

IT6702 DATA WARE HOUSING AND DATA MINING

UNIT I DATA WAREHOUSING

Data warehousing Components –Building a Data warehouse – Mapping the Data Warehouse to a Multiprocessor Architecture – DBMS Schemas for Decision Support – Data Extraction, Cleanup, and Transformation Tools –Metadata.

UNIT II BUSINESS ANALYSIS

Reporting and Query tools and Applications – Tool Categories – The Need for Applications – Cognos Impromptu – Online Analytical Processing (OLAP) – Need – Multidimensional Data Model – OLAP Guidelines – Multidimensional versus Multirelational OLAP – Categories of Tools – OLAP Tools and the Internet.

UNIT III DATA MINING

Introduction – Data – Types of Data – Data Mining Functionalities – Interestingness of Patterns – Classification of Data Mining Systems – Data Mining Task Primitives – Integration of a Data Mining System with a Data Warehouse – Issues –Data Preprocessing.

UNIT IV ASSOCIATION RULE MINING AND CLASSIFICATION

Mining Frequent Patterns, Associations and Correlations – Mining Methods – Mining various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining – Classification and Prediction - Basic Concepts - Decision Tree Induction - Bayesian Classification – Rule Based Classification – Classification by Back propagation – Support Vector Machines – Associative Classification – Lazy Learners – Other Classification Methods – Prediction.

UNIT V CLUSTERING AND TRENDS IN DATA MINING

Cluster Analysis - Types of Data – Categorization of Major Clustering Methods – K-means– Partitioning Methods – Hierarchical Methods - Density-Based Methods –Grid Based Methods – Model-Based Clustering Methods – Clustering High Dimensional Data - Constraint – Based Cluster Analysis – Outlier Analysis – Data Mining Applications.