B.E. Sem 8 (Rev.) Comp & I.T. Data Warehousing & Mining 15/12/08

Dec. PR 146

Con. 5563-08. **RC-7673**

(REVISED COURSE)

(3 Hours) **[Total Marks : 100**

N.B. : (1) Question No. **1** is **compulsory**.
(2) Attempt any **four** questions out of the remaining **six** questions.
(3) **All** question carry **equal** marks.

1. (a) Define Data Warehouse with feature. Explain the architecture warehouse with suitable block diagram. **10**
   (b) Explain Data mining as a step in KDD. Give the architecture of typical DM system. **10**

2. (a) Why metadata is important ? How to provide metadata ? **10**
   (b) Explain regression and association rules in Data mining along with example. **10**

3. (a) What are the type of OLAP Server ? Explain the different operation of OLAP with suitable example ? **10**
   (b) Explain the different methods of visualization with suitable example. **10**

4. (a) All Electronics company have sales dept. Sales consider four dimensions namely time, item, branch and location. The schema contains a central fact tables sales with two measures dollars–sold and unit sold. **10**
      (i) Define star schema and Snowflake schema for above case using DMQL.
      (ii) Design star schema and Snowflake schema for same.
   (b) List and describe five primitives for specifying data mining task. **10**

5. (a) Explain ETL of data warehousing in detail. **10**
   (b) What is clustering ? Explain k-means clustering algorithm. Suppose the data for clustering– { 2, 4, 10, 12, 3, 20, 11, 25 } **10**
   Consider k = 2, cluster the given data using above algorithm.

6. (a) State key issue to be considered while planning for data warehouse. Explain any four of them. **10**
   (b) Explain the general trend in data warehousing. **10**

7. Write short note on (any **two**) :– **20**
   (a) Web mining
   (b) Comparison between OLAP and OLTP
   (c) DMQL
   (d) Spatial clustering Algorithm.
