UNIT I INTRODUCTION TO DATA MINING (DM) Motivation for Data Mining - Data Mining-Definition and Functionalities – Classification of DM Systems - DM task primitives - Integration of a Data Mining system with a Database or a Data Warehouse - Issues in DM – KDD Process. UNIT II DATA PRE-PROCESSING Data summarization, data cleaning, data integration and transformation, data reduction, data discretization and concept hierarchy generation, feature extraction, feature transformation, feature selection, introduction to Dimensionality Reduction, CUR decomposition. UNIT III CONCEPT DESCRIPTION Mining Frequent Patterns, Associations and Correlations What is concept description? - Data Generalization and summarization-based characterization - Attribute relevance - class comparisons, Basic concept, efficient and scalable frequent item-set mining methods, mining various kind of association rules, from association mining to correlation analysis, Advanced Association Rule Techniques, Measuring the Quality of Rules. UNIT IV CLASSIFICATION AND PREDICTION Classification vs. prediction, Issues regarding classification and prediction, Statistical-Based Algorithms, Distance-Based Algorithms, Decision TreeBased Algorithms, Neural Network-Based Algorithms, Rule-Based Algorithms, Combining Techniques, accuracy and error measures, evaluation of the accuracy of a classifier or predictor. Neural Network Prediction methods: Linear and nonlinear regression, Logistic Regression Introduction of tools such as DB Miner / WEKA / DTREG DM Tools. UNIT V CLUSTER ANALYSIS Clustering: Problem Definition, Clustering Overview, Evaluation of Clustering Algorithms, Partitioning Clustering -K-Means Algorithm, KMeans Additional issues, PAM Algorithm; Hierarchical Clustering – Agglomerative Methods and divisive methods, Basic Agglomerative Hierarchical Clustering, Strengths and Weakness; Outlier Detection, Clustering high dimensional data, clustering Graph and Network data.