Engg. CSE 8th Semester

25 October 2017

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION

SUMMER SEMESTER, 2016-2017

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.

---

a) Define class model. How does a model change with the change of a classification problem but with the same input data? Explain with examples. 1+5

b) Astronomers have been cataloguing distant objects in the sky using long-exposure CCD images. The objects need to be labeled as star, galaxy or nebula. The data is highly noisy, and the images are also faint. Manual cataloguing can take decades to complete. To automate such task, describe the components of that pattern recognition system. 15

c) Compare between classification and regression. 4

a) Devise the decision rule for the Bayes classifier with minimum risk. Explain why this minimum risk classifier cannot always ensure minimum error. 1+4

b) In a particular 2-class hypothesis testing application, the conditional density for a scalar feature x given class ω_1

$$p(x \mid \omega_1) = k_1 \exp(-x^2/20)$$

Given class ω_2, the conditional density is

$$p(x \mid \omega_2) = k_2 \exp(-(x-6)^2/12)$$

i. Find k_1 and k_2. 2

ii. Assume that the prior probabilities of the two classes are equal, and that the cost for choosing correctly is zero. If the costs for choosing incorrectly are $\lambda_{12} = \sqrt{3}$ and $\lambda_{21} = \sqrt{5}$ (where λ_{ij} corresponds to predicting class i when it belongs to class j), what is the expression for the conditional risk? 6

iii. Find the equation of the decision boundary. Also find the decision regions which minimize the Bayes risk 12

a) Consider the three-dimensional normal distribution $p(x|\omega)$ with mean μ and covariance matrix Σ where 15

$$\mu = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 4 \\ 0 & 4 & 5 \end{bmatrix}$$

Compute all eigen-vectors and eigenvalues. Note: One of the eigenvalues is 9 and one of the eigenvectors is $[1\ 0\ 0]^T$

b) Find the probability density value at the point $x_0 = [5\ 6\ 3]^T$. 3

c) Construct an orthonormal transformation $y = \Phi^T x$, where the matrix Φ contains all eigenvectors. Show that for orthonormal transformations, Euclidean distances are preserved, i.e., $\|y\|^2 = \|x\|^2$. 7

4. a) Discuss qualitatively if samples from two categories are distinct (i.e., no feature point is labeled by both categories), there always exists a nonlinear mapping to a higher dimension that leaves the points linearly separable.

   b) Consider the following six data points:

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

   Are they linearly separable? Derive the decision boundary equation to separate them.

5. a) Define covariance. If the value of covariance between X and Y variables becomes zero, do you think X and Y are somehow dependent to each other? Justify your answer.

   b) Say there is a Pattern Recognition (PR) course offered at your university. The course instructors have observed that students get the most out of it if they are good at Math or Statistics. Over time, they have recorded the scores of the enrolled students in these subjects. Also, for each of these students, they have a label depicting their performance in the PR course: "Good" or "Bad", as shown in Figure 1. Upper right cluster shows "Good" and lower cluster as "Bad".

