

<b>Name of Faculty</b>	:	Faculty of Computer Science & Applications
<b>Name of Program</b>	:	Master of Computer Application in Data Science
<b>Course Code</b>	:	2MDS02
<b>Course Title</b>	:	Data Mining
<b>Type of Course</b>	:	Professional core
<b>Year of Introduction</b>	:	2023-24

<b>Prerequisite</b>	:	Knowledge of basic programming and statistics
<b>Course Objective</b>	CO1	It presents methods for mining frequent patterns, associations, and correlations.
	CO2	It then describes methods for data classification and prediction, and data-clustering approaches.
	CO3	It covers mining various types of data stores such as spatial, textual, multimedia, streams
<b>Course Outcomes</b>	:	Student can
	1	Ability to understand the types of the data to be mined and present a general classification of tasks and primitives to integrate a data mining system.
	2	Evaluate the accuracy of supervised and unsupervised models and algorithms.
	3	Choose and employ suitable data mining algorithms to build analytical applications
	4	Discover the role played by data mining in various fields.
	5	Extract interesting patterns from large amounts of data.

#### Teaching and Examination Scheme

Teaching Scheme (Contact Hours)			Credits	Examination Marks				
L	T	P		SEE	CIA	SEE	CIA	Total Marks
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	<b>Data Mining:</b> Data-Types of Data-, Data Mining Functionalities- Interestingness Patterns- Classification of Data Mining systems- Data mining Task primitives - Integration of Data mining system with a Data warehouse-Major issues in Data Mining-Data Preprocessing.	07	20%	CO1

2	<b>Association Rule Mining:</b> Mining Frequent Patterns–Associations and correlations – Mining Methods– Mining Various kinds of Association Rules– Correlation Analysis– Constraint based Association mining. Graph Pattern Mining, SPM.	07	20%	CO2 CO5
3	<b>Classification:</b> Classification and Prediction – Basic concepts–Decision tree induction–Bayesian classification, Rule-based classification, Lazy learner.	07	20%	CO1 CO2
4	<b>Clustering and Applications:</b> Cluster analysis–Types of Data in Cluster Analysis–Categorization of Major Clustering Methods–Partitioning Methods, Hierarchical Methods–Density-Based Methods, Grid-Based Methods, Outlier Analysis.	07	20%	CO3 CO5
5	<b>Advanced Concepts:</b> Basic concepts in Mining data streams–Mining Time-series data–Mining sequence patterns in Transactional databases–Mining Object- Spatial- Multimedia-Text and Web data – Spatial Data mining- Multimedia Data mining–Text Mining- Mining the World Wide Web.	08	20%	CO3 CO4

<b>Suggested Distribution of Theory Marks Using Bloom's Taxonomy</b>						
<b>Level</b>	Remembrance	Understanding	Application	Analyse	Evaluate	Create
<b>Weightage</b>	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**

<b>Sr. No.</b>	<b>Name of Experiment/Tutorial</b>	<b>Teaching Hours</b>
1	Demonstration of preprocessing on dataset student.arff	01
2	Demonstration of preprocessing on dataset labor.arff	01
3	Demonstration of Association rule process on dataset contactlenses.arff using apriori algorithm.	01
4	Demonstration of Association rule process on dataset test.arff using apriori algorithm.	02
5	Demonstration of classification rule process on dataset student.arff using j48 Algorithm.	01
6	Demonstration of classification rule process on dataset employee.arff using j48 algorithm.	01
7	Demonstration of classification rule process on dataset employee.arff using id3 algorithm.	01
8	Demonstration of classification rule process on dataset employee.arff	01

	using naïve bayes algorithm.	
9	Demonstration of clustering rule process on dataset iris.arff using simple k-means.	02
10	Demonstration of clustering rule process on dataset student.arff using simple kmeans.	01

**Major Equipment/ Instruments and Software Required**

Sr. No.	Name of Major Equipment/ Instruments and Software
1	Anaconda
2	I3/ I5 processor; 8GB RAM; 250GB HDD

**Suggested Learning Websites**

Sr. No.	Name of Website
1	<a href="https://www.javatpoint.com/data-mining">https://www.javatpoint.com/data-mining</a>
2	<a href="https://www.geeksforgeeks.org/data-mining/">https://www.geeksforgeeks.org/data-mining/</a>

**Reference Books**

Sr. No.	Name of Reference Books
1	Data Mining – Concepts and Techniques – Jiawei Han & Micheline Kamber, 3rd Edition Elsevier.
2	Data Mining Introductory and Advanced topics – Margaret H Dunham, PEA
3	Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques (Second Edition), Morgan Kaufmann, 2005.