Paper Code & Roll No. to be filled in your Answer Book

Roll No. | | | | | | | | | | |

Odd Semester Examination-2016

MCA (SEMESTER - V)

DATA WAREHOUSE & DATA MINING

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

Note : Attempt **all** questions. All questions carry **equal** marks.

1. Attempt **any four** parts of following: [5x4=20]

   (a) What is *data mining*? Describe the steps involved in data mining when viewed as a process of knowledge discovery.

   (b) How is a *data warehouse* different from a *database*? How are they similar?

   (c) What are the major differences between *operational database systems* and *data warehouses*?

   (d) What do you mean by *distance-based* algorithms?

[P.T.O.]

452 43 452

Scanned by CamScanner

(e) What are the different techniques of data *Visualization*? Explain in brief.

(f) Write a short note on Business Analysis.

2. Attempt any **four parts** of following: [5x4=20]

(a) What are the three layers of data warehouse architecture?

(b) What does *Data Mart* mean? Explain it, with its types.

(c) What are the different techniques of data warehouse modelling? Briefly describe the similarities and the differences of the two models, and then analyse their advantages and disadvantages with regard to one another.

(d) What are the differences between the three main types of data warehouse usage: *information processing*, *analytical processing*, and *data mining*?

(e) Identify the components of *distributed DBMS* environment.

(f) Explain briefly the concept of multimedia. How Data mining can be a part of it?

Scanned by CamScanner

3. Attempt **any four** parts of following: [5x4=20]

(a) Briefly outline the major steps of *decision tree classification*.

(b) Compare the advantages and disadvantages of *eager* classification versus *lazy* classification.

(c) Suppose that a data warehouse consists of the three dimensions *time, doctor*, and *patient*, and the two measures *count* and *charge*, where *charge* is the fee that a doctor charges a patient for a visit.

(i) Enumerate three classes of schemas that are popularly used for modelling data warehouses.

(ii) Draw a schema diagram for the above data warehouse using one of the schema classes listed in (i).

(d) What do you understand by *'Harvest System'* and *'Virtual Web View"*?

(e) What is the difference between two-tier and three-tier *Client/Server computing model*?

(f) Elaborate in brief about the applications of Data mining in contemporary world.

[P.T.O.]

Scanned by CamScanner

4. Attempt **any two** parts of following: [10x2=20]

(a) Describe three challenges to data mining regarding *data mining methodology* and *user interaction issues.*

(b) What is the relationship between *machine learning* and *data mining*?

(c) What is *Metadata*? Explain the role of metadata in data warehouse with its categories.

5. Attempt **any two** parts of following: [10x2=20]

(a) What do you mean by *web usage mining*? Explain its rules with examples.

(b) Explain N-N based algorithms.

(c) Briefly describe the classification processes using (a) *genetic algorithms*, (b) *rough sets*,and (c) *fuzzy sets.*

----- X -----

Scanned by CamScanner