

## **Module 1**

What is Machine Learning?

- Why Machine Learning?
- Machine Learning
- Statistics in Machine Learning
- Supervised and Unsupervised Learning

## **Module 2**

The Machine Learning Process

- Machine Learning Framework
- Machine Learning Approaches
- Machine Learning Techniques
- Machine Learning Algorithms
- Machine Learning Process
- CRISP-DM

## **Module 3**

Exploratory Data Analysis

- Exploratory Data Analysis (EDA)
- Sampling
- Data Profiling
- Descriptive Statistics
- Data Relationships
- Outliers and Anomalies
- Important Variables
- Output and Interpretation
- Feature Selection Methods

## **Module 4**

Models and Algorithms

- The Anatomy of a Model
- Classification
  - Decision Trees
  - Nearest Neighbor
  - Probability – Bayes Classification
  - Neural Networks
- Statistical Methods
- Clustering
- Association
- Anomaly Detection

- Application of Machine Learning Models

## **Module 5**

### Model Validation Techniques

- The Validation Process
- Fitting a Model
- Bias/Variance Tradeoff
- Validation Techniques
  - Confidence and Prediction Intervals
  - Statistical Significance
  - Classification Accuracy
  - Prediction Error Methods
  - Hold-out
  - Cross Validation