

50185

M.C.A. EXAMINATION, 2022

(Second Semester)

(Regular and Re-appear)

(From 2020-21 Onwards)

DATA WAREHOUSING AND MINING

Time : 3 Hours]

[Maximum Marks : 48

Note : Attempt *Five* questions in all, selecting *one* question from each Unit. Q. No. 1 is compulsory.

1. Briefly describe the following questions : $6 \times 2 = 12$
- (a) Which data mining method could be efficient to determine the impact of COVID-19 pandemic ?
 - (b) Differentiate between star schema and snowflake schema.

- (c) List out some social impact of data mining.
- (d) How data mining can be used for intrusion detection ?
- (e) What do you mean by Data Transformation ?
- (f) Describe factless fact table with an example.

Unit I

2. (a) Describe, how Data warehouse has solved the problem related to databases ? How virtual data warehouse is different from warehouse ? 4
- (b) With a neat diagram, explain the architecture and components of data warehouse. 5
3. What are the goals of data warehouse ? Describe the need of ETL process while creating data warehouse. List various methods of data cleaning and data integration. 9

Unit II

4. What do you mean by Partitioning and Aggregation ? Why partitioning is needed ? Explain various strategies used for partitioning. 9

5. (a) Describe Metadata. Differentiate between Data Mart and Data Warehouse. 5

(b) What is the necessity of tools in data warehouse ? Describe the role of data warehouse system manager and load manager. 4

Unit III

6. (a) Describe Kamber's model of knowledge discovery in database with the help of a diagram. 4

(b) What is Regression ? Define linear regression. Differentiate between simple regression and multiple regression. 5

7. Write short notes on the following :

(a) Data Discretization 2

(b) Classification by Support Vector Machine 3

(c) Confusion Matrix and ROC curves. 4

Unit IV

8. (a) What are Multi-dimensional association rules ? How these rules are mined from relational database ? 5

(b) What is Association rule Mining ? Illustrate usefulness of Constraint base association data mining. 4

9. What is meant by Clustering ? Explain Partitioning, Hierarchical and Grid based methods of clustering in detail. 9