

Roll No.: _____

Amrita Vishwa Vidyapeetham
Amrita School of Engineering, Coimbatore
B.Tech. Degree Examinations – April 2019
Sixth Semester
Computer Science and Engineering
15CSE361 Pattern Recognition

Time: Three hours

Maximum: 100 Marks

CO	Course Outcomes
CO01	Understand the working principles of pattern recognition system and algorithms
CO02	Apply statistical methods for decision making
CO03	Understand non parametric decision making system
CO04	Apply and evaluate Non-metric approach for real world problems
CO05	Apply and analyze unsupervised learning methods for real world problems.

Answer all Questions

(10 x 10 = 100 Marks)

1.
 - (a) An x-ray source emits 10 photons per second on the average, with a poisson distribution. What is the probability that exactly 10 photons will be emitted in a given second? [5][CO01]
 - (b) The phone lines to an airline reservation system are occupied 40% of the time. Assume that the events that the lines are occupied on successive calls are independent. Assume that 10 calls are placed to the airline.
 - (i) What is the probability that for exactly three calls the lines are occupied? [2][CO01]
 - (ii) What is the probability that for at least one call the lines are not occupied? [3][CO01]
2. Distance matrix for 5 samples is given below. Perform agglomerative clustering using complete linkage algorithm. [10] [CO05]

	1	2	3	4	5
1	0				
2	9	0			
3	3	7	0		
4	6	5	9	0	
5	11	10	2	8	0

3.
 - (a) A classifier has a 30 percent error rate. What is the probability that exactly three errors will be made in classifying 10 samples? [5] [CO04]
 - (b) Use a symmetric triangular window with a base of 4 to estimate the density at $x=2$, given samples at 0,1,1,2,3,5,8 and 9. [5] [CO04]

4.

- (a) In Bayes Decision Theory, what does the posterior probability capture? [2] [CO02]
- (b) How likelihood ratio is estimated in the Bayes Theorem? [2] [CO02]
- (c) In Bayes Decision Making, define the risk (i.e. expected loss) of a strategy (i.e. of a decision rule). [2] [CO02]
- (d) Describe limitations of Bayes decision making. [2] [CO02]
- (e) How do you estimate the density of the data that belongs to multiple features and normally distributed. [2] [CO02]

5. Consider the following data set consisting of the scores of two variables on each of seven individuals. [10] [CO05]

Subject	A	B
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

Perform a partitional clustering using the Forgy's algorithm by considering following mean values as initial seeds. Use Euclidean distance measure for finding the distance between samples.

	Individual	Mean Vector (centroid)
Group 1	1	(1.0, 1.0)
Group 2	4	(5.0, 7.0)

- 6. Two samples from class A are located at (0,0) and (1,0). Two samples from class B are located at (2,0) and (2,1). Find a linear discriminant function to equal to 1 for members of class A and -1 for members of class B. What set of three weights minimizes the squared error between the desired and the actual values of the discriminant function at the four samples? What is $D(x,y)$? Show the data and the lines $D=0$, $D=1$, and $D = -1$ on a graph. [10] [CO03]
- 7. Explain the steps involved in the Iterative Self-Organizing Data Analysis Technique Algorithm along with its merging and splitting methods. [10] [CO04]

8. Consider the following data set details the profitability of the software based on Age, Competition and Type parameters. [10] [CO05]

Age	Competition	Type	Profit
Old	Yes	Software	Down
Old	No	Software	Down
Old	No	Hardware	Down
Mid	Yes	Software	Down
Mid	Yes	Hardware	Down
Mid	No	Hardware	Up
Mid	No	Software	Up
New	Yes	Software	Up
New	No	Hardware	Up
New	No	Software	Up

Build a decision tree based on the above observations using Classification and Regression trees. Classify the following sample using your algorithm:

Age = Old, Competition = Yes, Type = Hardware.

9. Given the following statistics construct a confusion matrix and find, what is the probability that men has throat cancer if he has a positive tomography test result?
- 1% of Men have throat cancer
 - 80% of tomography tests detect throat cancer when it is there
 - 9.6% of tomography tests detect throat cancer when it's not there (and therefore 90.4% correctly return a negative result). [10] [CO02]
10. For class A, the probabilities that binary features x, y, and z are present are $\frac{1}{3}$, $\frac{1}{4}$, and $\frac{1}{5}$ respectively. Whereas for class B, these probabilities are $\frac{1}{4}$, $\frac{1}{5}$, and $\frac{1}{6}$. Suppose $P(A) = \frac{1}{3}$ and $P(B) = \frac{2}{3}$. What is the probability that the class is A, given that features x and y are present and z is absent? Assume the features are independent within each class. [10] [CO03]

* * * * *