7. (i) Explain how Market Basket analysis is performed by using the concept of Frequent pattern analysis.

(ii) Elaborate Multidimensional and Multilevel association rules with appropriate examples.

Unit-IV

8. (i) How supervised learning is different from unsupervised learning approaches ? Mention some pre requisites that are mandatory required before an approach can be called as supervised learning.

(ii) What do you mean by attribute selection methods ? Formulate any three feature selectors with their respective characteristics.

9. (i) How can Bagging and Boosting approaches enhance the performance of classifiers and predictors ?

(ii) Write down and explain any one method for implementing Hierarchical Clustering on the given data.

Roll No. : .....

Total No. of Questions : 9] [Total No. of Pages : 4

67208

MCA 5th Semester (Regular)
Examination, February-2022
(CBCS Scheme Current w.e.f. 2018-19)

Paper-18MCA35C3

DATA WAREHOUSING AND DATA MINING

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. Q. No. 1 is compulsory. All questions carry equal marks.

Compulsory Question

1. (i) List down any *four* features of Data Warehouse over the Database Systems.

(ii) Differentiate between Data Cube and Data Cuboid.

(iii) What type of users are meant for using the OLAP data sources ? Explain.

(iv) Explain the Pivot OLAP operation used in Data Warehouse.

(v) List down any four problems faced while working with real dataset.

(vi) How metadata can be used to solve Entity Identification problem while preprocessing ?

(vii) Draw a labelled diagram to represent Classification model for an educational organization.

(viii) Briefly discuss Bagging and Boosting techniques.

Unit–I

2. (i) Discuss how Data Warehouse handles heterogeneity of data, coming from different data sources, with examples.

(ii) What challenges are faced while transforming the data for making it most appropriate for data analysis ?

3. (i) Elaborate the multi-dimensional data model followed for creating and implementing Data Warehouse. Draw the labelled diagram for showing Lattice for three dimensions.

(ii) How dimension tables can be normalized in Snowflake schema for designing Data Warehouse ? Also mention its advantages.

Unit–II

4. (i) How OLTP is different from OLAP technology in terms of data type, operations and orientation ?

(ii) Why OLAP Servers are considered to be an integral part of Data Warehouse structure ? Define different types of OLAP servers.

5. What is Data Materialization ? Why is it always recommended to materialize the Data cuboid before actual Data Analysis. Explain in what scenarios, full materialization is preferable than partial materialization with proper examples.

Unit–III

6. (i) Explain any four methods for removing the missing values from the given data during data preprocessing.

(ii) Differentiate between data characterization and data discretization with proper examples.