



- Notes :
1. All questions carry marks as indicated.
 2. Solve Question 1 OR Questions No. 2.
 3. Solve Question 3 OR Questions No. 4.
 4. Solve Question 5 OR Questions No. 6.
 5. Solve Question 7 OR Questions No. 8.
 6. Solve Question 9 OR Questions No. 10.
 7. Solve Question 11 OR Questions No. 12.
 8. Due credit will be given to neatness and adequate dimensions.
 9. Assume suitable data whenever necessary.
 10. Illustrate your answers whenever necessary with the help of neat sketches.

1. a) Explain Three - tier architecture of Data warehouse in detail. 7
- b) Differentiate between Data Warehouse and operational Databases. 6

OR

2. a) Write short notes on: 6
 - i) Information Delivery Methods.
 - ii) Escalating need for strategic information.
- b) Explain various approaches to build Data Warehouse and Data Marts along with their advantages and disadvantages. 7
3. a) What is OLAP? What are different operations that can be performed on multidimensional data model. 8
- b) Explain how fast access and powerful calculations are performed using OLAP. 5

OR

4. a) Explain types of OLAP servers. 7
- b) List and explain initial Twelve guidelines of OLAP suggested by Dr. E.F. Codd. 6
5. a) What are various types of data? Explain their uses. 5
- b) Explain the following terms related to dimensional model: 9
 - i) FACT TABLE.
 - ii) DIMENSION TABLE.
 - iii) LARGE DIMENSION TABLE.

OR

6. a) Describe the STAR and SNOWFLAKE scheme with neat sketch. 8
- b) Write a detailed note on Data cubes of multidimensional data model. 6

7. a) Explain KDD process with neat diagram. **8**
- b) What do you mean by outlier? How outlier analysis is useful in brand detection? **6**

OR

8. a) Discuss the major issues in Data mining. **7**
- b) Explain data preprocessing process in detail. **7**
9. a) Consider following transactional dataset. find frequent item sets and association rules using priori algorithm, with support = 30% and confidence = 70% **9**

TID	List of Items IDS
T101	I ₁ , I ₂ , I ₅
T102	I ₂ , I ₄
T103	I ₂ , I ₃
T104	I ₁ , I ₂ , I ₄
T105	I ₁ , I ₃
T106	I ₂ , I ₃
T107	I ₁ , I ₃
T108	I ₁ , I ₂ , I ₃ , I ₅
T109	I ₁ , I ₂ , I ₃ ,
T110	I ₂ , I ₃ , I ₄

- b) Write in brief about constraint based association mining. **4**

OR

10. a) Define the following terms. **6**
- Frequent item sets.
 - Closed Item sets.
 - Association rules.
- b) Explain Bayesian classification and rule based classification in short. **7**
11. a) What do you mean by Hierarchical clustering approach? explain agglomerative and divisive hierarchical clustering. **8**
- b) Explain the computation of dissimilarity between object having following type of data: **5**
- Interval - Scaled variable.
 - Binary variable.

OR

12. a) Write short notes on: **any three**. **3**
- Data mining in Retail Industry. **3**
 - Web mining and its challenges. **3**
 - Principle of Memory Based Reasoning (MBR) **3**
 - Benefits of Data Mining. **3**
- b) How high dimensional data is clustered? Explain in detail. **4**
