

Con. 3276-08.

CO-4135

(REVISED COURSE)

(3 Hours)

[Total Marks : 100]

N.B. 1. Question no. 1 is compulsory.

2. Attempt any four questions out of the remaining six questions.

3. All Question carry equal marks.

4. Illustrate answers with neat sketches whenever required.

- Q1 (a) How are top-down and bottom up approaches for building data warehouse differ? Discuss the merits and limitation of each approach? 10
(b) Explain Data mining as a step in KDD. Give the architecture of typical DM system. 10
- Q2 (a) Give information package for recording information requirement for "Hotel Occupancy" considering dimensions like time, Hotel etc. Design star schema from information package. 10
(b) What is Association rule? Describe the suitable technique using FP trees of mining frequent pattern? 10
- Q3 (a) Describe features of web enabled Data warehouse? Why is data security a major concern for Web-enabled data warehouse? 10
(b) Define classification. Explain K-nearest neighbor classification algorithm. 10
- Q4 (a) what are the types of OLAP servers? Explain the different operations of OLAP with suitable example? 10
(b) List and describe five primitives for specifying data mining task. 10
- Q5. (a) Explain ETL of data warehousing in detail. 10
(b) What is clustering? Explain k – means clustering algorithm. Suppose the data for clustering - {2,4,10,12,3,20,11,25} 10
Consider $k = 2$, cluster the given data using above algorithm.
- Q6 (a) State key issues to be considered while planning for data warehouse. Explain any four of them. 10
(b) What is visualization? How can you mine descriptive statistical measures in large databases? 10
- Q 7. Write short notes on the following (any Two) 20
a) Comparison between Data warehouse and Data Mart
b) Applications and Trends in Data mining
c) Temporal mining
d) DMQL.