Data Preparation for Predictive Analytics

Learn via: **Classroom / Virtual Classroom / Online**

Duration: **1 Day**

https://bilginc.com/en/training/data-preparation-for-predictive-analytics-300-training/

Overview

This course will expose analytic practitioners, data scientists, and those looking to get started in predictive analytics to the critical importance of selecting, transforming, and properly preparing data ahead of model-building. The instructor will present the characteristics of varying data types, how to address data quality issues, and understanding data representations that are fitting to various project types.

Participants will learn that data outliers are often not errors in the data, but sometimes the data points of most interest. Live demonstrations will reinforce why problem context is required to understand how to deal with outliers and why undertreating extreme values can introduce model bias. This session will cover a wide range of data preparation exercises ranging from data sandbox construction to the creation of training, test, and validation data sets for model development.

Prerequisites

There are no prerequisites for this course.

Who Should Attend

- Analytic Practitioners
- Data Scientists
- IT Professionals
- Technology Planners
- Consultants
- Business Analysts
- Analytic Project Leaders

What You Will Learn

- Prepare a data sandbox for predictive analytics
- Detect and treat missing data and data quality issues
- Match data representations to fitting project types
- Construct various data transformations
- Handle data outliers without biasing model performance
- Build 'train / test / validation' data sets for model development
- Leave with resources, skills and plans to confidently process raw data for analytics