Predicting the Manner the Exercise Is Being Performed

Sue,

August 4, 2016

Introduction

The main goal of this project is to predict the manner in which 6 participants performed some exercise. This is the “classe” variable in the training set. By doing some cross validation the best model is chosen and uses to predict the 20 cases in the test dataset.

Data and Some Exploratory Analysis

In this project, the goal is to use data from accelerometers on the belt, forearm, arm, and dumbell of 6 participants. They were asked to perform barbell lifts correctly and incorrectly in 5 different ways. More information is available here (http://groupware.les.inf.puc-rio.br/har). You can find the training and test data for this project here:

- Trainset (https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv)
- Testset (https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv)

Loading and Cleaning Data

There are 160 variables in the original dataset many of which contain large number of NAs. Some other variables are highly correlated so we have to clean the data before further analysis.

Omitting Variables with High Amount of Missing Values

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```