# Syllabus: Predictive and Advanced Analytics

**OBJECTIVES**

The course enables students to:

• To learn, how to develop models to predict categorical and continuous outcomes, using such techniques as neural networks, decision trees, logistic regression, support vector machines and Bayesian network models.

• To know the use of the binary classifier and numeric predictor nodes to automate model selection.

• To advice on when and how to use each model. Also learn how to combine two or more models to improve prediction

**OUTCOMES**

The students will be able to:

• Understand the process of formulating business objectives, data selection/collection, preparation and process to successfully design, build, evaluate and implement predictive models for a various business application.
• Compare the underlying predictive modeling techniques.
• Select appropriate predictive modeling approaches to identify cases to progress with.
• Apply predictive modeling approaches using a suitable package such as SPSS Modeler

**Unit 1**: Introduction to Data Mining Introduction, what is Data Mining? Concepts of Data mining, Technologies Used, Data Mining Process, KDD Process Model, CRISP – DM, Mining on various kinds of data, Applications of Data Mining, Challenges of Data Mining.

**Unit 2**: Data Understanding and Preparation Introduction, Reading data from various sources, Data visualization, Distributions and summary statistics, Relationships among variables, Extent of Missing Data. Segmentation, Outlier detection, Automated Data Preparation, Combining data files, Aggregate Data, Duplicate Removal, Sampling DATA, Data Caching, Partitioning data, Missing Values.

**Unit 3**: Model development & techniques Data Partitioning, Model selection, Model Development Techniques, Neural networks, Decision trees, Logistic regression, Discriminant analysis, Support vector machine, Bayesian Networks, Linear Regression, Cox Regression, Association rules.

**Unit 4**: Model Evaluation and Deployment Introduction, Model Validation, Rule Induction Using CHAID, Automating Models for Categorical and Continuous targets, Comparing and Combining Models, Evaluation Charts for Model Comparison, MetaLevel Modeling, Deploying Model, Assessing Model Performance, Updating a Model.

**Text Book**: Predictive & Advanced Analytics (IBM ICE Publication)