

**Code No: C8209, C7810****JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****M.TECH I SEMESTER EXAMINATIONS APRIL/MAY-2012****DATA WAREHOUSING AND MINING****(COMMON TO COMPUTER NETWORKS, COMPUTER NETWORKS &  
INFORMATION SECURITY)****Time: 3hours****Max.Marks:60****Answer any five questions  
All questions carry equal marks**

- - -

- 1.a) Explain data mining as a step in the process of knowledge discovery.
- b) Differentiate operational database systems and data warehousing.
2. Write short notes for the following in detail:
  - a) Measuring the central tendency
  - b) Measuring the dispersion of data.
- 3.a) How can the data cube be efficiently constructed for discovery-driven Exploration?
- b) What is meant by dependent aggregates?
- c) Give one example for a 'complex query'.
- 4.a) Describe the data classification process with a neat diagram.
- b) How does the Naive Bayesian classification works? Explain.
- c) How to measure classifier accuracy?
- 5.a) How can we mine multilevel Association rules efficiently using concept hierarchies? Explain.
- b) Can we design a method that mines the complete set of frequent item sets without candidate generation? If yes, explain with example.
- 6.a) Given two objects represented by the tuples (22,1,42,10) and (20,0,36,8):
  - i) Compute the Euclidean distance between the two objects.
  - ii) Compute the Manhattan distance between the two objects.
  - iii) Compute the Minkowski distance between the two objects, using  $q = 3$ .
- b) Explain about Statistical-based outlier detection and Deviation-based outlier detection.
- 7.a) Give an example of generalization-based mining of plan databases by divide-and-conquer.
- b) What is sequential pattern mining? Explain.
- c) Explain the construction of a multilayered web information base.
8. Explain in detail
  - a) Web Content Mining
  - b) Web Linkage Mining
  - c) Web Usage Mining.