

IT-307-PATTERN RECOGNITION

Time: 03 Hours

Max. Marks: 40

Note: Attempt any *FIVE* questions.

Assume suitable missing data, if any.

Q.No.1

[8]

[a] Explain the role of Bayes theorem in decision making with the help of relevant Venn diagrams.

[b] Consider three firms manufacture smart phones and let the share of manufacture of these firms F1, F2, and F3 are 20%, 30% and 50%, respectively. Let these firms produce 2%, 3% and 5% defective smart phones, respectively. For an arbitrarily picked up defective smart phone, what is the probability that it is manufactured by F1, F2, or F3?

Q.No.2

[8]

[a] Consider a monovariate two class problem. Prior probabilities of the two classes are 0.8 and 0.2 respectively. Their mean and variances are 26, 22 and 2, 3 respectively. Find out a suitable decision boundary to partition the feature space. (Assume normal distribution of samples in each of the classes).

[b] Explain a single layer perceptron model which can classify a pattern (take own example) effectively. Explain the algorithm of nearest neighbor classifier. How do you measure the performance of nearest neighbor classifier? List out the limitations of this algorithm.

Q.No. 3

[8]

[a] Differentiate between Single linkage hierarchical clustering algorithm and Complete Linkage hierarchical clustering algorithms. How do you optimize classifier performance using these two algorithms?

[b] Consider a training set as given in the following table:

Name	Gender	Height in meters	Class
Arpit	Male	1.6	Short
Syamala	Female	2.0	Tall
Kanika	Female	1.9	Medium
Abhinav	Male	1.7	Medium
Ramaiah	Male	1.6	Short
Dileep	Male	2.2	Tall
Tarun	Male	2.1	Tall
Surya	Male	1.5	Short
Ambika	Female	1.8	Medium
Latha	Female	1.4	Short

Develop a K-nearest Neighbour algorithms, which can classify the instance (Shikha, Female, 1.95), the value of 'K' can be taken as 5.

Q.No. 4

[8]

- [a] Explain the construction of perceptron networks for linearly as well as nonlinearly separable data. Give a case study for each type of data and draw the network models with step by step justification. (Assume suitable weights and biasing parameter).
- [b] How root node influences the Minimal Spanning tree clustering algorithm? For a significantly large amount of data and multi-dimensional datasets, how this algorithm performs over other clustering techniques. (Compare with at least two other clustering algorithms).

Q.No. 5

[8]

- [a] The data set given as $X_1=[1,2]^T$, $X_2=[2,2]^T$, $X_3=[3,2]^T$, $X_4=[3,1]^T$
Compute the mean and covariance of data vector X.
- [b] Explain the role of the Covariance matrix in determining principal components using the PCA algorithm.

Q.No. 6

[8]

- [a] Explain the construction of suitable Hyperplane for two linearly separable classes using Support Vectors. Discuss all the constraints in optimizing hyperplane with necessary mathematical treatment.
- [b] What are the challenges of unsupervised learning over supervised Learning?