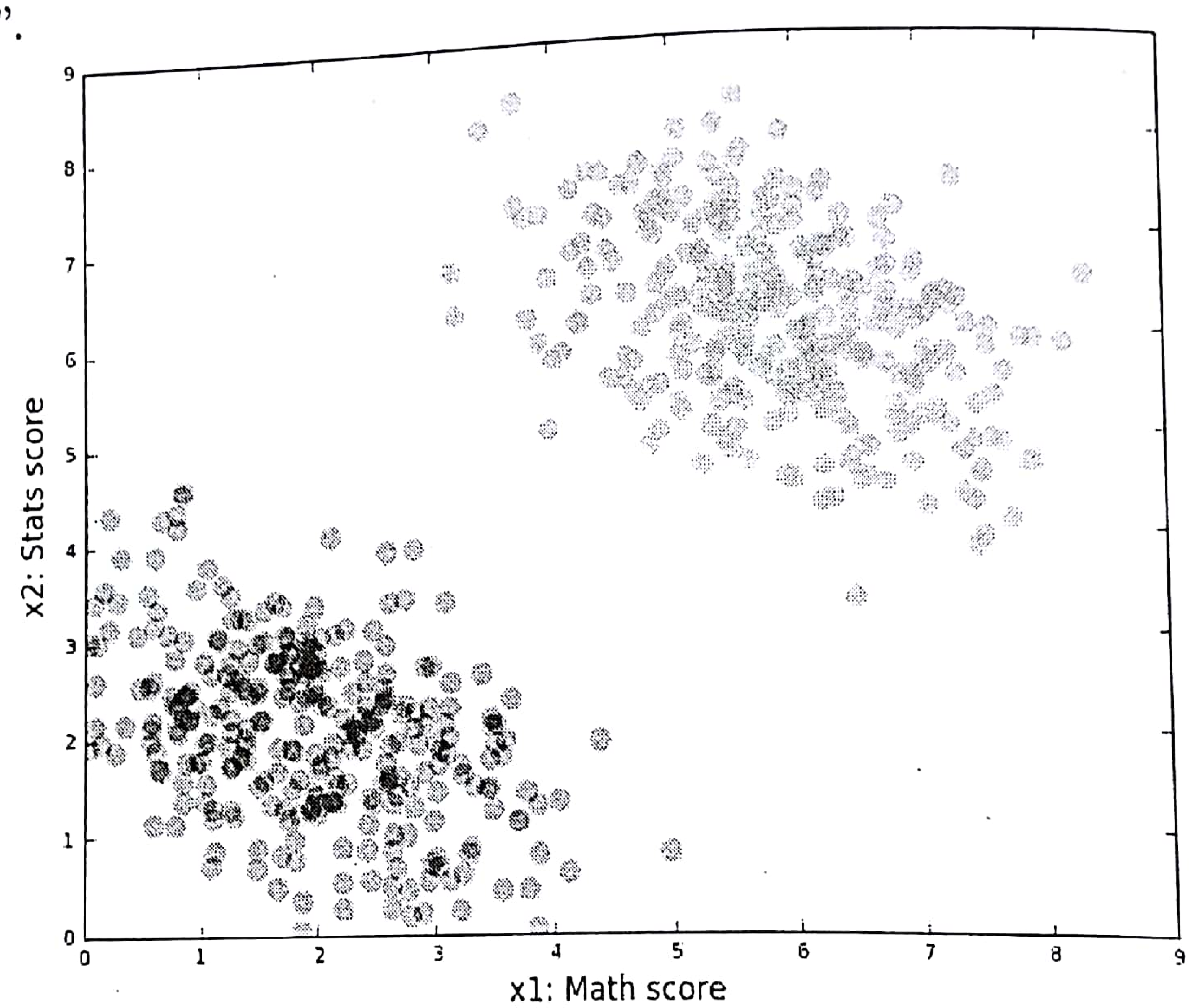

Figure 1

   Now they want to determine the relationship between Math and Statistics scores and the performance in the PR course. When a student requests enrollment, instructors would ask her to supply her Math and Statistics scores. Based on the data they already have, they would like to make an informal guess about his/her performance in the PR course. Design a linear classifier that would help them in making such guess and decide whether to let him/her to enroll or not. Remember that the worth of a classifier is not in how well it separates the training data. We eventually want it to classify yet-unseen data points (known as test data).

   c) Given

$$f(x,y) = y^2 - x$$

   Subject to

$$g_1(x,y) = 2x^2 + 2xy + y^2 - 1 = 0$$

   Find the extreme values using Lagrange Multipliers.

10 a) Suppose two discriminant functions g_1 and g_2 are defined based on the Bayes formula. Provide the equation of their decision boundary. What happens if the prior probabilities change their values, e.g., one class has higher prior probability than the other class. 3+4

15 b) Suppose the discriminant functions in Question 6.(a) are defined on a multivariate normal density function in the *d* dimensional feature space. If feature elements in all classes are statistically independent and have the same variance, then derive the general form of the discriminant function *g(x)*. Which conditions will convert this classifier to the 'minimum distance to class mean' classifier? 5+3

c) Show that 'minimum distance to class member' classifier can be viewed in terms of a linear machine for classifying subclasses of patterns. 10

5 a) Define the following terms along with their mathematical representation: 5

i. Recall or True Positive Rate (TRP)
ii. Precision
iii. False Positive Rate (FPR)
iv. Specificity
v. F-measure

10 b) Suppose the following data set is given, where 20 sample instances are given along with their true class label: *P* for positive class and *N* for negative class. If the scores are to be considered for classification into two classes, then calculate the TPR and FPR at five different thresholds on score: **0.75, 0.57, 0.53, 0.51** and **0.50**. Show the confusion matrix for each threshold value. 10

#	Class	Score
1	P	0.9
2	P	0.8
3	N	0.7
4	P	0.6
5	P	0.55
6	P	0.51
7	N	0.49
8	N	0.43
9	P	0.42
10	N	0.39
11	P	0.33
12	N	0.31
13	P	0.23
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

c) Suppose x have a uniform density 10

$$p(x|\theta) \equiv U(0,\theta) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{otherwise} \end{cases}.$$

If *n* samples D={$x_1, x_2, \ldots, x_n$} are drawn independently according to p(x| θ), then show that the maximum likelihood estimate for θ is *max*[D], i.e., the value of the maximum element in D.

8. A renowned factory produces very expensive and high quality chip rings that their qualities are measured in term of curvature and diameter. Result of quality control by experts is given in the table below:

Table 1: Curvature Diameter Quality Control Result

Curvature	Diameter	Result
2.95	6.63	Passed
2.53	7.79	Passed
3.57	5.65	Passed
3.16	5.47	Passed
2.58	4.46	Not Passed
2.16	6.22	Not Passed
3.27	3.52	Not Passed

As a consultant to the factory, you get a task to set up the criteria for automatic quality control. Then, the manager of the factory wants to test your criteria upon new type of chip rings that even the human experts argued to each other. The new chip rings have curvature 2.81 and diameter 5.46. Classify the new chip rings to either Passed or Not Passed. You have to use Linear Discriminant Analysis for feature transformation.