Roll No. | | | | | | | | | | | | | **Total No. of Pages: 02**

Total No. of Questions: 09

M.Sc.(IT) (2015 Onwards) (Sem. – 4)

DATA WAREHOSUING AND MINING

M Code: 74116

Subject Code: MSIT-404

Paper ID: [74116]

Time: 3 Hrs. **Max. Marks: 60**

INSTRUCTIONS TO CANDIDATES:

1. **SECTIONS-A, B, C & D contains TWO questions each carrying TEN marks and students have to attempt any ONE question from each SECTION.**
2. **SECTION-E is COMPULSORY consisting of TEN questions carrying TWENTY marks in all.**
3. **Use of non-programmable scientific calculator is allowed.**

SECTION A

1. a) Explain the use and importance of different kinds of relations for spatial systems. (5)

   b) What are various steps involved in the process for Data Pre-Processing? Explain. (5)

2. Explain different types of MDD Model with pictorial representation and suitable examples. (10)

SECTION B

3. Write note on:

   a) Temporality Data Types. (5)

   b) Temporal Hierarchies. (5)

4. a) Why there is a need for Temporal Extension of the Multi Dimensional Model? Explain. (5)

   b) Explain Conceptual Model for Temporal Data Warehouses. (5)

SECTION C

5. a) What are Bayesian classifiers? With an example, describe how to predict a class label using Naive Bayesian classification. (5)

   b) What is backpropogation? Describe backpropogation algorithm. (5)

6. a) Differentiate between Fuzzy Set and Rough Set approach with appropriate examples. (5)

   b) What is Pattern Classification? Describe about basic measures for text classification. (5)

SECTION D

7. a) Describe different types of variables used in cluster analysis with examples. (5)

   b) Briefly explain the hierarchical clustering. (5)

8. Explain the following with respect to DBSCAN:

   a) Algorithm (5)

   b) Strengths and weaknesses (5)

SECTION E

9. Give short answers of the following:

   a) Distinguish between Classification and Prediction.

   b) What is Cross Validation?

   c) List the major issues in data mining.

   d) What is the difference between OLTP and Data Warehouse?

   e) What is the difference between K-nearest Neighbor and Case based Reasoning?

   f) Discuss spatial objects.

   g) What is Multiple Regression?

   h) Discuss K-mean with respect to Clustering.

   i) What is the principle behind use of Genetic Algorithm?

   j) "Big data and Data Warehouses are interconnected" Justify the statement.