Course Project - HAR

Steve Chiu

13 June 2016

Executive Summary

This proejct relates to the quantification of how well of a particular activity people do. The goal is to use data from accelerometers on the belt, forearm, arm, and dumbell of 6 participants and then predict the manner in which they did the exercise, which is the “classe” variable in the training set.

The 19622 obervations from the dataset was randomly divided into two portion, 3 quarters for the learning of the model (using Random Forest), and the remaining on for the probing of the model.

Finally the model is used to make predication on 20 different test cases.

Background

Human Activity Recognition - HAR - has emerged as a key research area in the last years and is gaining increasing attention by the pervasive computing research community. More information please refer to http://groupware.les.inf.puc-rio.br/har (http://groupware.les.inf.puc-rio.br/har).

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, 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.

Exploratory Analysis

We shall first look into the data structure of the dataset, of which an extract is as per below:

```
download.file("https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.cs
v", "training.csv")
download.file("https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv
", "testing.csv")
training <- read.csv("training.csv", header = TRUE)
testing <- read.csv("testing.csv", header = TRUE)
head(training,3)
```

```
##   X user_name raw_timestamp_part_1 raw_timestamp_part_2   cvtd_timestamp
## 1 1  carlitos           1323084231               788290 05/12/2011 11:23
## 2 2  carlitos           1323084231               808298 05/12/2011 11:23
## 3 3  carlitos           1323084231               820366 05/12/2011 11:23
##   new_window num_window roll_belt pitch_belt yaw_belt total_accel_belt
```

```
## 1         no          11      1.41       8.07    -94.4                 3
## 2         no          11      1.41       8.07    -94.4                 3
## 3         no          11      1.42       8.07    -94.4                 3
##   kurtosis_roll_belt kurtosis_picth_belt kurtosis_yaw_belt
## 1
## 2
## 3
##   skewness_roll_belt skewness_roll_belt.1 skewness_yaw_belt max_roll_belt
## 1                                                                      NA
## 2                                                                      NA
## 3                                                                      NA
##   max_picth_belt max_yaw_belt min_roll_belt min_pitch_belt min_yaw_belt
## 1             NA                         NA             NA
## 2             NA                         NA             NA
## 3             NA                         NA             NA
##   amplitude_roll_belt amplitude_pitch_belt amplitude_yaw_belt
## 1                  NA                   NA
## 2                  NA                   NA
## 3                  NA                   NA
##   var_total_accel_belt avg_roll_belt stddev_roll_belt var_roll_belt
## 1                   NA            NA                NA            NA
## 2                   NA            NA                NA            NA
## 3                   NA            NA                NA            NA
##   avg_pitch_belt stddev_pitch_belt var_pitch_belt avg_yaw_belt
## 1             NA                NA             NA           NA
## 2             NA                NA             NA           NA
## 3             NA                NA             NA           NA
##   stddev_yaw_belt var_yaw_belt gyros_belt_x gyros_belt_y gyros_belt_z
## 1              NA           NA         0.00            0        -0.02
## 2              NA           NA         0.02            0        -0.02
## 3              NA           NA         0.00            0        -0.02
##   accel_belt_x accel_belt_y accel_belt_z magnet_belt_x magnet_belt_y
## 1          -21            4           22            -3           599
## 2          -22            4           22            -7           608
## 3          -20            5           23            -2           600
##   magnet_belt_z roll_arm pitch_arm yaw_arm total_accel_arm var_accel_arm
## 1          -313     -128      22.5    -161              34            NA
## 2          -311     -128      22.5    -161              34            NA
## 3          -305     -128      22.5    -161              34            NA
##   avg_roll_arm stddev_roll_arm var_roll_arm avg_pitch_arm stddev_pitch_arm
## 1           NA              NA           NA            NA                NA
## 2           NA              NA           NA            NA                NA
## 3           NA              NA           NA            NA                NA
##   var_pitch_arm avg_yaw_arm stddev_yaw_arm var_yaw_arm gyros_arm_x
## 1            NA          NA             NA          NA        0.00
## 2            NA          NA             NA          NA        0.02
## 3            NA          NA             NA          NA        0.02
##   gyros_arm_y gyros_arm_z accel_arm_x accel_arm_y accel_arm_z magnet_arm_x
## 1        0.00       -0.02        -288         109        -123         -368
## 2       -0.02       -0.02        -290         110        -125         -369
## 3       -0.02       -0.02        -289         110        -126         -368
##   magnet_arm_y magnet_arm_z kurtosis_roll_arm kurtosis_picth_arm
## 1          337          516
## 2          337          513
```

