05SCS242

USN | | | | | | | | | | |

NEW SCHEME

Second Semester M.Tech. Degree Examination, June 2007

Data Warehousing and Mining

Time: 3 hrs.] [Max. Marks:100

Note : ***Answer any FIVE full questions.***

1 a. Give the formal definition of a data warehouse and explain its characteristics. **(10 Marks)**

b. Compare operational data store with informational data. **(05 Marks)**

c. Discuss the need for data mining in organization. **(05 Marks)**

2 a. How does dependant data mart differ from independent data mart? Explain where they are implemented. **(08 Marks)**

b. Write a note on information delivery system. **(05 Marks)**

c. What are the guidelines and requirements for selecting OLAP systems? **(07 Marks)**

3 a. Explain MOLAP architecture with a neat diagram. **(07 Marks)**

b. State Baye's theorem. Give an instance of its application. **(05 Marks)**

c. What is meant by categorical predictors? Explain its classification. **(08 Marks)**

4 a. Differentiate between causality and collinearity with examples. **(08 Marks)**

b. Describe hypothesis testing taking any real-time application in detail. **(12 Marks)**

5 a. Explain contingency table used in statistics with an example. **(06 Marks)**

b. What do you mean by regression? Explain its purpose in business analysis. **(08 Marks)**

c. Discuss the terms used for measuring data mining effectiveness. **(06 Marks)**

6 a. What do you mean by application score card and business score card? Explain the same with respect to decision trees. **(10 Marks)**

b. Explain how ID3 algorithm used for handling high-cardenality predictors. **(10 Marks)**

7 a. Compare clustering with nearest-neighbor and also explain their usage. **(09 Marks)**

b. How do you classify clustering? Explain them in detail. **(11 Marks)**

8 a. Explain the different ways in which mutation can occur. **(12 Marks)**

b. Give the applications of generic algorithms in data mining. **(08 Marks)**

*****