Name. _____ **Printed Pages:01**

Student Admn. No.: _____

School of Computer Science and Engineering

Back Paper Examination Even Semester (Non - Graduating Batches) – June 2024

[Programme: B.Tech] [Semester: IV] Batch:]

Course Title: Data warehouse and Datamining **Max Marks: 100**

Course Code: E2UC406B / BTCS9213 **Time: 3 Hrs.**

Instructions:

1. *All questions are compulsory.*
2. *Assume missing data suitably, if any.*

		K Level	COs	Marks
	SECTION-A (15 Marks) 5 Marks each			
1.	Define data warehouse and discuss the basic characteristics of a data warehouse.			5
2.	Define support and confidence. Calculate support and confidence for A->D & C->A. Transactions: T1 {A, B, C}; T2 {A, C, D}; T3 {B, C, D}; T4 {A, D, E}; T5 {B, C, E}			5
3.	Discuss Data reduction in detail.			5
	SECTION-B (40 Marks) 10 Marks each			
4.	a) Compare OLTP and OLAP b) What are steps in designing the data warehouse? Explain			10
5.	Explain about the Three-tier data warehouse architecture with a neat diagram.			10
6.	For the following given Transaction Data-set, Generate Association Rules using Apriori Algorithm. Consider the values as Support=50% and Confidence=60% T1: I1, I2, I3 T2: I2, I3, I4 T3: I4, I5 T4: I1, I2, I4 T5: I1, I2, I3, I5 T6: I1,I2,I3,I4			10
7.	What is KDD? Explain about data mining as a step in the process of knowledge discovery.			10
	SECTION-C (45 Marks) 15 Marks each			
8.	Give Brief description of following: (a) Binning (b) regression (c) Clustering (d) Smoothing (e) Generalization (f) Aggregation			15

9.	a) Define Clustering? Explain about Types of Data in Cluster Analysis? b) Classify various Clustering methods. c) Write any one Partitioning based clustering methods.			15
10	a) what is a schema? Discuss various schemas used in data warehouse. b) Differentiate between star and snowflake schema and explain the advantages and disadvantages of snowflake schema.			15