B.E.(with Credits)-Regular-Semester 2012 - Computer Technology Sem VIII

CT802 - Data Warehousing and Datamining

P. Pages : 2 **GUG/W/16/7041**

Time : Three Hours * 4 8 2 9 * Max. Marks : 80

Notes : 1. Answer **five** questions
2. Due credit will be given to neatness and adequate dimensions.
3. Assume suitable data wherever necessary.
4. Illustrate your answers wherever necessary with the help of neat sketches.

1. a) Define datawarehouse and explain important features of datawarehouse. **8**

b) Differentiate between OLAP and OLTP. **8**

OR

2. a) Draw architecture of datawarehouse and explain its components. **8**

b) What is data cube? Give example of 3D data cube. Also draw neat sketch of data cube. **8**

3. a) What is KDD (knowledge discovery in databases) process? Explain **8**

b) Discuss datamining functionality. characterization and discrimination with suitable example. **8**

OR

4. a) Write in brief about various applications of datamining. **8**

b) Give classification of datamining system. **8**

5. a) Explain in brief measures of data dispersion **8**

b) Suppose that the dataanalysis includes the attribute age. the age values are: 13, 15, 16, 16, 19, 20, 22, 22, 25, 30, 36, 40, 40, 52, 70.
compute mean, mode, median and standard deviation of age. Also compute midrange and five number summary. **8**

OR

6. a) Write in brief about data objects and attributes. **8**

b) Write in brief about major tasks in data preprocessing. **8**

7. a) How does classification works? Explain. Also give justification for statement : classification is supervised learning. **8**

b) Write about attribute selection measures. **8**

OR

8.	a)	Explain association rule, support and confidence with reference to association rule mining. Also write two steps of association rule mining.	**8**
	b)	Explain Apriori property used for association rule mining.	**8**
9.	a)	Give overview of basic clustering methods	**8**
	b)	Explain k-means partitioning algorithms.	**8**

OR

10.	a)	What are the requirements of clustering algorithm as data mining tool. Write in brief.	**8**
	b)	Explain DBSCAN clustering algorithms.	**8**

****