

Course Code: 0A039G

Course Title: Advanced Machine Learning Models Using IBM SPSS Modeler (V18.2)

## Description:

This course presents advanced models available in IBM SPSS Modeler. The participant is first introduced to a technique named PCA/Factor, to reduce the number of fields to a number of core factors, referred to as components or factors. The next topics focus on supervised models, including Support Vector Machines, Random Trees, and XGBoost. Methods are reviewed on how to analyze text data, combine individual models into a single model, and how to enhance the power of IBM SPSS Modeler by adding external models, developed in Python or R, to the Modeling palette.

## Objectives:

- Introduction to advanced machine learning models
- Taxonomy of models
- Overview of supervised models
- Overview of models to create natural groupings
- Group fields: Factor Analysis and Principal Component Analysis
- Factor Analysis basics
- Principal Components basics
- Assumptions of Factor Analysis
- Key issues in Factor Analysis
- Improve the interpretability
- Factor and component scores
- Predict targets with Nearest Neighbor Analysis
- Nearest Neighbor Analysis basics
- Key issues in Nearest Neighbor Analysis
- Assess model fit
- Explore advanced supervised models
- Support Vector Machines basics
- Random Trees basics
- XGBoost basics
- Introduction to Generalized Linear Models
- Generalized Linear Models
- Available distributions
- Available link functions
- Combine supervised models
- Combine models with the Ensemble node
- Identify ensemble methods for categorical targets
- Identify ensemble methods for flag targets
- Identify ensemble methods for continuous targets
- Meta-level modeling
- Use external machine learning models
- IBM SPSS Modeler Extension nodes
- Use external machine learning programs in IBM SPSS Modeler
- Analyze text data
- Text Mining and Data Science
- Text Mining applications

- Modeling with text data

## Prerequisites:

- Knowledge of your business requirements
- Required: IBM SPSS Modeler Foundations (V18.2) course (0A069G/0E069G) or equivalent knowledge of how to import, explore, and prepare data with IBM SPSS Modeler v18.2, and know the basics of modeling.
- Recommended: Introduction to Machine Learning Models Using IBM SPSS Modeler (V18.2) course (0A079G/0E079G), or equivalent knowledge or experience with the product about supervised machine learning models (CHAID, C&R Tree, Regression, Random Trees, Neural Net, XGBoost), unsupervised machine learning models (TwoStep Cluster), and association machine learning models such as APriori.

## Duration:

8 Hrs

## Topics:

Introduction to advanced machine learning models

- Taxonomy of models
- Overview of supervised models
- Overview of models to create natural groupings
- Group fields: Factor Analysis and Principal Component Analysis
- Factor Analysis basics
- Principal Components basics
- Assumptions of Factor Analysis
- Key issues in Factor Analysis
- Improve the interpretability
- Factor and component scores
- Predict targets with Nearest Neighbor Analysis
- Nearest Neighbor Analysis basics
- Key issues in Nearest Neighbor Analysis
- Assess model fit
- Explore advanced supervised models
- Support Vector Machines basics
- Random Trees basics
- XGBoost basics
- Introduction to Generalized Linear Models
- Generalized Linear Models
- Available distributions
- Available link functions
- Combine supervised models
- Combine models with the Ensemble node
- Identify ensemble methods for categorical targets
- Identify ensemble methods for flag targets
- Identify ensemble methods for continuous targets
- Meta-level modeling
- Use external machine learning models
- IBM SPSS Modeler Extension nodes
- Use external machine learning programs in IBM SPSS Modeler
- Analyze text data
- Text Mining and Data Science

- Text Mining applications
- Modeling with text data

**Audience:**

- Data scientists
- Business analysts
- Experienced users of IBM SPSS Modeler who want to learn about advanced techniques in the software