**Data Warehousing & Mining Syllabus**

**Unit I**: Introduction: Fundamentals of data mining, Data Mining Functionalities,

Classification of Data Mining systems, Data Mining Task Primitives, Integration of

a Data Mining System with a Database or a Data Warehouse System, Major

issues in Data Mining. Data Preprocessing: Need for Preprocessing the Data,

Data Cleaning, Data Integration and Transformation, Data Reduction,

Discretization and Concept Hierarchy Generation.

**Unit II**: Data Warehouse and OLAP Technology for Data Mining: Data

Warehouse, Multidimensional Data Model, Data Warehouse Architecture, Data

Warehouse Implementation, Further Development of Data Cube Technology,

From Data Warehousing to Data Mining Data Cube Computation and Data

Generalization: Efficient Methods for Data Cube Computation, Further

Development of Data Cube and OLAP Technology, Attribute-Oriented Induction.

Unit III: Mining Frequent Patterns, Associations and Correlations: Basic

Concepts, Efficient and Scalable Frequent Item set Mining Methods, Mining

various kinds of Association Rules, From Association Mining to Correlation

Analysis, Constraint-Based Association Mining

**Unit IV**: Classification and Prediction: Issues Regarding Classification and

Prediction, Classification by Decision Tree Induction, Bayesian Classification,

Rule-Based Classification, Classification by Back propagation, Support Vector

Machines, Associative Classification, Lazy Learners, Other Classification Methods,

Prediction, Accuracy and Error measures, Evaluating the accuracy of a Classifier

or a Predictor, Ensemble Methods

**Unit V**: 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, Clustering High-Dimensional Data, Constraint-Based Cluster Analysis,

Outlier Analysis.

**Unit VI**: Mining Streams, Time Series and Sequence Data: Mining Data

Streams, Mining Time-Series Data, Mining Sequence Patterns in Transactional

Databases, Mining Sequence Patterns in Biological Data, Graph Mining, Social

Network Analysis and Multirelational Data Mining.

**Text Book:**

1. Data Mining – Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann Publishers, Elsevier, 2nd Edition, 2006.

**Reference Books:**

1. Data Mining Techniques, Arun K Pujari, 3rd edition, Orient Blackswan/Universities Press, 2013.

2. Data Warehousing Fundamentals, Paulraj Ponnaiah, John Wiley & Sons, 2001.