Total No. of Questions : 7]   [Total No. of Printed Pages :2

Roll No ....

IT-840

B.E. VIII Semester
Examination, June 2017
Data Mining and Warehousing
(Elective - IV)

Time : Three Hours

Maximum Marks : 70

Note: Total number of question 07. Attempt any five questions (Including all parts). All questions carry equal mark.s Assume missing data if any, suitably.

1. a) What are the differences between the three main types of data warehouse usage: Information processing, Analytical processing, and data mining?   7
   b) Briefly describe the similarities and differences between star and snowflake schemas.   7

2. a) Illustrate how each of the following functions may be implemented in MOLAP.   7
   i)   The generation of a data warehouse
   ii) Roll-up
   iii) Drill-down
   b) Explain in detail about Data Discretization and Concept Hierarchy generation.   7

3. a) Describe challenges to data mining regarding data mining methodology and user interaction issues.   7
   b) Explain in detail about Data Cube Aggregation and Attribute Subset Selection.   7

IT-840   PTO

4. a) Apply Apriori algorithm for the following transactions to find all frequent itemsets with min_sub = 3.   7

TID	10	20	30	40
Item - Bought	{A, C, D}	{B, C, E}	{A, B, C,E}	{B, E}

   b) Explain in detail about various dimensionality reduction techniques.   7

5. a) What are the advantages of decision tree induction?   3
   b) What are the two approaches to tree pruning?   4
   c) Define outlier and challenges of outlier detection? Evaluate what information is used by outlier detection method?   7

6. a) Evaluate the need for data preprocessing by taking example of an application of your choice.   7
   b) What are the features of Bayesian classification explain in detail?   7

7. Write short note on any two of the following: 7 each
   a) Data Mart
   b) Confusion matrix
   c) Metadata repository

******

IT-840