Practical machine learning: Project description

Hector Corrada Bravo and Rafael A. Irizarry

February 2010

Description

The goal of your project is to apply methods we have discussed in class to an interesting dataset and report results. Alternatively, you can explore a method we have discussed in class in more depth. You will be paired with a classmate to complete the project.

Proposal

An initial project description/proposal is due Friday 2/26.

Please describe the project in general terms: describe the dataset and what makes it interesting; what is the goal of analyzing this dataset (predictive performance, interpretation, visualization); what are specific challenges in using this data (large number of predictors/features, large number of observations, difficult to predict task); propose in very general terms what type of methods you might want to use.

If you are working on a specific method, rather than analyzing a dataset, describe the method you will be working on; characteristics of the method you find interesting; what type of questions you want to answer about the method (statistical or computational); what are your goals in working with this method.

Make this description no longer than 1.5 typed pages. Submit by email with subject [project proposal].

Project Writeup

In the writeup please include the following:

1. A description of the dataset and task (you can build on what you wrote for the initial description).
2. A description of any transformation, filtering, or procedure you performed on the data before analysis

3. A description of the methods you used, including a discussion of why they are suitable to the task

4. A report of results, make sure you address the task described in 1 as fully as possible

5. A discussion of results, and possible future directions for subsequent analysis

Make the writeup no longer than 5 typed pages. Submit by email with subject [project report].

The due date for the project is March 29, 2010.