Code :R5321205

R5

III B.Tech II Semester(R05) Supplementary Examinations, April/May 2011
DATA WAREHOUSING & DATA MINING
(Information Technology)

Time: 3 hours

Max Marks: 80

Answer any FIVE questions
All questions carry equal marks
⋆ ⋆ ⋆ ⋆ ⋆

1. (a) Discuss about Concept hierarchy.

   (b) Briefly explain about - classification of database systems.

2. (a) Briefly explain about data cleaning.

   (b) Why preprocessing of data is needed? Explain.

3. (a) Briefly discuss about Task-relevant data specification.

   (b) Explain the syntax for Task-relevant data specification.

4. (a) Explain about analysis of attribute relevance.

   (b) How is the analytical characterization performed? Explain with an example.

5. Propose and outline a level shared mining approach to mining multilevel association rules in which each item is encoded by its level position , and initial scan of the database collects the count for each item at each concept level, identifying frequent and sub frequent items. Comment on the processing cost of mining multilevel associations with this method in comparison to mining single level associations.

6. (a) Why is tree pruning useful in decision tree induction? What is a draw back of using a separate set of samples to evaluate pruning?

   (b) How rough set approach and fuzzy set approaches are useful for classification? Explain.

7. (a) What are the categories of major clustering methods? Explain.

   (b) Explain about outlier analysis.

8. (a) Discuss about multidimensional analysis and descriptive mining of complex data objects.

   (b) Explain text data analysis and information retrieval.

⋆ ⋆ ⋆ ⋆ ⋆