REVOLUTION ANALYTICS

Modeling in RRE: Overview and Agenda

Revolution Analytics

Overview

After completing this course, you will be able to:

- Conduct predictive analysis on your enterprise data using regression models.
- Implement models through embedded scoring functions in Revolution R Enterprise.
- Understand key concepts in coding big data functions efficiently.

Outline

Overview

The basic approaches we will cover:

- Linear Models
  - Linear Model with Ordinary Least Squares
  - Logistic Regression
  - Other Generalized Linear Models
- Model Specification and Prediction
  - Complex Formulas and Higher Order Terms
  - Stepwise Regression
  - Model Predictions and Cross-Validation

The Data

Throughout this course we will be using three data sets:

- Flight information regarding commercial flights, a subset of the 2009 airlines visualization competition 2007)
- Bank Marketing data set from the Machine Learning Repository at University of California, Irvine
- An internally created, randomly generated Churn data set for exercise purposes.

Outline

The Data: Airlines

The airlines dataset consists of information about a random subset of commercial flights.

Outline

The Data: Bank Marketing Data

The Bank Marketing Data Set, which we will refer to as the `Bank` data, concerns the relationship between direct marketing campaigns and subscription to a term deposit for a Portuguese bank.

S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier, 62:22-31, June 2014

Outline

The Data: Churn Data

In addition to the Bank Data, we will talk about Churn data, particularly for exercises.

One may imagine this data as that representing clients in a phone company, where variables such as `n.family.members` and `n.devices` refer to the number of family members and the number of devices of a particular client, respectfully.

We will explore more as we move forward…

Let’s get Started!

REVOLUTION ANALYTICS

Thank you

Revolution Analytics is the leading commercial provider of software and support for the popular open source R statistics language.

www.revolutionanalytics.com
1.855.GET.REVO
Twitter: @RevolutionR