

Project 2 (Final Project!)

Meta Learning

April 29th, 2017

Overview

How to submit this: Submit your code (.R or .ipynb) and your predictions (.csv) through CMS by 11:59pm May 17th.

Objectives: Understand and learn how to implement an ensemble learning method..

The Dataset

We'll be using the same dataset as project 1, description is re-stated below.

In 2013, 1,010 people aged 15-30 from Bratislava, Slovakia filled out a 150 question survey inquiring about their music preferences, movie preferences, hobbies & interests, phobias, health habits, personality traits, views on life, opinions, and spending habits. Most of the questions were yes or no, and people answered how much they agree on a scale of 1 to 5. You can read more about the dataset [here](#) and download the question description and answers [here](#).

Deliverables

You've learned about different prediction algorithms like Linear Regression, Logistic Regression, Decision Trees, K Nearest Neighbors, and you've also learned about ensemble strategies like boosting, bagging, and stacking. This project is to build on the skills you used in project 1, **you are expected to predict the values for one specific column in responses.csv, using an ensemble method of your choice, made from the models you used in project 1.**

Scoring

Using the same baseline methodology as project 1, you're predicted column must be 10% above the baseline.