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

05SCS242

Second Semester M. Tech. Degree Examination, May / June 08
Data Warehousing and Mining

Time: 3 hrs. Max. Marks:100

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

1 a. Discuss the need for data warehousing. **(05 Marks)**
b. Describe two tier data warehousing architecture and three tier data warehousing architecture by means of figures. **(08 Marks)**
c. Describe the challenges of operational data stores. **(07 Marks)**

2 a. Describe the different technological approaches to data warehouses data base. **(09 Marks)**
b. Elaborate on the issues to be dealt with by data sourcing, cleanup, extract, transformation and migration tools. **(06 Marks)**
c. Describe query and reporting tools in data warehousing environment. **(05 Marks)**

3 a. Explain ROLAP architecture by means of a diagram. How they differ from MOLAP? **(10 Marks)**
b. Why is online analytical processing (OLAP) needed? **(06 Marks)**
c. What are the approaches of deploying DSS tools? **(04 Marks)**

4 a. Describe summary statistics by means of an example. **(06 Marks)**
b. Differentiate between causality and collinearity using examples. **(08 Marks)**
c. Each of three identical jewelry boxes has 2 drawers. In each drawer of the first box there is a gold watch, in each drawer of the second box there is silver watch. If one drawer of the third box there is gold watch while in other there is silver watch. If we select a box at random open of the drawers and find it to contain a silver watch, what is the probability that the other drawer has got a watch? **(06 Marks)**

5 a. Describe business score card and algorithmic score card in data mining technologies. **(10 Marks)**
b. Explain prediction of wireless churns with CART. **(10 Marks)**

6 a. How does clustering and nearest neighbour work? **(10 Marks)**
b. Describe image recognition of human handwriting and solve bringing about strengths and weakness on its accuracy. **(10 Marks)**

7 a. What are genetic algorithms? **(05 Marks)**
b. How they can be used in business. **(05 Marks)**
c. Discuss the business scorecard and application scorecard for genetic algorithm. **(10 Marks)**

8 a. Explain the working of genetic algorithm. **(08 Marks)**
b. Discuss an application of genetic algorithm in data mining, for optimizing predictive customers segment. **(12 Marks)**

*****