Prediction for Quality of Activity

ljc

July 25, 2015

Background

Using devices such as Jawbone Up, Nike FuelBand, and Fitbit it is now possible to collect a large amount of data about personal activity relatively inexpensively. These type of devices are part of the quantified self movement – a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. 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 dumbell of 6 participants. They were asked to perform barbell lifts correctly and incorrectly in 5 different ways. More information is available from the website here: http://groupware.les.inf.puc-rio.br/har (http://groupware.les.inf.puc-rio.br/har) (see the section on the Weight Lifting Exercise Dataset).

Data

The training data for this project are available here:
https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv
(https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv)

The test data are available here: https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv
(https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv)

Data source: http://groupware.les.inf.puc-rio.br/har (http://groupware.les.inf.puc-rio.br/har)

```
setwd("~/Documents/Coursera_Data_Science/08_PML")
#download.file(https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv, "pml-training.zip", method="curl")
#unzip("./pml-training.zip")
trng_data<- read.csv("./pml-training.csv")
#download.file(https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv, "pml-testing.zip", method="curl")
#unzip("pml-testing.zip")
test_data<- read.csv("./pml-testing.csv")

dim(trng_data)
```

```
## [1] 19622   160
```

```
#head(trng_data)
```

```
dim(test_data)
```

```
## [1]  20 160
```

```
#head(test_data)
```

Preprocess Data

Split data into Training and Test datasets

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.1.3
```

```
## Loading required package: lattice
## Loading required package: ggplot2
```

```
library(kernlab)
inTrain <- createDataPartition(y=trng_data$classe, p=0.75, list=FALSE)
training <- trng_data[inTrain,]
testing <- trng_data[-inTrain,]
dim(training)
```

```
## [1] 14718   160
```

```
dim(testing)
```

```
## [1] 4904  160
```

Analysis

Fit a model

```
#set.seed(12345)
#modelFit <- train(classe ~.,data=training, method="rpart")
#modelFit
```