Data Mining in R

Course Objectives

* How to configure the RStudio environment and load packages
* How to use R basics such as basic math, data types, vectors, and calling functions
* How to use advanced data structures such as data frames, lists, and matrices
* How to use R base graphic packages
* How to do exploratory data analysis
* How to use R to support basic statistics, correlation, and covariance
* How to use linear models such as linear regression and logistic regression
* How to use models such as decision trees, Random Forest, and K nearest neighbor
* How to use clustering models such as K-means

Course Structure

* Modules:
  + Introduction to RStudio
  + R Basics
  + Introduction to Data Mining in R
  + Classification and Clustering Models in R
  + Summary

Module 1 Overview

* Topics:
  + What is R?
  + What is RStudio?
  + Why use RStudio?
  + Navigating RStudio
  + What are packages?
  + How to install packages
  + Hands-on Exercises

Module 2 Overview

* Topics:
  + R Math
  + Data Types
  + Working with Data
  + Loading Data
  + Writing Data
  + Data Structures
  + Hands-on Exercises

Module 3 Overview

* Topics:
  + Overview of Data Mining and Data Science
  + Exploratory Data Analysis
  + Base Graphics in R
  + Linear Regression
  + Logistic Regression
  + Hands-on Exercises

Module 4 Overview

* Topics:
  + Decision Trees
  + Clustering
  + Model Diagnostics
  + Hands-on Exercises

Module 5 – Summary

* Skills Review