Indraprastha College for Women

University of Delhi

WORK PLAN for ODD SEMESTER – 2025

Course Name:	**B. Sc. H Computer Science DSE**
Paper Title:	**Data Mining I**
Unique Paper Code:	**Not yet Available**
Semester:	**III**
Faculty(s):	**Dr. Sarabjeet Kaur**
Year:	**2025**

Unit No.	Learning Objective	Lecture No.	Topics to be Covered
1	Introduction to Data Mining: Motivation and Challenges for data mining,	1,2	1.1
1	Types of data mining tasks, Applications of data mining	3,4,5	1.2, 1.3, 1.4
1	Supervised vs. unsupervised techniques, Data measurements	6,7,8	1.4, 2.1
2	Data quality, Data preprocessing: aggregation, sampling	9, 10 11	2.2, 2.3.1, 2.3.2,
2	dimensionality reduction, feature subset selection	12 13 14	2.3.3 (introduction), 2.3.4 (introduction)
2	feature creation, variable transformation.	15 16 17	2.3.5 (introduction), 2.3.6 (Binarization and Discretization of Continuous attributes), 2.3.7,
2	Measures of similarity and dissimilarity (upto Pg 78)	18,19,20	2.4.2, 2.4.3 (excluding properties)
5	Classification: Naive Bayes classifier, Nearest Neighbour classifier	21 22 23	3 (up to 3.3.3), 3.4 (introduction)
5	decision tree, overfitting.	24 25 26	3.6, 6.3
5	confusion matrix, evaluation metrics and model evaluation	27 28 29	6.4, 6.11 (introduction, 6.11.2)
4	Association Rule Mining: Transaction data-set, frequent itemset, support measure, confidence of association rule, Apriori principle	30 31 32	4.1
4	Apriori algorithm,rule generation,	33 34	4.2 (upto 4.2.2)
3	Cluster Analysis: Basic concepts of clustering, measure of similarity	35 36 37	5.1.1, 5.1.2
			MID SEMESTER BREAK

3	types of clusters and clustering methods,	38 39 40	5.1.3 (well-separated and Density-based)
3	K-means algorithm	40 41	5.2 (upto Data in Euclidean Space),
3	measures for cluster validation, determine optimal number of clusters	42 43	5.5.1, 5.5.5
	REVISION	44 45	
	DISBERSAL OF CLASSES		

Unit	Contents/Syllabus
I	**Unit 1: Introduction to Data Mining:** Motivation and Challenges for data mining, Types of data mining tasks, Applications of data mining, Data measurements, Data quality, Supervised vs. unsupervised techniques
II	**Unit 2: Data Pre-Processing:** Data aggregation, sampling, dimensionality reduction, feature subset selection, feature creation, variable transformation
III	**Unit 3: Cluster Analysis:** Basic concepts of clustering, measure of similarity, types of clusters and clustering methods, K-means algorithm, measures for cluster validation, determine optimal number of clusters
IV	**Unit 4: Association Rule Mining:** Transaction data-set, frequent itemset, support measure, rule generation, confidence of association rule, Apriori algorithm, Apriori principle
V	**Unit 5: Classification:** Naive Bayes classifier, Nearest Neighbour classifier, decision tree, overfitting, confusion matrix, evaluation metrics and model evaluation.

S. No.	Name of Authors/Books/Publishers
1.	1. Introduction to Data Mining, Pang-Ning Tan, Michael St einbach, Vipin Kumar, Pearson Education.
2.	Data Mining: Concepts and Techniques, 3nd edition,Jiawei Han and Micheline Kamber
3.	Data Mining: A Tutorial Based Primer, Richard Roiger, Michael Geatz, Pearson Education 2003.
4.	. Introduction to Data Mining with Case Studies, G.K. Gupta, PHI 2006

5	Insight into Data mining: Theory and Practice, Soman K. P., DiwakarShyam, Ajay V., PHI 2006
	Paper Components

Credits	**Lecture (L)**	**Tutorial (T)**	**Practical (P)**

Assessment Scheme

S.No.	**Component**	**Marking Scheme**	**Total Marks**
1	Internal Assessment		**30**
	• Assignment/Quiz/Project/ Presentation	12	
	• Class Test	12	
	• Attendance	6	
2.	Continuous Assessment (**Tutorial)**		NA
	• Activity 1		
	• Activity 2		
	• Attendance		
3.	Practical		40
	• Continuous Assessment	10	
	• End Term Written/Practical Exam	20	
	• Viva	10	
4.	End Semester Examination		