Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

V SEMESTER B.TECH. (INFORMATION TECHNOLOGY / COMPUTER AND COMMUNICATION EGINEERING)

MAKEUP EXAMINATIONS, JANUARY 2019

SUBJECT: PROGRAM ELECTIVE - I PATTERN RECOGNITION [ICT 4020]

REVISED CREDIT SYSTEM
(01/ 01/ 2019)

Time: 3 Hours   MAX. MARKS: 50

1A. List the significances of cross validation techniques in pattern classification. Also, explain various cross validation techniques. **5**

1B. Draw the HMM for the data given in Table Q.1B and decode the outcome 'Happy-Happy-Sad' using Viterbi algorithm. Assume that at any random time instance, the state probabilities are P(Rainy) = 0.45, P(Cloudy) = 0.35 and P(Sunny = 0.2) respectively.

Table Q.1B

Transition Matrix		Present State			Emission Matrix	Happy	Sad
		Rainy	Cloudy	Sunny			
Previous State	Rainy	0.4	0.3	0.3	Rainy	0.4	0.6
	Cloudy	0.2	0.6	0.2	Cloudy	0.7	0.3
	Sunny	0.1	0.1	0.8	Sunny	1.0	0

3

1C. Explain the methods of improving the performance of classifier using ensemble techniques. **2**

2A. Consider the grammar $G = \langle T, N, R, S \rangle$ where T = {x, y}, N = {S}, R = {$S \rightarrow xSx \mid ySy \mid x \mid y \mid \epsilon$} and ϵ denotes empty string. Assume each production rule is associated with probability factor 0.2. Generate the palindrome sequence {xyx, xxyyxx, xxyxxyxx} using SCFG and compute the total probability. **5**

2B. Explain various feature subset selection algorithms used in pattern recognition. **3**

2C. Explain the influence of degree of dimensionality on the decision surface. **2**

3A. Predict the rent using KNN classifier for the data given in Table Q.3A. Assume k =3.

Table Q.3A.

Area (sq.ft)	1700	2600	1000	2000	2500	1500
Number_of_bedrooms	2	3	1	2	3	2
Floor_Number	3	5	2	4	3	2
Rent	7000	10000	3000	8500	11000	?

5

3B. How the following factors become a problem in pattern recognition? Explain using relevant examples.

   i.   Segmentation

   ii.   Mereology

   iii.   Projective Distortion   3

3C. Compute the dissimilarity matrix using Manhattan and Minkowski (h=3) distance metrics for the following dataset

$$\begin{bmatrix} 24 & 15 & 36 \\ 12 & 44 & 08 \\ 60 & 27 & 53 \\ 06 & 30 & 17 \end{bmatrix}$$

2

4A. Consider that the variation in stock price of two companies 'A' and 'B' is recorded in Table Q.4A for 10 minutes in a day. Assume the missing values are replaced by the corresponding average values. Use PCA and estimate the principal component (PC1).

Table Q.4A

Company A	0.44	0.98	-0.87	0.71		0.08	-1.2	-0.1	2.2	0.59
Company B	0.06		1.58	0.32	-0.80	0.53		-1.7	-2.3	0.75

5

4B. Explain Parzen Window Classifier using relevant equations.   3

4C. Compute the conditional risk involved for a three class category classification, using zero-one loss function, given $P(w1/x) = 0.3$, $P(w2/x) = 0.5$, $P(w3/x) = 0.2$.   2

5A. Explain the optimization algorithm that can be used to minimize the error of linear discriminant function.   5

5B. Explain the significance of the parameters mean and covariance matrix, to solve the problems of pattern recognition.   3

5C. Consider Table Q.5C, which presents product sales of two companies All_Electronics and High_Tech. Find the covariance between the two companies.

Table Q.5C

Time point	T1	T2	T3	T4	T5
All_Electronics	6	5	4	3	2
High_Tech	20	10	14	5	5

2