```
## 3          344          513
##   kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm skewness_yaw_arm
## 1
## 2
## 3
##   max_roll_arm max_picth_arm max_yaw_arm min_roll_arm min_pitch_arm
## 1           NA            NA          NA           NA            NA
## 2           NA            NA          NA           NA            NA
## 3           NA            NA          NA           NA            NA
##   min_yaw_arm amplitude_roll_arm amplitude_pitch_arm amplitude_yaw_arm
## 1          NA                 NA                  NA                NA
## 2          NA                 NA                  NA                NA
## 3          NA                 NA                  NA                NA
##   roll_dumbbell pitch_dumbbell yaw_dumbbell kurtosis_roll_dumbbell
## 1      13.05217      -70.49400    -84.87394
## 2      13.13074      -70.63751    -84.71065
## 3      12.85075      -70.27812    -85.14078
##   kurtosis_picth_dumbbell kurtosis_yaw_dumbbell skewness_roll_dumbbell
## 1
## 2
## 3
##   skewness_pitch_dumbbell skewness_yaw_dumbbell max_roll_dumbbell
## 1                                                             NA
## 2                                                             NA
## 3                                                             NA
##   max_picth_dumbbell max_yaw_dumbbell min_roll_dumbbell min_pitch_dumbbell
## 1                 NA                                 NA                  NA
## 2                 NA                                 NA                  NA
## 3                 NA                                 NA                  NA
##   min_yaw_dumbbell amplitude_roll_dumbbell amplitude_pitch_dumbbell
## 1                                       NA                       NA
## 2                                       NA                       NA
## 3                                       NA                       NA
##   amplitude_yaw_dumbbell total_accel_dumbbell var_accel_dumbbell
## 1                                          37                 NA
## 2                                          37                 NA
## 3                                          37                 NA
##   avg_roll_dumbbell stddev_roll_dumbbell var_roll_dumbbell
## 1                NA                   NA                NA
## 2                NA                   NA                NA
## 3                NA                   NA                NA
##   avg_pitch_dumbbell stddev_pitch_dumbbell var_pitch_dumbbell
## 1                 NA                   NA                NA
## 2                 NA                   NA                NA
## 3                 NA                   NA                NA
##   avg_yaw_dumbbell stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x
## 1               NA                  NA               NA                0
## 2               NA                  NA               NA                0
## 3               NA                  NA               NA                0
##   gyros_dumbbell_y gyros_dumbbell_z accel_dumbbell_x accel_dumbbell_y
## 1            -0.02                0             -234               47
## 2            -0.02                0             -233               47
## 3            -0.02                0             -232               46
##   accel_dumbbell_z magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z
```

```
## 1              -271              -559               293               -65
## 2              -269              -555               296               -64
## 3              -270              -561               298               -63
##   roll_forearm pitch_forearm yaw_forearm kurtosis_roll_forearm
## 1         28.4         -63.9        -153
## 2         28.3         -63.9        -153
## 3         28.3         -63.9        -152
##   kurtosis_picth_forearm kurtosis_yaw_forearm skewness_roll_forearm
## 1
## 2
## 3
##   skewness_pitch_forearm skewness_yaw_forearm max_roll_forearm
## 1                                                           NA
## 2                                                           NA
## 3                                                           NA
##   max_picth_forearm max_yaw_forearm min_roll_forearm min_pitch_forearm
## 1                NA                                NA                NA
## 2                NA                                NA                NA
## 3                NA                                NA                NA
##   min_yaw_forearm amplitude_roll_forearm amplitude_pitch_forearm
## 1                                      NA                     NA
## 2                                      NA                     NA
## 3                                      NA                     NA
##   amplitude_yaw_forearm total_accel_forearm var_accel_forearm
