



SWARNANDHRA

COLLEGE OF ENGINEERING & TECHNOLOGY

(AUTONOMOUS)

Accredited by National Board of Accreditation, AICTE, New Delhi, Accredited by NAAC with "A" Grade – 3.32 CGPA, Recognized under 2(f) & 12(B) of UGC Act 1956, Approved by AICTE, New Delhi, Permanent Affiliation to JNTUK, Kakinada Seetharampuram, W G DT, Narsapur -534280, (Andhra Pradesh)

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

TEACHING PLAN

Course Code	Course Title	Semester/Regulation	Branch	Contact Periods /Week	Academic Year	Date of commencement of Semester
20CD4TO1	DATA MINING AND DATA WAREHOUSING	IV	CSE-DS/AIDS	5	2023-24	03-01-2024

COURSE OUTCOMES

1	Summarize the architecture of data warehouse
2	Apply different preprocessing methods, Similarity, Dissimilarity measures for any given raw data.
3	Construct a decision tree and resolve the problem of model over fitting
4	Compare Apriori and FP-growth association rule mining algorithms for frequent itemset generation
5	Apply suitable clustering algorithm for the given data set

Data Warehouse and OLAP Technology

UNIT	Out Comes / Bloom's Level	Topic s No.	Topics/ Activity	Text Book/ Reference	Contact Hour	Delivery Method
I	CO1. Summarize the architecture of data warehouse	1.1	An Overview: Data Warehouse	T2	1	Chalk & Board
		1.2	A Multidimensional Data Model	T2	1	
		1.3	Data Warehouse Architecture	T2	1	
		1.4	Data Warehouse Implementation	T2	1	
		1.5	From Data Warehousing to Data Mining.	T2	1	
Revision of human factors					1	PPT
Total					07	

Data Mining

II	CO2. Apply different preprocessing methods, Similarity, Dissimilarity measures for any given raw data.	2.1.	Introduction to Data Mining	T1	1	Chalk & Board ppt
		2.2	Motivating challenges	T1	1	
		2.3	The origins of Data Mining	T1	1	
		2.4	Data Mining Tasks	T1	1	
		2.5	Types of Data	T1	1	
		2.6	Data Quality	T1	1	
		2.7	Data Preprocessing	T1	1	
		2.7.1	Aggregation	T1	1	
		2.7.2	Sampling	T1	1	
		2.7.3	Dimensionality Reduction	T1	1	
		2.7.4	Feature Subset Selection	T1	1	
		2.7.5	Feature creation	T1	1	
		2.7.6	Discretization and Binarization	T1	1	

		2.7.7	Variable Transformation	T1	1		
		2.8	Measures of Similarity and Dissimilarity	T1	1		
		Revision of human factors				1	
					Total	16	
Classification & Model Overfitting							
III	CO3. Construct a decision tree and resolve the problem of model overfitting	3.1	Basic Concepts	T1	1	Chalk & Board	
		3.2	General Approach to solving a classification problem	T1	1		
		3.3	Decision Tree Induction	T1	1		
		3.3.1	Working of Decision Tree	T1	1		
		3.3.2	Building a decision tree	T1	1		
		3.3.3	Methods for expressing an attribute test conditions	T1	1		
		3.3.4	Measures for selecting the best split	T1	1		
		3.3.5	Algorithm for decision tree induction.	T1	1		
		3.4	Model Over fitting		1		
		3.4.1	Due to presence of noise	T1	1		
		3.4.2	Due to lack of representation samples	T1	1		
		3.5	Evaluating the performance of classifier	T1	1		
		3.5.1	Holdout method	T1	1		
		3.5.2	Random sub sampling	T1	1		
		3.5.3	Cross-validation	T1	1		
		3.5.4	Bootstrap	T1	1		
		3.6	Bayes Theorem	T1	1		
		3.6.1	Naïve Bayes Classifier	T1	1		
						1	PPT
					Total	19	
Association Analysis							
IV	CO4. Compare Apriori and FP-growth association rule mining algorithms for frequent itemset generation.	4.1	Basic Concepts and Algorithms	T1	1	Chalk & Board PowerPoint presentations	
		4.2	Problem Definition	T1	1		
		4.3	Frequent Item Set Generation	T1	1		
		4.3.1	Apriori Principle	T1	1		
		4.3.2	Apriori Algorithm	T1	1		
		4.4	Rule Generation	T1	1		
		4.5	Compact Representation of Frequent Item sets	T1	1		
		4.6	FP-Growth Algorithm	T1	1		
						1	
					Total	09	

Cluster Analysis

V	C05: Apply suitable clustering algorithm for the given data set	5.1	Basic Concepts and Algorithms	T1	1	Chalk & Board Power point presentations		
		5.2	Overview, What Is Cluster Analysis?	T1	1			
		5.2.1	Different Types of Clustering	T1	1			
		5.2.2	Different Types of Clusters	T1	1			
		5.3	K-means: The Basic K-means Algorithm	T1	1			
		5.3.1	K- means Additional Issues	T1	1			
		5.3.2	Bisecting K-means	T1	1			
		5.3.3	Strengths and Weaknesses	T1	1			
		5.4	Agglomerative Hierarchical Clustering	T1	1			
		5.4.1	Basic Agglomerative Hierarchical Clustering Algorithm DBSCAN	T1	1			
		5.4.3	Traditional Density Center-Based Approach	T1	1			
		5.4.4	DBSCAN Algorithm	T1	1			
		5.4.5	Strengths and Weaknesses	T1	1			
		Discussion of previous year question paper						1
		Total					14	
CUMULATIVE PROPOSED PERIODS					65			

Text Books:

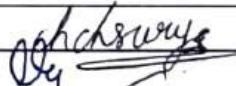
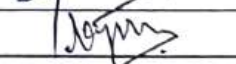


S. No.	AUTHORS, BOOK TITLE, EDITION, PUBLISHER, YEAR OF PUBLICATION
T1	Introduction to Data Mining : Pang-Ning Tan & Michael Steinbach, Vipin Kumar, Fifth Impression, Pearson, 2015.
T2	DataMining concepts and Techniques, 3 rd Edition, Jiawei Han, Michel Kamber, Elsevier, 2011


Reference Books:

S. No.	AUTHORS, BOOK TITLE, EDITION, PUBLISHER, YEAR OF PUBLICATION
R1	Data Mining Techniques and Applications: An Introduction, Hongbo Du Cengage Learning, 2010
R2	Data Mining : Introductory and Advanced topics : Dunham, First Edition, Pearson, 2020
R3	Data Warehousing Data Mining & OLAP, Alex Berson, Stephen Smith, TMH, 2008
R4	Data Mining Techniques, Arun K Pujari, Universities Press, 20011

Web Details:

1	NPTEL Online Course on Data Mining : https://onlinecourses.nptel.ac.in/noc18_cs14/preview
---	--

	Name	Signature
i. Faculty	Ch. Chandrika Surya	
ii. Course Coordinator	Dr G. Sudhakar	
iii. Module Coordinator	K. Jai Prakash	
iv. Program Coordinator	Dr B. Rama Krishna	


Principal