

Code No: A4001

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

M.Tech I Semester Examinations, March/April 2011

DATA WAREHOUSING AND MINING

(INFORMATION TECHNOLOGY)

Time: 3hours

Max. Marks: 60

Answer any five questions
All questions carry equal marks

- - -

1. a) Describe the identification of initial user requirements of a typical data ware house system.
b) What is preprocessing data? Why is it needed Attribute subset selection? [12]
2. a) Distinguish between operational data base systems and Data ware houses.
b) How are ROLAP servers different from MOLAP and HOLAP servers? [12]
3. Explain the following:
 - a) Generating association rules from frequent items.
 - b) Improving the efficiency of apriori.
 - c) Constraint based association mining. [12]
4. a) Briefly out line how to compute the dissimilarity between objects described by the following types of variables:
 - i) Asymmetric binary variables
 - ii) Nominal variables
 - iii) Ratio-scaled variables
 - iv) Numeric (interval-scaled) variables.
 b) Explain about DBSCAN. [12]
5. a) What are the issues regarding classification and prediction? Explain.
b) Discuss about Bayesian classification. [12]
6. a) List and describe any four primitives for specifying a data mining task.
b) Briefly discuss the data smoothing techniques. [12]
7. a) Write BIRCH algorithm.
b) Write short notes on Attribute-oriented induction. [12]
8. a) Explain similarity search in multimedia data bases and time-series data bases.
b) How mining WWW is different or similar to traditional data mining? [12]
