

Roll No.

67208

**MCA 5th Sem. CBCS Scheme
(w.e.f. 2018-19)**

Examination – May, 2019

DATA WAREHOUSING AND DATA MINING

Paper : 18MCA35C3

Time : Three Hours] [Maximum 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 examination.

Note : Question No. 1 is *compulsory*. In addition to Question No. 1, attempt *four* more questions by selecting *one* question from each Unit. All questions carry equal marks.

1. (i) Name different layers/tiers in the Data Warehouse architecture.
- (ii) List different principles followed for designing a fact table.
- (iii) What is dropping condition used in Attribute Oriented Induction ?
- (iv) What is the purpose of Pivot OLAP operation ?

67208- 550 -(P-4)(Q-9)(19)

P. T. O.

- (v) Explain the SRSWOR sampling technique used for data reduction.
- (vi) What do you mean by the interestingness measures used in association mining ? Name any two such measures.
- (vii) Differentiate between the categorical, ratio-scaled & ordinal data.
- (viii) What do you understand by Dendrogram in reference to clustering approaches ?

UNIT – I

2. (i) Explain how data warehouse can handle the homogeneous as well as heterogeneous data.
- (ii) Elaborate the multidimensional data model followed by the data warehouse with all its components and features.
3. (i) Differentiate between Virtual and Enterprise data warehouse.
- (ii) What do you understand by the term Schema ? Explain the schema which is mostly preferred for smaller sized data warehouse.

UNIT – II

4. (i) Explain the concept of bitmap indexing used for efficient implementation of data cube in multidimensional data model.
- (ii) What role does the back-end tools-utilities & metadata repository play in the data warehouse architecture ?
5. (i) What do you mean by pre-computation/materialization of data cuboids ?

67208- -(P-4)(Q-9)(19) (2)

Explain different types of materialization with examples.

- (ii) Elaborate the use of Rollup, Drill down & Slice OLAP operations with their advantages.

UNIT – III

- 6. (i) Why it is recommended to preprocess the given data before data analysis ? Explain how Karl Pearson method can be used for solving the redundancy problem in the data.
- (ii) Apply the z-score normalization as a data transformation technique with the following data :
mean = Rs. 50,000 standard deviation = Rs. 15,000 and the targeted value = Rs. 70,000.
- 7. (i) Enumerate the concept of multidimensional and multi level association rules derived from association mining.
- (ii) What do you understand by the data generalization ? Explain how the generalization process can be controlled, with examples.

UNIT – IV

- 8. (i) Differentiate between the classifier & predictor used for data analysis. Also describe the classification model including the learning & testing steps.
- (ii) What are attribute selection measures ? Why they hold an important position in data mining ? Explain the information Gain method with

appropriate example for applying attribute selection.

- 9. (i) Explain the k-medoids Centroid based technique used for clustering the given data.
- (ii) What is the neighbourhood, core object density reachable & directly density reachable of an object used in Density based clustering methods.

