

Total No. of printed pages = 4

63/2 (SEM-4) MCA 4.2

2022

MCA

(Theory Paper)

Paper Code : MCA 4.2

(Data Mining and Warehousing)

Full Marks – 75

Time – Three hours

**The figures in the margin indicate full marks
for the questions.**

1. Answer the following questions : 1×5=5

(a) Where a data warehouse used ?

(i) Decision support system

(ii) Transaction system

(iii) Logical system

(iv) None of these

[Turn over

(b) What are the functions of Data Mining ?

- (i) Association and correlational analysis classification
- (ii) Prediction and characterization
- (iii) Cluster analysis and Evolution analysis
- (iv) All of the above

(c) _____ describe the data contained in the data warehouse.

- (i) Relational data
- (ii) Operational data
- (iii) Meta data
- (iv) Informational data

(d) Which of the following is an essential process in which the intelligent methods are applied to extract data patterns ?

- (i) Warehousing
- (ii) Data mining
- (iii) Text mining
- (iv) Data selection

(e) The source of all data warehouse data is the _____.

- (i) Operational environment
- (ii) Informal environment
- (iii) Formal environment
- (iv) Technology environment.

2. Answer the following questions : $2 \times 5 = 10$

- (a) How to compute confidence for an association rule $X \rightarrow Y$?
- (b) What is a data warehouse ? How it is differs from DBMS ?
- (c) What is the role of meta data repository.
- (d) Define data cube. Give an example.
- (e) Differentiate between ROLAP and HOLAP.

3. Answer any *six* of the following questions :

$5 \times 6 = 30$

- (a) What is the importance of data marts in data warehouse ?
- (b) How does PAM algorithm work ? Explain.
- (c) Give example for defining different schemas of data warehouse.
- (d) Explain the requirements of cluster analysis.

- (e) Describe about a three-tier data warehouse architecture.
 - (f) Compare K-mean with K-medoids algorithms for clustering.
 - (g) Describe about the major issues in data mining.
4. Answer any *three* of the following questions :
10×3=30
- (a) With necessary diagrams and examples of data cubes explain various OLAP operations.
 - (b) Write DBSCAN clustering algorithm and estimate time and space complexity. State and explain the issues in cluster evaluation.
 - (c) Describe the various phases in knowledge discovery process with a neat diagram.
 - (d) Discuss about the data mining architecture with necessary diagram.