



MARRI LAXMAN REDDY

INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

COURSE CONTENT

DATA MINING LABORATORY								
VI Semester: CSM								
VII Semester: CSE								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
24X0586	Core	L	T	P	C	CIA	SEE	Total
		0	0	2	1	40	60	100
Contact Classes: Nil	Tutorial Classes: Nil	Practical Classes:30			Total Classes: 30			
Prerequisites: Database Management Systems Laboratory								

Course Overview:

This Data Mining Lab focuses on practical implementation of data preprocessing techniques such as cleaning, normalization, integration, and data partitioning using Weka/Python tools. Students learn data warehousing concepts including schemas, OLAP operations, and ETL processes for managing and analyzing large datasets. The lab provides hands-on experience in association rule mining and classification methods like Apriori, FP-Growth, Decision Trees, Bayesian, and K-Nearest Neighbor algorithms. It also covers clustering techniques such as K-Means, BIRCH, PAM, and DBSCAN to discover patterns and group similar data effectively.

Course Objectives:

- To understand and apply various data preprocessing techniques including cleaning, transformation, and integration to prepare datasets for analysis.
- To explore data partitioning and warehouse schemas to efficiently organize and manage large-scale data.
- To implement data mining algorithms such as association rule mining, decision trees, and clustering methods to extract meaningful patterns.
- To perform classification and prediction using techniques like Bayesian classification and K-nearest neighbor methods.
- To develop skills in OLAP operations and data cube construction for multidimensional data analysis and decision-making.

Course Outcomes: After Completion of the Course, Students should be able to

- Apply data preprocessing techniques to clean, transform, and integrate datasets for effective analysis.
- Implement data warehousing solutions using different schemas and data partitioning methods.
- Execute data mining algorithms like Apriori, FP-Growth, and decision trees to discover patterns and relationships in data.
- Perform data classification and clustering using methods such as Bayesian, KNN, K-Means, BIRCH, PAM, and DBSCAN.
- Perform OLAP operations to analyze multidimensional data for informed decision-making.

List of Experiments: Experiments using Weka/ Python

1. Data Processing Techniques:
 - (i) Data cleaning
 - (ii) Data transformation – Normalization
 - (iii) Data integration
2. Partitioning - Horizontal, Vertical, Round Robin, Hash based
3. Data Warehouse schemas – star, snowflake, fact constellation
4. Data cube construction – OLAP operations
5. Data Extraction, Transformations & Loading operations
6. Implementation of Attribute oriented induction algorithm
7. Implementation of apriori algorithm
8. Implementation of FP – Growth algorithm
9. Implementation of Decision Tree Induction
10. Calculating Information gain measures
11. Classification of data using Bayesian approach
12. Classification of data using K – nearest neighbour approach
13. Implementation of K – means algorithm
14. Implementation of BIRCH algorithm
15. Implementation of PAM algorithm
16. Implementation of DBSCAN algorithm

TEXT BOOKS:

1. Data Mining – Concepts and Techniques - JIAWEI HAN & MICHELINE KAMBER, Elsevier.
2. Data Warehousing, Data Mining & OLAP- Alex Berson and Stephen J. Smith- Tata McGraw-Hill Edition, Tenth reprint 2007

REFERENCE BOOKS:

1. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Anuj Karpatne, Introduction to Data Mining, Pearson Education

ELECTRONIC RESOURCES:

<https://cse20-iiith.vlabs.ac.in/exp/mst-based/>

<https://www.geeksforgeeks.org/data-science/data-mining/>

<https://www.kaggle.com/code/thisray/data-mining-lab1-2018-news-data-mining>

<https://www.hackerrank.com/blog/getting-started-with-data-science-and-machine-learning-2/>

MATERIALS ONLINE:

1. Lab Manual
2. Open-ended experiments