NR

Code No: B5801

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
M.Tech II Semester Examinations, October/November 2011
DATA WAREHOUSING AND MINING
(COMPUTER SCIENCE & ENGINEERING)

Time: 3hours **Max. Marks: 60**

Answer any five questions
All questions carry equal marks

- - -

1) What are the major challenges of mining a large amount of data? Differ with mining a small amount of data? [12]

2) What is missing values? Discuss the various methods to handle the problem of missing values? How would you differ from data cleaning? [12]

3) Write short notes on data mining primitives and data mining query languages? [12]

4) a) Differences between the data generalization and summarization?
   b) Define discriminating? How would you use it in different classes? [12]

5) a) Explain the various kinds of association rules in mining?
   b) Give an example to show that items in a strong association rule may actually be negatively correlated? [12]

6) a) Explain how association classification can be used for document classification?
   b) Define boosting and classification? [12]

7) What are the strengths and weaknesses of K-Means in comparison with the K-Medoids algorithms? [12]

8) a) Differentiate the precision and recall quality measures in IRS?
   b) Can we construct a data cube for multimedia data analysis? How? [12]

****