

NR

Code No: B6203

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

M.Tech II Semester Examinations, October/November 2011

DATA MINING AND WEB MINING

(WEB TECHNOLOGIES)

Time: 3hours

Max. Marks: 60

**Answer any five questions
All questions carry equal marks**

- - -

1. a) What are the various issues in data mining? Explain each one in detail.
b) Explain the OLAP Technology for data mining. [12]
2. What is data cleaning? How would it differ from missing values? [12]
3. Briefly describe one efficient incremental frequent item set mining method which can incorporate newly added transaction with adjustment (without redoing from scratch). [12]
4. What are the major challenges for mining long patterns for FP-Growth and Apriori? [12]
- 5.a) What is time series algorithm in data mining?
b) What are the various challenges in mining complex types of data? [12]
6. What is relevance ranking? How would you use relevance ranking in information retrieval? [12]
- 7.a) Discuss the importance of similarity measures in clustering.
b) Write a short note on Bayesian learners. [12]
- 8.a) Explain the various issues in the web crawling.
b) Discuss the examples of web crawlers. Explain each one in a brief. [12]
