jumping rivers

Machine Learning with R

This two-day course is aimed at not only teaching an understanding of some of the most common machine learning techniques, but also the approach to implementing machine learning. During this course, attendees will learn how to define a problem and prepare data, the range of techniques available for solving common problems and the approaches to take to evaluate models and achieve the best results possible.

Machine Learning can be applied to data in a whole range of fields from Finance to Pharmaceutical, Retail to Marketing, Sports to Travel and many, many more! This course is aimed at anyone interested in applying machine learning methods to their data in order to: gain deeper insight, make better decisions or build data products.

Course Outline

- Introduction to analytics: a general introduction into analytics and some of the techniques that are in common use.
- Simple regression problems: simple and multiple linear regression and model diagnostics.
- Model selection and assessment: Cross validation and bootstrapping. Penalised regression and shrinkage.
- Classification: KNN, clustering, logistic regression, Linear Discriminant analysis and associated diagnostics.
- Advanced regression techniques: polynomial regression, splines, local regression, GAMs, trees and random forests.
- Workflow development: Throughout we will develop a workflow of training, testing and assessing models that can be extended to techniques not directly covered.

Learning Outcomes

By the end of the course participants will be able to...

- have a thorough understanding of popular analytical techniques practised in industry today
- assess model validity through bootstrapping and cross validation
- understand which technique applies to their own data
- efficiently and effectively analyse their own data using said techniques in R
- apply the developed workflow practice to other modelling techniques

Course Structure

- Day 1:
  - Introduction to analytics
  - Simple regression problems
  - Model selection and assessment
- Day 2:
  - Model selection and assessment
  - Classification
  - Advanced regression techniques

Prior Knowledge

It will be assumed that participants are familiar with R. For example inputting data, basic visualisation, basic data structures and use of functions. Attending the introduction to R course will provide a sufficient background, but the programming with R will be helpful.