JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, March - 2017
DATA WAREHOUSING AND DATA MINING
(Computer Science and Engineering)

Time: 3 Hours

Max. Marks: 80

Answer any Five Questions
All Questions Carry Equal Marks
- - -

1.a) Explain the process of knowledge discovery in databases with necessary diagram.
  b) Discuss principle component analysis as a preprocessing activity.
  c) How to handle missing values? [16]

2.a) Define data warehouse. Compare and contrast data warehouse with operational database system.
  b) Write and explain BUC algorithm for data cube computation. [8+8]

3.a) Explain designing graphical user interface based on data mining query language.
  b) Discuss the DMQL syntax for task relevant data and visualization. [8+8]

4.a) Differentiate between characterization and analytical characterization. Quote suitable examples.
  b) How to use box-plot summary for measuring the data dispersion?
  c) What is a scatter plot? What is meant by loess curve? [16]

5.a) What is a FP-tree? What is its significance in association rule mining? Explain with an example data set.
  b) Discuss Association rule clustering system for mining quantitative association rules from relational database. [8+8]

6.a) How to Baye's theorem is used for classification? Explain in detail.
  b) What are the measures for classifier's accuracy?
  c) Discuss the classification process. Compare it with prediction. [16]

7.a) Discuss the similarity measures used in clustering.
  b) Explain the merits and demerits of partitioning methods. [8+8]

8.a) Explain the utilization of latent semantic indexing in text mining.
  b) Give the algorithm for web structure mining.
  c) What will be content of spatial data cube? [16]

--ooOoo--