Assignment

Matteo Gambera

7 aprile 2020

Summary

One thing that people regularly do is quantify how much of a particular activity they do, but they rarely quantify how well they do it. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbbell of 6 participants. The goal of your project is to predict the manner in which they did the exercise. This is the "classe" variable in the training set. You may use any of the other variables to predict with. You should create a report describing how you built your model, how you used cross validation, what you think the expected out of sample error is, and why you made the choices you did. You will also use your prediction model to predict 20 different test cases.

Protocol

Reproducibility

the pseudo-random number generator seed was set at 1 for all code.

Model

Participants were asked to perform one set of 10 repetitions of the Unilateral Dumbbell Biceps Curl in five different fashions:

- Class A: exactly according to the specification,
- Class B: throwing the elbows to the front,
- Class C: lifting the dumbbell only halfway,
- Class D: lowering the dumbbell only halfway,
- Class E: throwing the hips to the front.

Class A corresponds to the specified execution of the exercise, while the other 4 classes correspond to common mistakes.

Cross-Validation

Will be performed by subsampling the training data set randomly without replacement into 2 subsamples: subTraining data (70% of the original Training data set) and subTesting data (30%). Our models will be fitted on the subTraining data set, and tested on the subTesting data. Once the most accurate model is choosen, it will be tested on the original Testing data set.

Data

The training data for this project are available here:

- **training**

The test data are available here:

- **testing**

The data for this project come from this source: **link**. If you use the document you create for this class for any purpose please cite them as they have been very generous in allowing their data to be used for this kind of assignment.

```
library(data.table)

## Data training
url_training <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"
training <- fread(url_training)

## Data testing
url_testing <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"
testing <- fread(url_testing)
```