## 1                                        36                NA
## 2                                        36                NA
## 3                                        36                NA
##   avg_roll_forearm stddev_roll_forearm var_roll_forearm avg_pitch_forearm
## 1               NA                   NA                NA                NA
## 2               NA                   NA                NA                NA
## 3               NA                   NA                NA                NA
##   stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## 1                   NA                NA               NA
## 2                   NA                NA               NA
## 3                   NA                NA               NA
##   stddev_yaw_forearm var_yaw_forearm gyros_forearm_x gyros_forearm_y
## 1                 NA              NA            0.03            0.00
## 2                 NA              NA            0.02            0.00
## 3                 NA              NA            0.03           -0.02
##   gyros_forearm_z accel_forearm_x accel_forearm_y accel_forearm_z
## 1           -0.02             192             203            -215
## 2           -0.02             192             203            -216
## 3            0.00             196             204            -213
##   magnet_forearm_x magnet_forearm_y magnet_forearm_z classe
## 1              -17              654              476      A
## 2              -18              661              473      A
## 3              -18              658              469      A
```

```
str(training)[1:10]
```

```
## 'data.frame':    19622 obs. of  160 variables:
##  $ X                       : int  1 2 3 4 5 6 7 8 9 10 ...
```

```
##  $ user_name               : Factor w/ 6 levels "adelmo","carlitos",..: 2 2 2 2
2 2 2 2 2 2 ...
##  $ raw_timestamp_part_1    : int  1323084231 1323084231 1323084231 1323084232 1
323084232 1323084232 1323084232 1323084232 1323084232 1323084232 ...
##  $ raw_timestamp_part_2    : int  788290 808298 820366 120339 196328 304277 368
296 440390 484323 484434 ...
##  $ cvtd_timestamp          : Factor w/ 20 levels "02/12/2011 13:32",..: 9 9 9 9
9 9 9 9 9 9 ...
##  $ new_window              : Factor w/ 2 levels "no","yes": 1 1 1 1 1 1 1 1 1 1
...
##  $ num_window              : int  11 11 11 12 12 12 12 12 12 12 ...
##  $ roll_belt               : num  1.41 1.41 1.42 1.48 1.48 1.45 1.42 1.42 1.43
1.45 ...
##  $ pitch_belt              : num  8.07 8.07 8.07 8.05 8.07 8.06 8.09 8.13 8.16
8.17 ...
##  $ yaw_belt                : num  -94.4 -94.4 -94.4 -94.4 -94.4 -94.4 -94.4 -94
.4 -94.4 -94.4 ...
##  $ total_accel_belt        : int  3 3 3 3 3 3 3 3 3 3 ...
##  $ kurtosis_roll_belt      : Factor w/ 397 levels "","#DIV/0!","-0.016850",..:
1 1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_picth_belt     : Factor w/ 317 levels "","#DIV/0!","-0.021887",..:
1 1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_yaw_belt       : Factor w/ 2 levels "","#DIV/0!": 1 1 1 1 1 1 1 1 1
1 ...
##  $ skewness_roll_belt      : Factor w/ 395 levels "","#DIV/0!","-0.003095",..:
1 1 1 1 1 1 1 1 1 1 ...
##  $ skewness_roll_belt.1    : Factor w/ 338 levels "","#DIV/0!","-0.005928",..:
1 1 1 1 1 1 1 1 1 1 ...
##  $ skewness_yaw_belt       : Factor w/ 2 levels "","#DIV/0!": 1 1 1 1 1 1 1 1 1
1 ...
##  $ max_roll_belt           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ max_picth_belt          : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ max_yaw_belt            : Factor w/ 68 levels "","#DIV/0!","-0.1",..: 1 1 1
1 1 1 1 1 1 1 ...
##  $ min_roll_belt           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_pitch_belt          : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_yaw_belt            : Factor w/ 68 levels "","#DIV/0!","-0.1",..: 1 1 1
1 1 1 1 1 1 1 ...
##  $ amplitude_roll_belt     : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ amplitude_pitch_belt    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ amplitude_yaw_belt      : Factor w/ 4 levels "","#DIV/0!","0.00",..: 1 1 1 1
1 1 1 1 1 1 ...
##  $ var_total_accel_belt    : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_roll_belt           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_roll_belt        : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_roll_belt           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_pitch_belt          : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_pitch_belt       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_pitch_belt          : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_yaw_belt            : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_yaw_belt         : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_yaw_belt            : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ gyros_belt_x            : num  0 0.02 0 0.02 0.02 0.02 0.02 0.02 0.02 0.03 .
..
```

```
##  $ gyros_belt_y           : num  0 0 0 0 0.02 0 0 0 0 0 ...
##  $ gyros_belt_z           : num  -0.02 -0.02 -0.02 -0.03 -0.02 -0.02 -0.02 -0.
02 -0.02 0 ...
