

Soft Core	STS 557 : DATA MINING TECHNIQUES	No. of credits :3
-----------	-------------------------------------	-------------------

Course Outcomes:

CO1: Design data warehouse with dimensional modelling and apply OLAP operations.

CO2: Gain knowledge about basic concepts of Machine Learning and Identify machine learning techniques suitable for a given problem

CO3: Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining

CO4: Apply Dimensionality reduction techniques.

CO5: To assess the strengths and weaknesses of the different algorithms, identify the application area of algorithms and apply them.

CO6: Apply data mining techniques as well as methods in integrating and interpreting the data sets and improving effectiveness, efficiency and quality for data analysis.

Unit I

Data Mining – motivations and importance, Knowledge Discovery in Databases (KDD) process - search, induction, querying, approximation and compression. Kinds of data considered for data mining, basic data mining tasks, data mining issues, Data Mining models - predictive and descriptive, inter-connections between Statistics, Data Mining, Artificial Intelligence and Machine Learning. Applications of data mining. (10 hrs)

Unit II

Data marts, databases and data warehouses - OLTP systems, multidimensional models – data cubes, OLAP operations on data cubes, multidimensional schemas. Data pre-processing – data cleaning, data integration, data transformation and data reduction. Visualisation techniques for multidimensional data - scatter plot matrix, star plots, Andrews plots, Chernoff faces, parallel axis plots. (10 hrs)

Unit III

Supervised learning – classification and prediction, statistical classification-Linear Discriminants-Mahalanobis' linear discriminant, Fisher's linear discriminant; Bayesian classifier, Regression based classification, k-NN(nearest neighbour) classifier. Tree classifiers-decision trees, ID3 algorithm CART. (08 hrs)