28 NOV 2023

D Y PATIL UNIVERSITY — RAMRAO ADIK INSTITUTE OF TECHNOLOGY

RAMRAO ADIK INSTITUTE OF TECHNOLOGY, NERUL

(D Y Patil Deemed to be University)

Program: B.Tech in Computer Engineering
End Semester Examination: B.Tech. Semester VI
Course Code: CECDLO6032 Course Name: Data Warehousing and Mining
Time: 2 hour
Max. Marks: 60

Instructions: 1. All three questions are compulsory

Que. No.	Question	Max. Marks	CO	BT
Q1	Solve any **Four**			
i)	Illustrate how data mart is important in data warehouse architecture?	5	CO1	BT4
ii)	What are the techniques used for Data Loading?	5	CO2	BT2
iii)	Enlist issues in data mining. Discuss anyone issue in detailed with an example.	5	CO3	BT4
iv)	Differentiate between multilevel and multi-dimensional association rules.	5	CO4	BT4
v)	Differentiate between Agglomerative and Divisive Clustering in Hierarchical Methods.	5	CO5	BT4
vi)	What are the different techniques used in web content mining?	5	CO6	BT2

Que. No.	Question	Max. Marks	CO	BT
Q2 A	Solve any **Two**			
i)	Define Web content, structure and Usage mining in short.	5	CO6	BT4
ii)	How data mining is important in Global Warming data analysis?	5	CO3	BT4
iii)	Justify, importance of market basket analysis in today's Business Analytics?	5	CO4	BT4
iv)	Enlist OLAP operations. Explain ROLL up operation with an example.	5	CO2	BT4
Q 2 B	Solve any **One**			
i)	Discuss : i. The steps in KDD process. ii. The architecture of typical Data Mining system	10	CO3	BT2
ii)	Discuss Association Rule Mining and Apriori Algorithm. Apply Apriori Rule mining to find all frequent item sets and association rules for the following dataset with minimum support count is 2 and 70% minimum confidence.	10	CO4	BT5

Transaction ID	Items
100	1,2,5
200	2,4
300	2,3

D Y PATIL UNIVERSITY — RAMRAO ADIK INSTITUTE OF TECHNOLOGY, NERUL

RAMRAO ADIK INSTITUTE OF TECHNOLOGY, NERUL

(D Y Patil Deemed to be University)

		400 | 1,2,4 500 | 1,3 600 | 1,3 700 | 1,3,2,5 800 | 1,3 900 | 1,2,3			

Que. No.	Question	Max. Marks	CO	BT
Q3	Solve any **Two**			
i)	For a supermarket chain consider the following dimensions namely product, store, time and promotion. The schema contains a central fact table for sales. i. Design star schema for above example. ii. Design snowflake schema for the same.	10	**CO1**	**BT6**
ii)	Discuss different steps involved in Data Preprocessing.	10	**CO3**	**BT4**
iii)	Explain k-mean clustering algorithm. Apply k-mean algorithm for the following dataset with two clusters. Data Set = {1,2,6,7,8,10,15,17,20}	10	**CO5**	**BT5**

==========

Course Outcomes (CO) -Learner will be able to:

CO1. Understand Data Warehouse fundamentals with dimensional modelling
CO 2. Understand OLAP operations in Multidimensional Data Model
CO 3. Understand Data Mining and Data Pre-processing steps.
CO 4. Explore frequent patterns and Association mining algorithms.
CO 5. Apply various classification and clustering techniques on real world scenario.
CO 6. Describes social network in Web Mining and apply web mining algorithm.

BT1- Remembering, BT2- Understanding, BT3- Applying, BT4- Analyzing, BT5- Evaluating, BT6- Creating