BE(IT) Sem VIII (R) 29/05/09
Data Warehousing & Mining 11-2

Con. 3162-09. VR-4314

**(REVISED COURSE)**

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

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

1. (a) Name any five types of activities that are part of ETL process. Which of these are time consuming ? Explain any three.
   (b) What are techniques and application of data mining ?

2. (a) How are top-down and bottom up approaches for building data warehouse differ ? Discuss the merits and limitation of each approach. Also discuss the practical approach.
   (b) Explain different algorithm for spatial mining.

3. (a) Describe feature of web enabled data warehouse. Why data security a major concern for web-enabled data warehouse ?
   (b) Explain Temporal mining.

4. (a) Give information package for recording information requirement for "Hotel Occupancy" considering dimensions like time, Hotel etc. Design Star schema from information package. If possible draw Snowflake schema.
   (b) Explain Web mining. Explain web content mining with reference to Crawlers, Harvest system, Virtual web view and personalization ?

5. (a) What is MDDB ? What types of business requirements determine use of MDDB in Data warehouse ?
   (b) Explain what is classification ? What are the issues in classification ? Apply statistical based algorithm to obtain the actual probabilities of each event to classify the new type as a tall. Use the following data :—

| Person ID | Name | Gender | Height | Class |
|---|---|---|---|---|
| 1 | Kristina | Female | 1.6m | short |
| 2 | Jim | Male | 2 m | Tall |
| 3 | Maggie | Female | 1.9 m | medium |
| 4 | Martha | Female | 1.85m | medium |
| 5 | John | Male | 2.1 m | Tall |
| 6 | Bob | Male | 1.7 m | short |
| 7 | Clinton | Male | 1.8 m | medium |
| 8 | Nyssa | Female | 1.6 m | short |
| 9 | Kathy | Female | 1.65 m | short |

[TURN OVER

6. (a) State and Explain key issues to be considered while planning for data warehouse for Information Technology department.

   (b) Explain Association Rule mining. Explain Aprioi algorithm with example.

7. Write short notes on (any **two**) :—

   (a) DMQL

   (b) Different operation of OLAP

   (c) Trends in Data mining

   (d) Fact constellation, Factless Fact Table, Updation to the Dimension Tables.