

(REVISED COURSE)

(3 Hours)

[Total Marks : 100

N.B. (1) Question No. 1 is compulsory.

(2) Attempt any four out of remaining six questions.

(3) Figures to the right indicate full marks.

(4) Illustrate answers with neat sketches wherever required.

1. (a) Define Data warehouse with features ? Explain the architecture of data warehouse with suitable block diagram. 10
(b) Draw KDD process diagram ? Explain in detail ? 10
2. (a) What are the types of OLAP servers ? Explain the different operation of OLAP with suitable example ? 10
(b) Explain the different methods of visualization with suitable example ? 10
3. (a) Differentiate between star schema, snowflake schema and fact constellation. 10
(b) Define classification ? Explain any two algorithms with suitable example ? 10
4. (a) Name any five types of activities that are part of ETL process ? Which of these are time consuming ? Explain any three ? 10
(b) What is Association Rule mining ? Give the Apriori Algorithm. Apply association Rule to find all frequent item sets from following table ? 10

Trans-id	Items
100	1, 3, 4, 6
200	2, 3, 5, 7
300	1, 2, 3, 5, 8
400	2, 5, 9, 10
500	1, 4

(Let min-support = 60% and min-confidence = 80)

5. (a) Why metadata is important ? How to provide metadata ? 10
(b) What are techniques and application of data mining ? 10
6. (a) What is MDDDB ? What types of business requirements determine use of MDDDB in Data warehouse ? 10
(b) Explain K-mean clustering algorithm ? Suppose the data for clustering is { 1, 3, 5, 15, 23, 11, 25 }. Consider k = 2, cluster the given data using above algorithm. 10
7. Write short notes on the following (any four) :- 20
 - (a) Data warehouse project planning and management
 - (b) Temporal mining
 - (c) Comparison between OLAP and OLTP
 - (d) Web structure mining
 - (e) Trends in data mining.