Code: IT6T4

III B.Tech-II Semester–Regular/Supplementary Examinations–March 2018

DATA MINING & DATA WAREHOUSING
(INFORMATION TECHNOLOGY)

Duration: 3 hours   Max. Marks: 70

PART – A

Answer *all* the questions.   All questions carry equal marks

$$11x\ 2 = 22\ M$$

1. a) Describe about transactional database.

   b) What is data mining?

   c) What is dimensionality reduction?

   d) Describe snowflake schema.

   e) What are the differences between data matrix and

      dissimilarity matrix?

   f) Define support and confidence.

   g) Define a frequent set.

   h) What is classification?

   i) List out methods to improve accuracy of classification.

   j) What is cluster analysis?

   k) Write a note on grid based clustering.

PART – B

Answer any *THREE* questions. All questions carry equal marks.

$$3 \times 16 = 48 \text{ M}$$

2. a) Describe different data mining functionalities. 8 M

   b) Explain different data pre processing techniques. 8 M

3. Explain different OLAP operations in the multi dimensional data model. 16 M

4. a) Write Apriori algorithm for generating frequent item sets. 8 M

   b) Differentiate between maximal and closed frequent item sets. 8 M

5. a) Explain about rule based classification. 8 M

   b) What is meant by Bayesian classifier? How Bayes theorem is used for classification. 8 M

6. a) Write algorithm for DBSCAN clustering. 8 M

   b) Describe agglomerative hierarchical clustering algorithm? 8 M