USN | | | | | | | | | |

MCA552

Fifth Semester MCA Degree Examination, June-July 2009
Data Mining and Warehousing

Time: 3 hrs. Max. Marks:100

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

1 a. Define Data Warehouse. Explain its features. **(04 Marks)**
b. Explain OLAP operations in the Multidimensional Data model with neat diagram. **(08 Marks)**
c. How to index OLAP data bitmap indexing and join indexing. Explain with example. **(08 Marks)**

2 a. With neat block diagram explain an integrated OLAM and OLAP architecture. **(08 Marks)**
b. What is data cleaning? Explain data cleaning method for filling missing values. **(08 Marks)**
c. Write and explain any four data reduction strategies. **(04 Marks)**

3 a. Give DMQL syntax for specifying the kind of knowledge to be mined. **(10 Marks)**
b. Explain primitives for specifying a data mining task. **(10 Marks)**

4 a. Write Apriori algorithm for discovering frequent item sets for mining boolen association rules. **(10 Marks)**
b. Discuss the different approaches to mining multilevel association rules. **(10 Marks)**

5 a. What are Bayesian classifiers? Explain working of simple Bayesian classifier. **(12 Marks)**
b. Explain Back propagation algorithm. **(08 Marks)**

6 a. Explain typical requirements of clustering in data mining. **(10 Marks)**
b. Write and explain categories of major clustering methods. **(10 Marks)**

7 a. Discuss Density-Based Spatial Clustering of Applications with Noise (DBSCAN). **(10 Marks)**
b. What is a distance-based outlier? Write a note on efficient algorithms used for mining distance-based outliers. **(10 Marks)**

8 a. What are the different criteria used in selection of a data mining system? Explain. **(10 Marks)**
b. Describe the trends in data mining. **(10 Marks)**

* * * * *