

INDIAN STATISTICAL INSTITUTE

(Pattern Recognition)

Back Paper Final EXAMINATION (2026)

(MSQMS, SQC)

Duration: 180 minutes

Maximum Marks: 70

(ANSWER ALL)

1. Describe the K-medoid and DBSCAN clustering algorithms. [10]
2. Compare the Advantages and disadvantages of K-medoid and DBSCAN algorithms. [10]
3. Differentiate between the principal component analysis (PCA) and linear discriminant analysis? Describe the step-wise process of finding the PCA of a data set. [6+4]
4.
 - a. What is ROC curve? Describe the method of finding and interpretation of this curve. [5]
 - b. Define the relationship between Euclidean, normalised Euclidean and Mahalanobis distance with equations. [5]
5. If X is a data sets with 9 samples and 2 features each. Find the Euclidean and Mahalanobis distances between 2nd and 5th samples of X. Explain the reasons for the different distance values. [10]
$$XX = \begin{bmatrix} 4 & 5 \\ 3 & 2 \\ 3 & 2 \\ 1 & 6 \\ 0 & 2 \\ 5 & 7 \\ 7 & 9 \\ 3 & 7 \\ 2 & 1 \end{bmatrix};$$
6. What is Bayes decision rule for classification of a data set. Discuss the advantages and disadvantages of this rule. [10]
7.
 - a. List the common properties of a distance METRIC. [2]
 - b. Describe the whole steps of operation of KNN classification model. Discuss the advantages and disadvantages of this model. [4+4]

=====**End of Question Paper**=====