

## **Predictive Analytics:**

### **An Intensive Overview of Strategy, Application, and Best Practices for Data Mining**

---

#### **Session I - Strategic Overview**

This session introduces participants to the conceptual foundation of data mining projects. It is intended to give an overview of the types of problems that are appropriate for data mining, offer an approach that is realistic for applied model development in a business environment, and explain why traditional approaches are insufficient.

- Why Build Models?
- Data Mining: What it is, and What it isn't
- Matching Technologies and Problem Types
- Belief in Degree of Set Membership vs Right Answers
- Why Traditional Statistics is Not Enough
- Where Data Mining Works
- Defining Goals for Better Performance

#### **Session II – It's About the Data!**

Improved performance generally comes from one of two sources: getting more information content from the available data, or getting better data. This session explores sources of data, the strengths and weaknesses of various types of data, and techniques for manipulating data to extract information content.

- Types of Data
- Sources of Data
- Data Errors
- Missing Data
- Outliers
- Normalizing Data
- Derived Variables
- Variable Reduction
- Data Transformation Ideas

#### **Session III – Conceptual Introduction to Core Modeling Technologies**

The third session introduces a number of the advanced technologies commonly used in data mining. Discussion of these technologies in Level I is focused on conceptual understanding, and the strengths and weaknesses of the various methods. A technical drill-down into techniques for the various algorithms is reserved for the Level II offering.

- Linear Regression
- Logistic Regression
- Clustering
- Classification Trees
- Chaos
- Neural Networks
- Genetic Algorithms

**Session IV – Making Predictive Analytics Work!**

The final session addresses the Experimental Design and Project Definition aspects of Data Mining projects. These areas are where most data mining projects fall short of their potential.

- Experimental Design
- The Data Mining Process
- Case Studies