

Roll No.

67162

**M.C.A. 4th Semester
Examination-May, 2013
(w.e.f. May, 2013)
Data Warehousing & Mining**

Paper-MCA-402

Time : 3 hours

Max. Marks : 80

Before answering the questions, candidates should ensure that they have been supplied the correct and complete question paper. No complaint in this regard will be entertained after the examination.

Note : Attempt **five** questions in all selecting at least **one** question from each unit.

Unit-I

1. (a) How schema is an integral part of Data warehouse ? Explain various types of schema with their advantages and disadvantages.
- (b) Explain the need of ETL process while creating data warehouse.

67162-1400-(P-3)(Q-8)(13)

(1)

[Turn Over

2. (a) Explain how Data warehouse has solved the problems related to databases. How virtual Data warehouse is different from warehouse ?
- (b) Why cleaning of data is important in warehouse process ? Explain any two such techniques.

Unit-II

3. (a) Differentiate OLTP and OLAP. Discuss different OLAP operation in multidimensional data model.
- (b) Discuss different views related to the design of Data warehouse. Also explain the steps for warehouse design process.
4. (a) Why three-tier data warehouse architecture is required ? Discuss different data warehouse model based on architectural view.
- (b) How computation of data cubes and indexing of OLAP data is useful for data warehouse implementation ?

Unit-III

5. (a) What are different methods for generation of concept hierarchies for categorical data ? Also give an example of concept hierarchy generation based

on number of distinct values per attribute.

- (b) What is data discretization ? Discuss its different categories. How entropy is considered as measure of discrimination ?
6. (a) Explain various approaches used for mining multilevel association rules.
- (b) What is constraint based association mining ? Discuss various constraints used for it. Also explain mining guided by additional rule constraints.

Unit-IV

7. (a) How classification is done by backpropagation ? What is the role of multilayer feed-forward neural network in it ?
- (b) How accuracy can be checked and estimated for any classification technique ?
8. (a) Discuss hierarchical method of cluster analysis with its different categories. Also differentiate BIRCH, ROCK and chameleon algorithms.
- (b) How mining of multimedia data on web is performed ?