(ii) What are lazy learners ? Explain anyone such method that is most popularly used for classification purpose.

**9.** (i) Explain how the classifiers are evaluated in terms of sensitivity, specificity & precision with appropriate examples.

(ii) Elaborate anyone hierarchical based algorithm for clustering the data effectively.

---



*Roll No.* .....

# 67208

## MCA 5th Semester CBCS Scheme

## w. e. f. 2018-19

## Examination – December, 2018

### 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 :* Attempt ***five*** questions in all, selecting ***one*** question from each Unit. Question No. **1** is ***compulsory.*** All questions carry equal marks.

**1.** (i) How Virtual warehouse are designed and created ?

(ii) Differentiate between fact data and dimension data.

(iii) What do mean by climbing condition used in Attribute Oriented Induction ?

(iv) Write a short note on join indexing used in data warehouse.

P. T. O.

(v) What are the general optimization strategies used for Cube Computation ?

(vi) Name different constraints that can be applied to Association mining.

(vii) Differentiate between the categorical, ratio-scaled & ordinal variables.

(viii) What is the concept of Clustering Feature (CF) used in Clustering algorithm.

## UNIT – I

**2.** (i) Why digitization has increased the need of data warehouse ? Explain. How data warehouse solves different issues related to handling & storing of data ?

(ii) Elaborate how Data Warehouse is constructed by discussing all tiers of its architecture with the help of labelled diagram.

**3.** (i) Differentiate between Snowflake and Galaxy schema of data warehouse with their respective advantages and disadvantages.

(ii) What do you mean by partitioning of data set ? Differentiate between Vertical and Horizontal partitioning with examples . :

## UNIT – II

**4.** (i) How Online Analytical Processing (OLAP) is different from Online Transaction Processing (OLTP) ?



(ii) Write an explanatory note on different components & features of multidimensional data model followed by the data warehouse.

**5.** (i) Define Summarization. How different degrees of summarization is calculated and represented in the form of Lattice.

(ii) What is Concept Hierarchy ? Discuss why it is known as backbone of the multidimensional model based data warehouse.

## UNIT – III

**6.** (i) What do you mean by the contingency table ? Explain how the correlation can be evaluated for categorical data in pre processing phase.

(ii) Discuss how Histograms can be used for performing numerosity reduction with appropriate examples.

**7.** (i) Discuss why Attribute Oriented Induction is known as query oriented generalization approach ? Briefly explain the generalization process.

(ii) State the Apriori property used in association mining algorithm. Explain how it is applied on the given data set for association mining.

## UNIT – IV

**8.** (i) State the Bayes' theorem. How Naive Bayesian classification is performed to classify the numeric & categorical data.