for approximating the unknown parameters. 9