



ARSD College, University of Delhi

Model Course Handout/Lesson Plan

Course Name : B.Sc. (Hons.) Computer Science						
Semester	Course Code	Course Title	Lecture (L)	Tutorial (T)	Practical (P)	Credit (C)
VI	32347611	DSE-3 (BSCH17B)- DATA MINING-Theory	4	0	4	6
Teacher/Instructor(s)		Dr. Shalini Gupta				
Session		2021-22				

Course Objective:

- This course introduces data mining techniques and enables students to apply these techniques on real-life datasets.
- The course focuses on three main data mining techniques: Classification, Clustering and Association Rule Mining tasks.

Course Learning Outcomes:

On successful completion of the course, students will be able to do following:

1. Pre-process the data, and perform cleaning and transformation.
2. Apply suitable classification algorithm to train the classifier and evaluate its performance.
3. Apply appropriate clustering algorithm to cluster data and evaluate clustering quality
4. Use association rule mining algorithms and generate frequent item-sets and association rules

Lesson Plan:

Unit No.	Learning Objective	Lecture No.	Topics to be covered
1.	Introduction to Data Mining	1-2	What is Data Mining?
		3-4	Challenges in Data Mining
		5-6	Data Mining origins
		7-8	Data Mining tasks
2.	Data Mining Techniques	9-10	Data preprocessing, Types of data
		11-12	Data quality
		13-14	Data preprocessing
		15-16	Measures of similarity and dissimilarity
		17-18	Data cleaning, Data integration
3.	Introduction to Classification	19-20	Data reduction, Data transformation
		21-22	Classification introduction
		23-24	General approach to solving a classification

	technique		problem	
		25-26	Decision Tree induction	
		27-28	Evaluating the performance of a classifier	
		29-30	Rule based classifier	
		31-32	Nearest Neighbor classifiers	
		33-34	Introduction of Bayes Theorem	
		35-36	Bayesian Classifiers	
		37-38	Bayesian classifiers for continuous Attributes (Introduction) Class Assignment	
4.	Association Mining	Rule	39-40	Introduction
			41-42	Problem Definition
			43-44	Frequent Itemsets
			45-46	Frequent item set generation, Rule generation
			47-48	Class Test
5.	Clustering		49-50	Clustering Analysis
			51-52	Partition based clustering
			53-54	k-means clustering, k-medoids
			55-56	Hierarchical based clustering
			57-58	DBSCAN
			59-60	Revision and doubt session

Evaluation Scheme:

No.	Component	Duration	Marks
1.	Internal Assessment		25
	• Quiz		
	• Class Test		
	• Attendance		
	• Assignment		
2.	End Semester Examination	3 hrs	75

Details of the Course

Unit	Contents	Contact Hours
I	What Is Data Mining? Challenges, Data Mining Origins, Data Mining Tasks	8
II	Types of data, Data Quality, Aggregation, Sampling, Dimensionality reduction, Feature subset selection, Feature creation, Discretization, variable transformations, Dissimilarity among data objects, similarity among data objects	12
III	Preliminaries, General Approach to Solving a Classification Problem, Decision Tree Induction, Evaluating the Performance of a Classifier Rule Based Classifier, Nearest Neighbor Classifiers, Bayesian Classifiers, Alternative Metrics	18

IV	Problem definition, Frequent item set generation, Rule generation	10
V	Basic concepts of clustering analysis, k- Means, Agglomerative Hierarchical Clustering, DBSCAN	12
	Total	60
Suggested Books:		
Sl. No.	Name of Authors/Books/Publishers	Year of Publication/Reprint
1.	Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education.	2006
2.	Data Mining: Concepts and Techniques, 3rd edition, Jiawei Han and Micheline Kamber	2012
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
Mode of Evaluation:		Internal Assessment / End Semester Exam