Automating Business Modeling with the `AutoModelR` Package

Derek McCrae Norton[*]

March 31, 2008

Abstract

Many issues arise in the business environment, like other fields, which must be addressed. Many of these issues have been addressed in other fields separately, but need to be jointly addressed in a business environment. The objective of `AutoModelR` is to attempt to address these issues in an automatic manner.

The issues addressed are: Exploratory Data Analysis, Dimension Reduction, and Automatic Initial Modeling. To address these issues, a data set is passed to `AutoModelR` consisting of a dependant variable and one or more (often many more) independent variables. The first step is to remove variables which have zero variation, missing value percentages above a threshold, and variables with one unique value accounting for more than some threshold percentage. A report is then generated using `Sweave` which gives tables of descriptive statistics for numeric and factor data separately as well as graphical displays of the data. The next step is an application of a filter type dimension reduction to arrive a smaller data subset for initial modeling. The last step is to automatically fit various simple models determined by the type of dependent variable and report on those fits.

`AutoModelR` is an attempt to automate some of the repetitive steps in modeling, so that more time can be spent on advanced modeling.

---

[*]InterContinental Hotels Group, USA