



COMPUTER ENGINEERING



POSITIVE QUADRANT
TECHNOLOGIES
SERVING INFORMATION WORLDWIDE

SEM V

DATA WAREHOUSING AND MINING

Programming & development

Course Curriculum



DATA WAREHOUSING AND MINING SEM V

Module 1 : Data Warehousing Fundamentals

- Introduction to Data Warehouse
- Data warehouse architecture
- Data warehouse v/s Data Marts
- E-R Modeling v/s Dimensional Modeling
- Information Package Diagram
- Data Warehouse Schemas
 - Star Schema
 - Snowflake Schema
 - Flactless Fact Table
 - Fact Constellation Schema
 - Update to the dimension Tables
- Major Steps in ETL Process
- OLTP V/S OLAP
- OLAP operations
 - Slice
 - Dice
 - Rollup
 - Drilldown and Pivot



Module 2 : Data Mining

- Introduction to Data Mining
- Data Mining Task Primitives
- Architecture
- KDD process
- Issues in Data Mining
- Applications of Data Mining

Module 3 : Data Exploration

- Types of Attributes
- Statistical Description of Data



- Data Visualization

Module 4 : Data Pre-processing

- Descriptive data summarization
- Cleaning
- Integration & transformation
- Data reduction
- Data Discretization and Concept hierarchy generation

Module 5 : Classification

- Basic Concepts
- Decision tree induction
- Native Bayesian Classification
- Accuracy and Error measures
- Evaluation the Accuracy of a Classifier
 - Holdout & Random Subsampling
 - Cross Validation
 - Bootstrap

Module 6 : Clustering

- Types of data in Cluster analysis
- Partitioning Methods
 - k-Means
 - k-Medoids
- Hierarchical Methods
 - Agglomerative
 - Divisive

Module 7 : Mining frequent patterns and associations

- Market Basket Analysis
- Frequent Item sets
- Closed Item sets



- Association Rule
- Frequent Pattern Mining
- Apriori Algorithm
- Association Rule Generation
- Improving the Efficiency of Apriori
- Mining Frequent Itemsets without candidate generation
- Introduction to Mining Multilevel Association Rules and Mining Multidimensional Association Rules.

Module 8: Web Mining

- Introduction to Web Mining
- Web Content Mining
 - Crawlers
 - Harvest System
 - Virtual Web View
 - Personalization
- Web Structure Mining
 - Page Rank
 - Clever
 - Web Usage Mining

