

[Time:2.30 Hrs]

[Marks:75]

- Please check whether you have got the right question paper.**
- N.B:**
1. All question are compulsory.
 2. Figures to the right indicate full marks.

- Q.1 Attempt any Three of the following: 15**
- a. Explain clustering algorithms.
 - b. Explain the Partitioning Around Medoids algorithm used in K-Medoids.
 - c. What is the CLARA algorithm?
 - d. Explain the DBSCAN algorithm.
 - e. Describe the K-Means clustering algorithm.
 - f. What are the different linkage criteria used in AGNES?
- Q.2 Attempt any Three of the following: 15**
- a. What are the common challenges encountered in classification tasks?
 - b. Discuss the role of the parameter 'k' in the K-NN algorithm.
 - c. What is a k-Dimensional Tree (k-D Tree), and how is it constructed?
 - d. How do rule-based classifiers differ from decision trees?
 - e. What is Bayesian classification, and how does it apply Bayes' theorem?
 - f. How can a confusion matrix be used to calculate accuracy and precision?
- Q.3 Attempt any Three of the following: 15**
- a. Define eigenvalues and eigenvectors in the context of linear algebra.
 - b. What is CUR Decomposition?
 - c. State and explain the eigenvalue equation: $Av = \lambda v$.
 - d. Write a short note on singular values in SVD?
 - e. What are the advantages of using PCA for dimensionality reduction?
 - f. What is Principal Component Analysis (PCA), and why is it used in data analysis?
- Q.4 Attempt any Three of the following: 15**
- a. Explain the mathematical formulation of the PageRank algorithm.
 - b. Define link spam and its impact on search engine rankings.
 - c. What is content-based filtering in recommendation systems?
 - d. Explain the difference between content-based and collaborative filtering approaches.
 - e. How can Topic-Sensitive PageRank be applied to improve search engine results?
 - f. What are the challenges associated with collaborative filtering?
- Q.5 Attempt any Three of the following: 15**
- a. What is the difference between distance and similarity in clustering?
 - b. What is the DIANA algorithm?
 - c. What is a confusion matrix, and how is it used to evaluate classification models?
 - d. What is the role of the covariance matrix in PCA?
 - e. What are the limitations of dimensionality reduction techniques?
 - f. How does the damping factor influence the PageRank computation?

*****END*****