(Following Paper ID and Roll No. to be filled in your Answer Book)

PAPER ID : 1452 Roll No. | | | | | | | | | | |

M.C.A.

(SEMESTER-IV) THEORY EXAMINATION, 2011-12

DATA WAREHOUSING & MINING

Time : 3 Hours] *[Total Marks : 100*

Note : Attempt **all** Sections as directed.

SECTION – A

Attempt **all** parts. $10 \times 2 = 20$

1. Short answer type :
   (a) Define Data-Mart.
   (b) Write the Inmon definition of Data warehouse.
   (c) Explain meta data and its importance.
   (d) Explain the term crossover, mutation in Genetic algorithm.
   (e) What do you mean by data processing and data cleaning ?
   (f) Discuss for what purpose statistics is used in data mining.
   (g) Is web data analysis and web mining different ? Comment.
   (h) Define Classification and association techniques used to mine the data.
   (i) What is Bottom-up design approach to build a data warehouse ?
   (j) Define the fact table and dimension table of data warehouse.

SECTION – B

2. Attempt any **three** questions. $3 \times 10 = 30$
   (a) What do you mean by client – server model ? What is the need of client – server architecture ? Also, explain 2-tier and 3-tier architecture in detail.
   (b) Draw the data warehouse architecture and briefly explain its various components.
   (c) Explain the following :
      (i) Data ETL tools.
      (ii) Reporting & query tools and applications.
   (d) What is a decision tree, show by example ? Describe ID3 algorithm of the decision tree.
   (e) What is data visualization ? How can data visualization help in decision making ?

SECTION – C

Attempt **all** questions.

3. Attempt any **two** parts. **2 × 5 = 10**

   (a) Compare and contrast OLTP system and data warehouse system.

   (b) Explain multidimensional model of a data warehouse with help of an example.

   (c) Explain the following :

   (i) Shared – Disk architecture

   (ii) Shared – Nothing architecture

4. Attempt any **two** parts. **2 × 5 = 10**

   (a) What is data warehousing ? What can a data warehouse do ? List five common data warehousing applications.

   (b) Explain with help of suitable example/diagram the three basic conceptual DBMS schemas.

   (c) What are the different types of database parallelism and what are the data partioning techniques ?

5. Attempt any **one** part. **1 × 10 = 10**

   (a) Briefly explain and give Dr. E. F. Codd 12 guidelines and requirement as the basis for selecting OLAP systems.

   (b) What is an OLAP ? Explain MOLAP, ROLAP and Managed Query Environment (MQE), also give architecture for each.

6. Attempt any **one** part. **1 × 10 = 10**

   (a) Define data mining. List and draw the KDD process and explain the steps of the process. Also give application of data mining.

   (b) Explain the following algorithm of clustering.

   (i) K – Means

   (ii) Nearest Neighbour

7. Write short notes on any **two** of the following : **2 × 5 = 10**

   (a) Multimedia Data-Mining

   (b) Web Data-Mining

   (c) Neural Networks