Time: 2 ½ Hrs **[Total Marks: 75]**

N.B

1) All questions are compulsory
2) Figures to right indicate marks
3) Illustration, depth answers and diagrams will be appreciated.
4) Mixing of sub-questions are not allowed.
5) Each question carries 5 Marks.

Q 1. Attempt any three of the following **[15]**

a) Explain 5 'V' characteristics of big data.
b) What are the modern data analytic tools? Explain in brief.
c) What is ARIMA Model? Explain.
d) What is Neural Network? List any five application of NN.
e) Explain the membership function with example.
f) What is need of Principal components analysis in Neural Network?

Q 2. Attempt any three of the following **[15]**

a) The distributed systems shall meet the some requirements. List them.
b) Explain the Elements of Hadoop Ecosystem.
c) List relational-algebra operations in brief.
d) Explain sorting and searching mathematical algorithms with MapReduce.
e) Distributed File System? How is DFS extended in the Hadoop Distributed File System?
f) Write a program to implement the matrix-multiplication algorithm using any one programming language.

Q 3. Attempt any three of the following **[15]**

a) Explain Jaccard Similarity of Sets.
b) Eplain the term K-shingles.
c) How to perform Matrix Representation of Sets?
d) Explain a) Jaccard Distance b) Cosine Distance
e) Explain the three conditions that we need for a family of functions.
f) List the applications of locality-sensitive hashing.

Q 4. Attempt any three of the following **[15]**

a) Explain the concept stream data model with suitable diagram.
b) Explain the Issues in Stream Processing.
c) What is The Bloom Filter? Explain the purpose of Bloom Filter.
d) How to deal with Infinite Streams?
e) What is Stream Queries? Explain two ways that queries get asked about streams.
f) How to perform selection, or filtering on streams?

Q 5. Attempt any three of the following **[15]**

a) Explain the term: Simulated Annealing
b) What is Apache Hadoop? What are the characteristics of a Distributed File System?
c) What is the role of mapper function and combiner function in MapReduce?
d) Compute the Jaccard similarities of each pair of the following three sets: {1, 2, 3, 4}, {2, 3, 5, 7}, and {2, 4, 6}.
e) List the methods for high degrees of similarity.
f) List DGIM Conditions.

****