

Subject Code: 1ET4050502	Subject Title: DATA MINING AND WAREHOUSING
Pre-Requisite	The knowledge of programming languages like Python, Perl, C/C++, SQL, and Java and basics of database

RATIONALE:

COURSE OBJECTIVE:

TEACHING AND EVALUATION SCHEME:

Teaching Scheme (Hours per week)				Evaluation Scheme (Marks)			
Lecture	Tutorial	Practical	Credit	Theory		Practical	
				University Assessment	Continuous Assessment	University Assessment	Continuous assessment
3	-	2	4	60	40	20	30

DETAILS OF COURSE CONTENTS:

Sr. NO.	TOPIC	Total Hrs.	WEIGHTAGE
1	Fundamentals of data mining 1.1 Basics of Data Mining 1.2 History, strategies, techniques, applications 1.3 challenges of data mining, Future of data mining 1.4 Types of Data 1.4.1 Database Data 1.4.2 Data Warehouses 1.4.3 Transactional Data 1.4.4 Other Kinds of Data	07	16%
2	Objects, Attributes & Statistical Description of Data 2.1 Data Attribute 2.1.1 Nominal Attributes 2.1.2 Binary Attributes 2.1.3 Ordinal Attributes 2.1.4 Numeric Attributes 2.1.5 Discrete versus Continuous Attributes 2.2 Mean, Median, and Mode 2.3 Measuring the Dispersion of Data: Range, Quartiles, Variance, Standard Deviation, and Interquartile Range	08	19%
3	Data Preprocessing 3.1 Preprocess the Data 3.2 Major Tasks in Data Preprocessing 3.3 Data Cleaning 3.3.1 Missing Values	08	19%

	3.3.2 Noisy Data 3.3.3 Data Cleaning as a Process 3.4 Data Integration		
4	Classification 4.1 Decision tree 4.2 Probability based solving 4.3 Concepts of Clustering	08	20%
5	Data Warehouse 5.1 Data Warehouse 5.2 Differences between Operational Database Systems and Data Warehouses	06	14%
6	Data Warehouse 6.1 Basic of WEKA 6.2 Installing WEKA 6.3 WEKA data file format 6.4 Data visualization in WEKA 6.5 Data filtering 6.6 Using the concepts of data mining with WEKA	05	12%

COURSE OUTCOME:

The course content should be taught and implemented with an aim to develop different skills leading to the achievement of the following competencies and course learning outcomes:

1. To understand the concept of Data Mining & its attributes
2. To apply the concept of data mining components and techniques in designing data mining systems
3. To Learn how to pre-process data before applying data mining techniques
4. Able to Choose the data-mining task (classification, regression, clustering, Association Rules etc.)
5. Install and Configure WEKA Tool

The practical should be carried out in such a manner that students are able to acquire different learning outcomes in cognitive, psychomotor and affective domain to demonstrate course learning outcomes.

LIST OF PRACTICAL:

SR. No.	Experiment
1	Perform practical about basic python programming.
2	Study and perform basic libraries like NumPy, Pandas, Matplotlib in python.
3	Demonstrate the use of ARFF files taking input and display the output of the files.
4	Create your own excel file. Convert the excel file to .csv format and prepare it as ARFF files.

5	Perform Preprocessing, Classification techniques on any dataset.
6	Perform Clustering technique on dataset.
7	Classify the dataset using decision tree.
8	Perform Association technique on any dataset.
9	Install and Configure WEKA Tool
10	Demonstration of Weka Explorer, Mining techniques and Attribute Relation File Format (ARFF).

LIST OF REFERENCES:

Sr.no.	Title	Author	Publication
1	Data Mining Concepts and Techniques	Jiawei Han and Micheline Kamber	Kaufmann Publishers, 2011
2	Data Mining Techniques	Arun K Pujari	Orient Longman Publishers
3	Fundamentals of Data Warehouses	M.Jarke, M Lenzerni	
4	Principles of Data Mining	David Hand, Heikki Mannila, Padhraic Smyth,	PHI
5	Data Mining: Methods and Techniques	A B M Shawkat Ali, Saleh A, Wasimi	CENGAGE Learning

LINK OF LEARNING WEB RESOURCE:

1	http://www.tutorialspoint.com/data_mining/
2	https://www.guru99.com/data-mining-tutorial.html
3	https://www.javatpoint.com/data-mining