DATA & WEB MINING

Module 1 **15Hrs**

Introduction to Data mining: Role Data in Data Mining, Data Mining functionalities, patterns in data mining, Type of patterns, Classification of Data Mining Systems, Major issues in Data Mining; **Mining Association Rules in Large Databases** : Association Rule Mining, Mining Single-Dimensional Boolean Association Rules from Transactional Databases, Mining Multilevel Association Rules from Transaction Databases, Mining Multidimensional Association Rules from Relational Databases and Data Warehouses, From Association Mining to Correlation Analysis, Constraint-Based Association Mining. **Classification and Prediction:** Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, Classification by Backpropagation, Classification Based on Concepts from Association Rule Mining, Other Classification Methods, Prediction, and Classifier Accuracy.**Cluster Analysis Introduction** : Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical methods, Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Outlier Analysis.

Module 2 **10Hrs**

Introduction to WWW, Information Retrieval and Web Search: Basic Concepts, IR models, Relevance Feedback, Evaluation Measures, Text and Web Page Pre-Processing, Link Analysis: Graph Mining, Social Network Analysis, Co-Citation and Bibliographic Coupling, Page Rank, HITS, Community Discovery, Web Crawling: Basic and Universal Crawlers, Structured Data Extraction: Wrapper Generation: Wrapper Induction, Automatic Wrapper Generation: Problems, String Matching and Tree Matching, Information Integration: Pre-Processing for Schema Matching, Domain and Instance-Level Matching.

Module 3 **5Hrs**

Opinion Mining: Sentiment Classification, Feature-Based Opinion Mining and Summarization, Opinion Search, Opinion Spam, Web Usage Mining: Data Collection and Pre-Processing, Data Modeling for Web Usage Mining, Discovery and Analysis of Web Usage Patterns, Privacy Preserving Data Mining: Issues and Solutions.

Text Books:

1. J. Han & M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2nd ed, 2006. (Module 1)
2. Bing Liu. Web Data Mining, Exploring Hyperlinks, Contents and Usage Data, Springer Publishers (Module 2 and Module 3)

References:

1. Margret H Dunham,Data Mining Introductory and advanced topics, Pearson Education, 6th ed,2009,
2. Shawkat Ali and Saleh Wasimi,Data Mining: Methods and Techniques, Cengage Learning, Indian Edition,2009,