B.E.-Computer Technology Sem VIII

CT802 - Data Warehousing and Data Mining

P. Pages : 2 **GUG/S/18/7041**

Time : Three Hours * 2 1 5 7 * Max. Marks : 80

Notes : 1. All questions are compulsory.
2. All questions carry equal marks.
3. Due credit will be given to neatness and adequate dimensions.
4. Assume suitable data wherever necessary.

1. a) How datawarehouse is used for processing of an information explain? **8**

b) Write short notes on Enterprise warehouse Data mart and virtual warehouse. **8**

OR

2. a) Explain OLAP operations with examples. **8**

b) What are the different characteristics of datawarehouse. **8**

3. a) What is data mining functionality? Explain different types of data mining functionality with examples. **10**

b) Discuss the issues in data mining in detail. **6**

OR

4. a) Explain the KDD process in detail. What is data mining? **8**

b) What is data mining? Write short notes on classification of Data mining. **8**

5. a) Discuss issues to be considered during data integration. **6**

b) Describe the different methods for data cleaning. **10**

OR

6. a) Given two objects represented by tuples (22, 1, 42, 10) and (20, 0, 36, 8) **8**

i) Compute Euclidean distance

ii) Compute Manhattan distance

iii) Compute Minkowski distance
(Using q = 3)

b) Explain how statistical description of data can be given. **8**

7. a) Explain mining multilevel association rules from transactional data bases. **8**

b) Write in brief about Decision Tree. **8**

OR

8. a) Write an algorithm for k-nearest neighbour classification given k, the nearest number of neighbour and n the number of attributes descripting each type. **8**

b) Write in detail attribute selection measure. **8**

9. a) Explain partitioning methods along with k-means and k-medoid methods. **8**

b) What is outlier analysis? Why outlier mining is important? Explain statistical based outlier detection? **8**

OR

10. a) Explain the requirements of clustering in data mining. **8**

b) Explain grid based clustering methods. **8**

****