Format-3

IIMTU-NEP IMPLEMENTATION

Second Year/ IV Semester

Programme: Certificate **Class: BCA**		Year: Second Semester: Fourth
Credits Theory:4Cr	**Subject:** Data Mining	
Course Code: BCA-NEP-404	**Title:** Data Mining	
Course Objectives: CO1: To introduce students to basic applications, concepts, and techniques of data mining. CO2: To develop skills for using recent data mining software to solve practical problems in a variety of disciplines. CO3: To extract knowledge from data repository for data analysis, frequent pattern, classification and prediction. CO4: Understand and implement classical models and algorithms in data warehouses and data mining CO5: Master data mining techniques in various applications like social, scientific and environmental context.		
Nature of Paper: Discipline Specific Elective		
Minimum Passing Marks/Credits: 40 (ISE +ESE)		
L:4 T:0 P:4(In Hours/Week) Theory - 1 Hr. = 1 Credit Practical- 2 Hrs.=1Credit(4Hrs./Week=4Credits)		

Unit	**Contents**	**No. of Lectures Allotted**
I	**Introduction:** Data Mining - Overview, Motivation, Definition & Functionalities, Major issues in Data Mining, Integration of Data Mining System with Data Warehouse System. **Data Preprocessing**: Descriptive Data Summarization, Data Cleaning-Missing Values, Noisy Data, Data Integration and Transformation, Data Reduction-Data Cube Aggregation, Attribute Subset Selection, Dimensionality Reduction, Numerosity Reduction, Discretization and Concept Hierarchy.	12

II	**Association Rules:** Introduction, Frequent Itemsets, Closed Itemsets, Methods to Discover Association Rules, Apriori Algorithm, Multilevel Association Rule Mining, and Rule Evaluation Metrics.	12
III	**Classification and Prediction:** Classification Techniques- Decision Tree, Rule-Based Classification, Bayesian Classification, k-Nearest-Neighbor Classifier, Linear Regression, Accuracy and Error Measures	12
IV	**Cluster Analysis:** Introduction, Types of Data, Partitioning Methods- k-Means and k-Medoids, Hierarchical Clustering- Chameleon, Density Based Methods-DBSCAN, OPTICS. Grid Based Methods- STING, Model Based Methods- Neural Network Approach, Outlier Analysis.	12
V	**Recent Trends and Applications:** Web Mining, Spatial Data Mining, Text Mining, Multimedia Data Mining, Applications of data mining in finance, business, social networks.	12

Text Books:

1. Jiawei Han, Jian Pei, Micheline Kamber, *"Data Mining: Concepts and Techniques"*, Elsevier.

Reference

1. Margaret H. Dunham, "*Data Mining: Introductory and Advanced Topics",* Pearson Education.
2. Arun K. Pujari, *"Data Mining Techniques"*, Universities Press
3. Pieter Adriaans&DolfZantinge, *"Data Mining"*, Pearson Education

Evaluation/Assessment Methodology

Max. Marks 100

1) Class tasks/ Sessional Examination 2) Presentations /Seminar 3)Assignments 4)Research Project Report Seminar On Research Project Report 5) ESE	15 10 75
Total:	100
Prerequisites for the course: DBMS	
Course Learning Outcomes: CO1: Understand the functionality of the various data mining and data warehousing component. CO2: Appreciate the strengths and limitations of various data mining and data warehousing models. CO3: Explain the analyzing techniques of various data CO4: Describe different methodologies used in data mining and data ware housing. CO5: Compare different approaches of data ware housing and data mining with various technologies.	

Dr. Deepa Sharma
Chairperson- CBCS Committee