Machine Learning in R

Advanced Level Course

Duration: 1 Day (9am – 4pm)

Course Pre-Requisites: Knowledge of Programming in R, a foundational grasp of R syntax and an understanding of statistics is recommended for this course.

Course Overview

By the end of the course, attendees will be adept at extracting insights from datasets using cluster analysis and constructing predictive models via random forests. They will be competent in evaluating the predictive accuracy of models and leveraging these models for informed decision-making.

1. Introduction

Participants will be introduced to the realm of machine learning, with a focus on the differences between supervised and unsupervised learning paradigms. They will also gain a holistic view of the machine learning process.

2. Cluster Analysis

Students will grasp the core concepts and applications of cluster analysis, focusing on the k-means algorithm. They will comprehend how the algorithm operates, the nuances of data preparation, determining the right number of clusters, and practical implementation in R.
A significant portion of this section is dedicated to customer segmentation analysis, a prevalent business application of clustering.

3. Random Forests

This section introduces learners to the fundamentals of tree-based models, highlighting the distinctions between classification and regression trees. Students will understand the evolution from individual decision trees to the ensemble method of random forests. Key topics encompass ensemble learning with bagging, data preparation, splitting data into training and test sets, model training, accuracy assessment, optimisation, and prediction.

4. R Scripts in Power BI - Demonstration

This hands-on demonstration will guide students through the process of setting up R scripts in Power BI. They will learn to run R scripts within Power Query and create visually appealing reports and dashboards in Power BI using R visuals.