

Module 1 - Predictive Analytics Concepts

- What and Why of Predictive Analytics
 - Predictive Analytics Defined
 - Business Value of Predictive Analytics
- The Foundation for Predictive Analytics
 - Statistical Foundation
 - Data Mining Foundation
 - Machine Learning Foundation
 - Data Science Foundation
 - Describing Data Science
 - The Changing Landscape of Data Sources
- Predictive Analytics in BI Programs
 - Predictive Analytics in the BI Stack
 - Predictive Analytics in the BI Roadmap
 - Business, Technical, and Data Dependencies
- Becoming Analytics Driven
 - Business Driven
 - Grass Roots Driven
- Common Applications for Predictive Analytics
 - What Business Needs to Predict
- The Language of Predictive Analytics
 - Making Sense of the Terminology

Module 2 - Models and Statistics

- Predictive Models
 - What Are Models?
 - Using Models
 - Categories of Models
 - Model Development
 - Contributing Communities
- Descriptive Statistics
 - Variables
 - Frequencies and Summaries
 - Distribution and Skew
 - Relationships
 - Dependent and Independent Variables
- Inferential Statistics
 - Modeling the Population
- Probability
 - Estimating Likelihood
 - Calculating Probability
 - Calculating Odds
 - Logit Transformation
 - Probability Distribution Models

- Probability Distribution Examples

Module 3 - Regression Model Examples

- Regression Models
 - Overview
- Linear Regression Models
 - Overview
 - Example Business Case
 - Example Model
- Logistic Regression Models
 - Overview
 - Example Business Case
 - Steps for Creating the Model
 - Example Model
 - Predictors and Classifiers
 - Classifier Example

Module 4 - Building Predictive Models

- Model Building Processes
 - Data Mining Projects
 - CRISP-DM
 - SEMMA
 - CRISP-DM and SEMMA Compared
- Implementation and Operations Teams
 - A Team Effort
 - Roles and Responsibilities
- Predictive Techniques
 - Probability Values
 - Classification and Clustering
 - Segmentation
 - Association
 - Sequencing
 - Forecasting
- Technology
 - Features and Functions Overview
 - The Tools Landscape
- Model Building Algorithms
 - What and Why
 - Some Examples

Module 5 - Implementing Predictive Capabilities

- Introductory Concepts
 - Distribution View
 - Model Types View
 - Process View
 - Process Overview
- Business Understanding

- Activities and Deliverables
 - Pragmatics
- Data Understanding
 - Activities and Deliverables
 - Pragmatics
- Data Preparation
 - Activities and Deliverables
 - Pragmatics
- Modeling
 - Activities and Deliverables
 - Pragmatics
- Evaluation
 - Activities and Deliverables
 - Pragmatics
- Deployment
 - Activities and Deliverables
 - Pragmatics

Module 6 - Human Factors in Predictive Analytics

- Analytics Culture
 - Executive Buy-In
 - Strategic Positioning
 - Enterprise Range and Reach
 - Decision Processes
- People and Predictive Analytics
 - The Team
 - The Range of People
 - The Range of Knowledge
 - Readiness
 - Trust and Motivation
 - Expectations and Intent
 - Getting from Analytics to Impact
- Ethics and Predictive Analytics
 - Why Ethics Matters
 - Data and Ethics

Module 7 - Getting Started with Predictive Analytics

- Predictive Analytics Readiness
 - Readiness Checklist
 - Executive Commitment
 - Organizational Buy-In
 - Data Assets
 - Human Assets
 - Technology Assets
- Predictive Analytics Roadmap
 - A Plan to Evolve
 - An Evolving Plan

Day 2

Advanced Analytics: Leveraging Data Science and Machine Learning Techniques to Gain Data Insights

Offered by Deanne Larson through TDWI Onsite Education

Module 1 – What is Analytics?

- What is Analytics and Data Science?
- Statistics in Analytics
- Machine Learning
- Supervised and Unsupervised Learning

Module 2 – The Analytics Process

- Analytics Framework
- Analytics Approaches
- Analytics Techniques
- Analytics Algorithms
- Analytics 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
 - Support Vector Machine
- Prediction
 - Regression
 - Classifiers as Regression
- Ensemble Methods
- Clustering
- Association
- Anomaly Detection
- Application of Analytics 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

Module 6 – Summary

- Key Summary Points

Day 3

Hands-on Data Mining With R

Offered by Deanne Larson through TDWI Onsite Education

Attendees should have some coding experience, basic statistics, and will need to bring a laptop computer with RStudio installed prior to the session. In advance of the class attendees will receive detailed instructions for download and installation of RStudio.

Module 1 – Introduction to RStudio

- What is R?
- What is RStudio?
- Why use RStudio?
- Navigating RStudio
- What are packages?
- How to install packages
- Hands-on Exercises

Module 2 - R Basics

- R Math
- Data Types
- Working with Data
- Loading Data
- Writing Data
- Data Structures
- Hands-on Exercises

Module 3 – Introduction to Data Science in R

- Overview of Data Mining and Data Science
- Exploratory Data Analysis
- Base Graphics in R
- Linear Regression
- Logistic Regression
- Hands-on Exercises

Module 4 – Classification and Clustering Models in R

- Decision Trees
- Clustering
- Model Diagnostics
- Hands-on Exercises

Module 5 –Model Validation

- 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

Module 5 –Summary

- Skills Review