Time : Three Hours]

[Maximum Marks : 79

INSTRUCTIONS TO CANDIDATES

(1) All questions carry marks as indicated.

(2) Solve Question No. 1 OR Question No. 2.

(3) Solve Question No. 3 OR Question No. 4.

(4) Solve Question No. 5 OR Question No. 6.

(5) Solve Question No. 7 OR Question No. 8.

(6) Solve Question No. 9 OR Question No. 10.

(7) Due credit will be given to neatness and adequate dimensions.

(8) Assume suitable data whenever necessary.

(9) Illustrate your answers wherever necessary with the help of neat sketches.

1. (a) Describe various OLAP operations in the multidimensional Data Model. 8

   (b) Compare between :

   (i) Data mart & Data warehouse.

   (ii) OLTP & OLAP. 6

OR

2. (a) Enumerate three classes of Schemes that are popularly used for modeling Data Warehouse. Write features of each Schema. 7

   (b) List different types of OLAP servers. Describe any two in detail. 7

3. (a) Describe KDD architecture with neat sketch. 8

   (b) State the need of Data Preprocessing. Describe different steps involved in Data Preprocessing. 6

OR

4. (a) Discuss the major issues in data mining. 8

   (b) Describe different kinds of data sources used for data mining. 6

MI—11433 1 (Contd.)

5. (a) Differentiate between classification and prediction. 6

   (b) Write the steps involved in decision tree algorithm. State its advantages and disadvantages. 8

OR

6. (a) Define Clustering. Describe various requirements of clustering and how it differs from classification. Also mention application area of both, classification and clustering. 8

   (b) Differentiate between K-means and K-medoids. 6

7. (a) What is Market Basket Analysis ? Explain in detail. 7

   (b) What do you mean by frequent item set, and closed item set ? Elaborate with example. 7

OR

8. (a) Generate FP – Tree for following transaction database. Assume min – sup = 30% :

TID	List of Items IDS
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

   6

   (b) A database has five transactions. Let min sup = 60% and min con f = 80%. Find all frequent item sets using Apriori algorithm : https://www.rtmnuonline.com

TID	Items bought
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, K, I, E}

MI–11433

9 (a) Explain the techniques for mining time series data. 7

(b) Discuss the basic measure for text retrieval. Explain various retrieval methods. 7

OR

10. (a) Graph and network mining have become increasingly important and heavily researched. Justify the same. 7

(b) Explain the concept of sequence pattern in detail. 7