M.E. ELECTRONICS AND TELE-COMMMUNICATION ENGINEERING
FIRST YEAR SECOND SEMESTER-2019

PATTERN RECOGNITION **Time: 3 hours** **Full Marks: 100**

Answer **Q. 1** and **any five** from **the rest.**

1. Answer the following questions: **5 x 3**
   a) If x is uniformly distributed in [1, 10], find the probability that x lies between 0.2 and 0.5.
   b) Define a positive semi-definite matrix and prove that it has non-negative eigenvalues.
   c) Explain the importance of smooth kernels in the context of non-parametric density estimation.
   d) What are support vectors? What is their importance?
   e) Distinguish between hierarchical and partitional clustering.

2. a) Discuss the basic building blocks of a pattern recognition system. **3**
   b) For a random variable X, prove that $\sigma_X^2 = E(X^2) - (\mu_X)^2$, where, $E(.)$ is the expectation operator. **4**
   c) Find the variance of a uniform random variable x in $[a, b]$. **3**
   d) What is meant by maximum likelihood estimate? Find the maximum likelihood estimate for μ in a normal distribution. **2+5**

3. a) State Bayes' Theorem. Explain how it can be used for parametric decision making. **2+3**
   b) Four samples from class A have values x = 2, -2, 1, -1 while four samples from class B have values 1, 2, 2 and 3. Assuming the classes are normally distributed and have equal prior probabilities, estimate the probability that an unknown sample at x = 0.5 came from class A. **6**
   c) Consider a two-class problem where the reward for correctly classifying a sample from class A is Rs. 6 and that from class B is Rs. 8. The penalty for misclassifying a sample from A is Rs. 15 and that from B is Rs. 30. For a sample with feature value x, $P(A|x)$ = 0.6 and $P(B|x)$ = 0.4. Find the maximum likely class as well as the least risk class for this sample. **6**

4. a) Discuss the basic structure of the Hidden Markov Models (HMM). Clearly explain what is hidden in HMM. **4+1**
   b) Explain the decoding problem in HMM. Provide a computationally efficient solution to this problem and justify the efficiency of your solution. **2+5**
   c) Obtain the overall probability of classification error using a model based error estimation strategy. **5**

5. a) Differentiate between parametric and non-parametric density estimation techniques. Show that p(x) = k/NV, as a general formulation of the non-parametric density estimation problem (symbols have their usual meanings). **2+5**
   b) How Parzen window can be used for non-parametric density estimation? Given dataset $X = \{4, 5, 6, 10, 10, 12, 14, 15, 16\}$, estimate the density of $p(x)$ using Parzen window at $x = 11$ using window sizes of 4 and 6. **2+4**
   c) Discuss how smooth kernel functions can improve density estimations as compared to that of Parzen windows. **4**

6. a) Obtain a bound for the nearest neighbor error rate. **5**
   b) A sample from class A is located at $(x, y, z) = (1, 2, 3)$, a sample from class B is located at (7, 4, 5) and a sample from class C is located at (6, 2, 1). How can you classify a sample (3, 4, 5) using a suitable algorithm and Euclidean distance? **4**
   c) Analyze the importance of Wolfe dual in the context of SVM. **5**
   d) Consider the following mapping: $x \in R^2 \rightarrow y \in R^3$ where **3**

$$y = \begin{bmatrix} x_1^2 \\ \sqrt{2}\, x_1 x_2 \\ x_2^2 \end{bmatrix}$$

   Show that $y_i^T y_j = (x_i^T x_j)^2$.

7. a) Explain Karhunen-Loeve transform of a vector. Why it is important for pattern recognition? **4+1**
   b) Show how Karhunen-Loeve transform can lead to dimensionality reduction. Justify the name PCA in this context. **5+1**
   c) Argue that the optimum hyperplane classifier of a SVM is unique. **6**

8. a) Consider a clustering problem where the clusters can be of arbitrary shapes. Discuss a suitable strategy to solve this problem. **5**

b) Perform a hierarchical clustering of the following data using the average linkage algorithm and city block distance. Show the distance matrices and the dendogram. **8**

Sample	x	y
1	0.0	0.0
2	0.5	0.0
3	0.0	2.0
4	2.0	2.0
5	2.5	8.0
6	6.0	3.0
7	7.0	3.0

c) Compare and contrast K-means and Forgy's algorithms. **4**