

Name of Faculty	:	Faculty of Engineering & Technology
Name of Program	:	Master of Technology (M.Tech.) - Artificial Intelligence and Data Science
Course Code	:	1MAI03
Course Title	:	Data Mining and Warehousing
Type of Course	:	Professional Core
Year of Introduction	:	2023-24

Prerequisite	:	To learn large size data warehousing methods and their needs in business systems.
Course Objective	:	Recognize and analyze diverse data domains, understand data mining processes, apply mathematical data mining techniques cross-disciplinarily, and build and assess visualization systems.
Course Outcomes	:	At the end of this course, students will be able to:
	CO1	Identify a number of common data domains and corresponding analysis tasks, including multivariate data, networks, text and cartography
	CO2	Comprehend the key processes of data mining, data warehousing and knowledge discovery process
	CO3	Implement data mining techniques to solve problems in other discipline in a mathematical way
	CO4	Exercise building and evaluating visualization systems

Teaching and Examination Scheme

Teaching Scheme (Contact Hours)			Credits	Examination Marks				
L	T	P		Theory Marks		Practical Marks		Total Marks
SEE	CIA	SEE	CIA					
3	0	2	4	70	30	30	20	150

Legends: L-Lecture; T-Tutorial/Teacher Guided Theory Practice; P - Practical, C - Credit, SEE - Semester End Examination, CIA - Continuous Internal Assessment (It consists of Assignments/Seminars/Presentations/MCQ Tests, etc.)

Course Content

Unit No.	Topics	Teaching Hours	Weightage	Mapping with CO
1	Introduction: Introduction to Data Mining, Importance of Data Mining, Data Mining functionalities, Classification of Data mining systems, Data mining architecture, Major Issues in Data Mining, Applications of Data	08	12%	CO1

	Mining, Social impacts of data mining.			
2	Introduction to Data Warehouse and OLAP Technology: Introduction to Data Warehouse and OLAP Technology for Data Mining, Multidimensional data Model, Data warehouse Data Model, Data warehouse Architecture, Data warehouse Implementation, Development of Data Cube Technology, From Data warehousing to Data Mining. Data Pre-processing, Data cleaning, Data Integration and Transformation, Data reduction, Discretization and Concept Hierarchy Generation	12	12%	CO2
3	Data Mining primitives, Languages and System Architectures, Concept description: Characterization and Comparison, Analytical Characterization, Mining Class Comparison.	06	16%	CO2
4	Association Rule Mining: Association Rule Mining Mining of Single dimensional Boolean association rules, Multilevel association rules and Multidimensional association rules, Correlation Analysis, Constraint based association Mining	08	10%	CO3
5	Classification and Predication: Basic issues regarding classification and predication, Classification by Decision Tree, Bayesian classification, classification by back propagation, Associative classification, Prediction, Classifier accuracy	06	20%	CO3
6	Cluster Analysis: Cluster Analysis, basic issues, clustering using partitioning methods, Hierarchical methods, Density based methods, Grid based methods and model based methods, Algorithms for outlier analysis.	10	15%	CO4
7	Mining complex Types of data: Multidimensional analysis and descriptive mining of complex data objects, Introduction to spatial mining, multimedia mining, temporal mining, text mining and web mining with related algorithms.	10	15%	CO4

Suggested Distribution of Theory Marks Using Bloom's Taxonomy						
Level	Remembrance	Understanding	Application	Analyse	Evaluate	Create
Weightage	20	30	30	20	0	0

NOTE: This specification table shall be treated as a general guideline for the students and the teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Suggested List of Experiments/Tutorials

Sr. No.	Name of Experiment/Tutorial	Teaching Hours
1	Install software WeKA	04
2	Install software XML Miner	04
3	Create one real time example for classification and prediction	06
4	Implement clustering algorithm	04
5	Implement any real time example for classification and accuracy algorithm	04
6	OLAP in standard SQL Server	04
7	MS Excell & Datamining plugin	04

Major Equipment/ Instruments and Software Required

Sr. No.	Name of Major Equipment/ Instruments and Software
1	WeKA
2	XML Miner
3	Orange
4	R

Suggested Learning Websites

Sr. No.	Name of Website
1	https://www.javatpoint.com/data-mining-cluster-vs-data-warehousing
2	https://www.researchgate.net/publication/325216437_Good_references_on_data_mining_and_analytics
3	https://www.topcoder.com/thrive/articles/data-warehousing-and-data-mining
4	https://www.oreilly.com/library/view/data-mining-for/9780470526828/bi01.html

Reference Books

Sr. No.	Name of Reference Books
1	Introduction to data mining by Pang - Ning Tan, Michel Steinbach, Vipin kumar - Pearson
2	Data Mining by Vikram Puri And P.RadhaKrishana -Oxfrod Press.
3	Mordern Data Warehousing, Data Mining and Visualization by George M. Marakas - Pearson.
4	Data Mining concepts and Techniques by Jiawei Han, Micheline Kamber -Elsevie