



MACHINE LEARNING Professional

Delivery mode: self-paced, onsite, live online

Course length: 2 day

Target Audience: Business Analyst, Data Scientist, Domain Expert

Prerequisite courses: none



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator

Practice building processes in RapidMiner that implement both supervised and unsupervised machine learning.

Overview

This course provides the opportunity to directly work on creating and modifying RapidMiner Processes to train, validate and apply models to new data. It covers the most common and useful techniques in machine learning.

Consider taking both *Applications & Use Cases Professional* as well as *Data Engineering Professional* before this course.

Course Objectives

- Classification and Regression
- Split Validation
- Scoring
- Correlations
- Feature Importance
- Clustering and Association Analysis



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator



Machine Learning Professional Topic Guide:

Review the Basics of Data Science (DS) and Machine Learning (ML).

Learn common terminology and basic types of DS and ML so you have the context to understand the rest of this course and the other courses.

Use common Machine Learning algorithms for Classification and Regression including:

k Nearest Neighbor, Naïve Bayes, Decision Tree, Linear Regression, Logistic Regression, General Linear Model, Neural Network.

Review the basics of Scoring and Deployment, and how to implement them.

Learn about validation including how to implement and interpret the results of Split Validation.

Learn when and how to inspect Correlations between attributes.

Practice basic Feature Engineering, and learn when it is needed.

This includes practice using and interpreting the results of Feature Importance and Weighting.

Practice implementing Clustering with k-Means and X-Means

Learn when and how to perform Association Analysis. This includes:

Understanding the data prep for a variety of potential data structures.

Identification of frequent item-sets with FP-Growth.

Rule generation with the Create Association Rules operator.

Interpretation of different criteria including support, confidence, and lift.

Be able to use Auto Model effectively.



MACHINE LEARNING Professional



Delivery mode: self-paced, onsite, live online
Course length: 2 days



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator

We designed our program with the most common user personas and the required areas of expertise for applied data science in mind:

