B.Tech
(SEM VII) ODD SEMESTER THEORY EXAMINATION 2009-10
DATA MINING & DATA WAREHOUSING

Time: 3 Hours] [Total Marks: 100

Note: Attempt all questions.

1. Attempt any four parts: 5 \times 4 = 20

(a) Explain the data mining process with neat diagram.
(b) What do you mean by data cleaning?
(c) Explain clustering and regression with example.
(d) What is Z-score normalization?
(e) Distinguish between dimensionality reduction and numerosity reduction.
(f) Explain Histogram. The following data are a list of prices of commonly sold items at a company. The number have been sorted 1, 1, 5, 5, 5, 8, 8, 10, 10, 15, 15, 15, 20, 20, 20, 20. Make a histogram for price using singleton buckets.
2 Attempt any four parts:
(a) What do you understand by the terms data characterization in content to concept description?
(b) With the help of suitable example, explain data discrimination in brief.
(c) List out the reasons, why we perform attribute relevance analysis?
(d) What are the main purposes of statistics, used in data mining?
(e) What do you understand by outliers?
(f) What do you mean by association rules, for what purposes it is being used? Explain with example.

3 Answer any two parts:
(a) What are the different classification techniques? Discuss issues regarding classification and prediction.
(b) What do you mean by neural network? Explain multilayer Feed-Forward neural network. Differentiate between Feed-Forward and Feedback system.
(c) What do you mean by decision tree? Describe ID3 algorithm of the decision tree. Why it is unsuitable for data mining applications?
4 Attempt any two parts:

(a) Define the data warehousing with suitable example, why we need a separate data warehouse? Differentiate between OLAP and OLTP.

(b) What is a multidimensional data model? How we convert tables and spreadsheets to Data cubes? Convert 2-D tables into 3-D data cubes.

(c) (i) Explain Snow-Flake schema with an example.

(ii) Explain fact constellation with an example.

5 Attempt any two parts:

(a) Explain OLAP functions and tools in brief. What are the main features of OLAP servers?

(b) What do you mean by aggregation? Explain in brief, how the OLAP handles aggregation? Write the differences between MOLAP and HOLAP.

(c) Write short notes on:

(i) Slice and Dice operations

(ii) Testing of data warehouses.