



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

VI Semester: CSM

VII Semester: CSE

Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P		CIA	SEE	Total
24X0518	Core	3	0	0	3	40	60	100
Contact Classes: 45	Tutorial Classes: Nil	Practical Classes: Nil			Total Classes: 45			
Prerequisites: Database Management System								

Course Overview:

This course introduces the fundamentals of Data Mining, including data types, preprocessing techniques, statistical analysis, and data visualization methods. It covers association analysis techniques such as Apriori and FP-Growth to discover patterns and relationships in large datasets. The syllabus also focuses on classification and clustering methods, including decision trees, neural networks, SVM, k-means, hierarchical clustering, and outlier detection. Finally, it explores advanced topics like web mining, spatial mining, and temporal data mining for analyzing complex real-world data patterns over time and space.

Course Objectives:

- To introduce the fundamental concepts of data mining, knowledge discovery, and data preprocessing techniques.
- To familiarize students with association analysis and pattern mining algorithms such as Apriori and FP-Growth.
- To develop the ability to implement and evaluate classification techniques including Decision Trees, Bayes Classifiers, and SVMs.
- To provide knowledge of clustering methods, outlier detection, and their applications in analyzing complex datasets.
- To expose students to advanced data mining concepts such as Web Mining, Spatial Mining, and Temporal Mining for real-world applications.

Course Outcomes: After completion of the course, students should be able to

- Understand fundamental concepts of data mining, data preprocessing, and similarity measures for real-world datasets.
- Implement association analysis techniques like Apriori and FP-Growth for discovering meaningful patterns in data.
- Apply classification algorithms including Decision Trees, Bayes Methods, SVM, and Neural Networks to solve predictive problems.
- Demonstrate clustering techniques, outlier detection, and evaluation methods for grouping and analyzing complex data.
- Analyze advanced data mining concepts such as Web, Spatial, and Temporal Mining for domain-specific applications.

Module- I **[10]**

Introduction to Data Mining: What Data mining? Kinds of Data, Knowledge Discovery process, Data Mining Functionalities, Kinds of Patterns, Major Issues in Data Mining. Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity and Dissimilarity, Data Pre-processing: Major Tasks in Data Pre-processing, Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization.

Module- II **[10]**

Association Analysis: Basic Concepts, Market Basket Analysis, Apriori Algorithm, FP-growth, From Association Analysis to Correlation Analysis, Pattern Mining in Multilevel Associations and Multidimensional Associations.

Module- III **[10]**

Classification: Basic Concepts, Decision Tree Induction, Bayes Classification Methods, Rule-Based Classification, Metrics for Evaluating Classifier Performance, Ensemble Methods, Multilayer Feed- Forward Neural Network, Support Vector Machines, k-Nearest-Neighbor Classifiers.

Module- IV **[10]**

Cluster Analysis: Requirements for Cluster Analysis, Overview of Basic Clustering Methods, Partitioning Methods-k-Means, k-Medoids, Hierarchical Methods-AGENES, DIANA, BIRCH, Density- Based

Method-DBSCAN, Outlier Analysis: Types of Outliers, Challenges of Outlier Detection, and Overview of Outlier Detection Methods

Module- V

[8]

Advanced Concepts: Web Mining- Web Content Mining, Web Structure Mining, Web Usage Mining, Spatial Mining- Spatial Data Overview, Spatial Data Mining Primitives, Spatial Rules, Spatial Classification Algorithm, Spatial Clustering Algorithms, Temporal Mining- Modeling Temporal Events, Time Series, Pattern Detection, Sequences, Temporal Association Rules.

Text Books:

1. Jiawei Han, Micheline Kamber, Jian Pei., Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann/Elsevier, 2012.
2. Margaret H Dunham, Data Mining Introductory and Advanced Topics, 2nd Edition, Pearson Education, India, 2006.

Reference Books:

1. Data Mining Techniques, Arun K Pujari, 3rd Edition, Universities Press.
2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne and Vipin Kumar, Introduction to Data Mining, 2nd Edition, Pearson Education India, 2021.
3. Amitesh Sinha, Data Warehousing, Thomson Learning, India, 2007.

ELECTRONIC RESOURCES:

- <https://www.geeksforgeeks.org/data-science/data-mining/>
- <https://nptel.ac.in/courses/106105174>
- https://onlinecourses.nptel.ac.in/noc21_cs06/preview
- https://www.tutorialspoint.com/data_mining/index.htm
- https://www.w3schools.com/ai/ai_data.asp

MATERIALS ONLINE:

1. Course template
2. Tutorial question bank
3. Tech talk and Concept Video topics

4. Open-ended experiments
5. Definitions and terminology
6. Assignments
7. Model question paper – I
8. Model question paper – II
9. Lecture notes
10. E-Learning Readiness Videos