

(3 Hours)

Total Marks: 80

- N.B. :** 1) Question No.1 is **compulsory**.
 2) Attempt any **THREE** from the remaining questions.
 3) Figures to the right indicate full marks.

- Q1.** Write a short note on:
- (a) Star and Snowflake schema. [5]
 - (b) Data preprocessing [5]
 - (c) Differentiate between ORDBMS and OODBMS [5]
 - (d) Text Retrieval Methods [5]
- Q2.** (a) Explain Distributed database architecture [10]
 (b) Explain the KDD process in detail. [10]
- Q3.** (a) Differentiate the following, [10]
 1. Agglomerative and Divisive clustering
 2. OLAP Vs Data Mining
- (b) Apply Apriori algorithm to the following data set to find out strong association rule with Support= 50% and Confidence=60%. [10]

Transaction ID	Items
100	Dairymilk, Chips, cold drink, Juice
200	Dairymilk, Chips, cold drink
300	Dairymilk, Biscuits, Mayonnaise
400	Dairymilk, cold drink, Biscuits

- Q4.** (a) Explain Data warehouse architecture in detail. [10]
 (b) Why clustering is an unsupervised technique? Explain K-means algorithm. [10]
- Q5.** (a) Explain the following, [10]
 1. Web mining
 2. Bayesian classification
 (b) Explain the Decision tree used in classification. Explain ID3 algorithm with a suitable example. [10]
- Q6.** (a) What is the Abstract Data type. Explain with suitable example. [10]
 (b) Explain Data reduction techniques. [10]
