



Programme	B.Sc. IT Honours (Data Science)			Branch	Computer Applications				
Semester	V			Version	1.0.0.0				
Effective from Academic Year		2026-27		Effective for the batch Admitted in		June 2024			
Subject code	U75A3DM		Subject Name		DATA MINING				
Teaching scheme					Examination scheme(Marks)				
(Per week)	Lecture (DT)		Practical (Lab.)		Total	CCE	SEE	Total	
	L	TU	P	TW					
Credit	2	-	2	-	4	Theory	50	50	100
Hours	2	-	4	-	6				

Objective:

Students will be able to understand basic concepts and techniques of data mining and develop skills to solve problems using data mining tools

Pre-requisites:

Students should have basic knowledge of programming (preferably WEKA), SQL, and fundamental statistics. Familiarity with data structures, algorithms, and basic concepts of machine learning

Learning Outcome:

Name of CO	Description
CO1	Understand the data mining concepts and techniques
CO2	Understand how to handle missing and noisy data and apply normalization, feature selection, and extraction techniques using tools like WEKA
CO3	Develop classification models using decision trees, Bayesian classifiers, and rule-based classifiers and assess prediction accuracy.
CO4	Use algorithms like K-Means and hierarchical clustering and generate association rules
CO5	Analyze big data, text mining, and web mining concepts and perform practical case studies using open-source datasets and tools.

Mapping of CO and PO:

Cos	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	2	2	1	2	1	1	2	1	0	2	0
CO2	2	3	2	2	3	1	1	1	1	1	0	0
CO3	2	3	3	2	3	0	0	0	0	0	1	1
CO4	2	2	3	2	3	0	0	0	0	0	1	1
CO5	2	2	2	3	3	1	2	2	2	2	2	2

Content:

Unit	Content	Hrs.
1	Introduction to Data Mining: Definition, Need and Applications of Data Mining, KDD Process (Knowledge Discovery in Databases), Data Mining vs Machine Learning, Types of Data and Patterns	06
2	Data Preprocessing and Transformation: Data Cleaning, Integration, Reduction, and Transformation, Normalization Techniques Handling Missing Values and Noisy Data, Feature Selection and Extraction Techniques	06
3	Classification and Prediction: Classification by Decision Tree, Bayesian classification, Associative classification, Prediction,	06

	Statistical Based Algorithms, Rule-Based Classification	
4	Clustering and Association Rule Mining: Similarity and Distance Measures, K-Means Algorithm, Hierarchical Algorithms, Partitioned Algorithms, Clustering Large Databases, Clustering with Categorical Attributes, Outlier Detection, Clustering high dimensional data, clustering Graph and Network data.	06
5	Advanced Topics and Applications Introduction to Big Data and Data Mining, Introduction to Web Mining and Text Mining, Dataset handling using WEKA. Case studies: Kaggle datasets,	06
Practical Content:		
List of practical specified by subject teacher based on above mentioned topics		
Reference Books:		
1	Data Mining: Concepts and Techniques by Jiawei Han & Micheline Kamber	
2	Data mining: Concepts, models, methods and algorithms by M. Kantardzic	
3	Introduction to Data Mining, Michael Steinbach, Anuj Karpatne, Vipin Kumar Pearson Education, 2nd Edition	
Web Reference:		
1	https://www.cs.waikato.ac.nz/ml/weka/documentation.html	
2	https://www.geeksforgeeks.org/data-mining/	
3	https://www.kaggle.com/	
MOOC/Certificate Course:		
1	https://www.coursera.org/specializations/data-mining	
2	https://nptel.ac.in/courses/106105174	
3	https://www.udemy.com/course/data-mining-with-python/	
Question Paper Scheme:		
	End Semester Examination Duration: (2 Hours Theory Examination)	
	Note for Examiner: - Q-1 Any Five out of Seven (25 Marks) Q-2 Any Two out of Three (06 Marks) Q-3 Mandatory question (05 Marks) Q-4 Any Two out of Three (08 Marks) Q-5 Any Two out of Three (06 Marks)	
	*The question paper must comprehensively address all Course Outcomes (COs), align with Bloom's Taxonomy levels, and ensure complete syllabus coverage	