

Time :2 ½ Hours**Marks: 75**

- N.B.** 1) All questions are compulsory.
 2) Figures to the right indicate marks.
 3) Illustration, I depth answers and diagram will be appreciated.
 4) Mixing of sub-questions is not allowed.
 5) Each question carries 5 marks.

Q.1 Attempt Any Three of the following.**15 Marks**

- What is partitioning? Explain the term Centroid.
- List and explain working of K-Medoid.
- Write a note on CLARA. List its advantages and disadvantages.
- Explain K-Mean Algorithm? Write Advantage and Disadvantages.
- What is DBSCAN? Explain the fundamentals of DBSCAN.
- Distinguish between agglomerative and divisive algorithm.

Q.2 Attempt Any Three of the following.**15 Marks**

- What is classification? What are the challenges in classification? State the ways to overcome these challenges.
- Explain how support vector machine are used for classification in Big Data application. Give suitable example.
- Explain in detail CHAID.
- What is Regression? Explain different types of regression.
- What is support vector machine? Explain in detail different types of SVM.
- Explain Confusion matrix in detail.

Q.3 Attempt Any Three of the following.**15 Marks**

- Write a short note on Collaborative Filtering and its use in big data.
- Describe the dimensionality using Singular-Value decomposition.
- Explain the architecture of a spam farm.
- Describe Spam Mass.
- How will you represent transition matrices?
- Define the term SVD. Discuss its working.

Q4. Attempt Any Three of the following.**15 Marks**

- Explain how PageRank iterations can be implemented using MapReduce.
- State the steps to follow to integrate topic-sensitive Page Rank into a search engine.
- Write a short note on Content- Based Recommendation.
- Explain the Principal Component Analysis.
- Write a note on Singular value decomposition.
- Describe the dimensionality reduction using SVD.

Q5. Attempt Any Three of the following.**15 Marks**

- Briefly explain the k nearest neighbour algorithm .
- Describe in brief topic sensitive Page-Rank.
- What is ROC Curve? What is the significance of ROC Curve? What does the shape of curve indicate?
- Explain the Hierarchical Algorithms: Agglomerative (AGNES), Divisive (DIANA)
- Explain with an example.
 - K-Means
 - K-Medoids
- Briefly explain working of early search engine. State the issues related to these search engines. Explain how Google tries to overcome it.
