

Roll No.

Total No. of Pages : 2

MCA/MX

5267

Data Warehousing and Data Mining

Paper : MCA-402

Time : Three Hours]

[Maximum Marks : 80

Note :- Attempt **FIVE** questions in all. Question No. **1** is compulsory. In addition to compulsory question, attempt **FOUR** more questions selecting **ONE** question from each Unit.

1. (a) Differentiate between ROLAP and MOLAP.
(b) What do you mean by frequent item-set ?
(c) Explain the term 'Lattice of Cuboid'.
(d) What are the various reasons for noisy data ?
(e) What should be the criteria for pattern interestingness ?
(f) Difference between classification and clustering.
(g) Explain data integration.
(h) What is data preprocessing ? 3 each

UNIT-I

2. (a) Why are operational databases not particularly suited for decision support system ? 6
(b) What is difference between OLAP and OLTP systems ? How a data warehouse is related to OLAP systems ? 8
3. Discuss various schemas used for modelling a data warehouse using suitable examples. 14

UNIT-II

4. (a) Is data mining another hype or a fact ? What are the factors responsible for evolution of data mining ? 8
(b) Describe various data mining tasks using examples. 6

5. (a) Discuss various strategies used for handling missing data. Which method is more preferable and why ?
(b) List and describe five data mining primitives for specifying a data mining task.

UNIT-III

6. (a) What is 'Association Rule' ? Discuss various types of association rules.
(b) What is meant by 'Association Rule Mining' ? Explain by giving a suitable example.
7. Describe various methods used for 'Data Classification and Prediction'.

UNIT-IV

8. (a) Give classification of various clustering methods.
(b) Define K-means algorithm by using an example.
9. Write short notes on the following :
(a) Outlier Analysis
(b) Tools for Data Mining.