##  $ accel_belt_x           : int  -21 -22 -20 -22 -21 -21 -22 -22 -20 -21 ...
##  $ accel_belt_y           : int  4 4 5 3 2 4 3 4 2 4 ...
##  $ accel_belt_z           : int  22 22 23 21 24 21 21 21 24 22 ...
##  $ magnet_belt_x          : int  -3 -7 -2 -6 -6 0 -4 -2 1 -3 ...
##  $ magnet_belt_y          : int  599 608 600 604 600 603 599 603 602 609 ...
##  $ magnet_belt_z          : int  -313 -311 -305 -310 -302 -312 -311 -313 -312
-308 ...
##  $ roll_arm               : num  -128 -128 -128 -128 -128 -128 -128 -128 -128
-128 ...
##  $ pitch_arm              : num  22.5 22.5 22.5 22.1 22.1 22 21.9 21.8 21.7 21
.6 ...
##  $ yaw_arm                : num  -161 -161 -161 -161 -161 -161 -161 -161 -161
-161 ...
##  $ total_accel_arm        : int  34 34 34 34 34 34 34 34 34 34 ...
##  $ var_accel_arm          : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_roll_arm           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_roll_arm        : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_roll_arm           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_pitch_arm          : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_pitch_arm       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_pitch_arm          : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ avg_yaw_arm            : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stddev_yaw_arm         : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ var_yaw_arm            : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ gyros_arm_x            : num  0 0.02 0.02 0.02 0 0.02 0 0.02 0.02 0.02 ...
##  $ gyros_arm_y            : num  0 -0.02 -0.02 -0.03 -0.03 -0.03 -0.03 -0.02 -
0.03 -0.03 ...
##  $ gyros_arm_z            : num  -0.02 -0.02 -0.02 0.02 0 0 0 0 -0.02 -0.02 ..
.
##  $ accel_arm_x            : int  -288 -290 -289 -289 -289 -289 -289 -289 -288
-288 ...
##  $ accel_arm_y            : int  109 110 110 111 111 111 111 111 109 110 ...
##  $ accel_arm_z            : int  -123 -125 -126 -123 -123 -122 -125 -124 -122
-124 ...
##  $ magnet_arm_x           : int  -368 -369 -368 -372 -374 -369 -373 -372 -369
-376 ...
##  $ magnet_arm_y           : int  337 337 344 344 337 342 336 338 341 334 ...
##  $ magnet_arm_z           : int  516 513 513 512 506 513 509 510 518 516 ...
##  $ kurtosis_roll_arm      : Factor w/ 330 levels "","#DIV/0!","-0.02438",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_picth_arm     : Factor w/ 328 levels "","#DIV/0!","-0.00484",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_yaw_arm       : Factor w/ 395 levels "","#DIV/0!","-0.01548",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ skewness_roll_arm      : Factor w/ 331 levels "","#DIV/0!","-0.00051",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ skewness_pitch_arm     : Factor w/ 328 levels "","#DIV/0!","-0.00184",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ skewness_yaw_arm       : Factor w/ 395 levels "","#DIV/0!","-0.00311",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ max_roll_arm           : num  NA NA NA NA NA NA NA NA NA NA ...
```

```
##  $ max_picth_arm           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ max_yaw_arm             : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_roll_arm            : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_pitch_arm           : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_yaw_arm             : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ amplitude_roll_arm      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ amplitude_pitch_arm     : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ amplitude_yaw_arm       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ roll_dumbbell           : num  13.1 13.1 12.9 13.4 13.4 ...
##  $ pitch_dumbbell          : num  -70.5 -70.6 -70.3 -70.4 -70.4 ...
##  $ yaw_dumbbell            : num  -84.9 -84.7 -85.1 -84.9 -84.9 ...
##  $ kurtosis_roll_dumbbell  : Factor w/ 398 levels "","#DIV/0!","-0.0035",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_picth_dumbbell : Factor w/ 401 levels "","#DIV/0!","-0.0163",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ kurtosis_yaw_dumbbell   : Factor w/ 2 levels "","#DIV/0!": 1 1 1 1 1 1 1 1 1
1 ...
##  $ skewness_roll_dumbbell  : Factor w/ 401 levels "","#DIV/0!","-0.0082",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ skewness_pitch_dumbbell : Factor w/ 402 levels "","#DIV/0!","-0.0053",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ skewness_yaw_dumbbell   : Factor w/ 2 levels "","#DIV/0!": 1 1 1 1 1 1 1 1 1
1 ...
##  $ max_roll_dumbbell       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ max_picth_dumbbell      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ max_yaw_dumbbell        : Factor w/ 73 levels "","#DIV/0!","-0.1",..: 1 1 1
1 1 1 1 1 1 1 ...
##  $ min_roll_dumbbell       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_pitch_dumbbell      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ min_yaw_dumbbell        : Factor w/ 73 levels "","#DIV/0!","-0.1",..: 1 1 1
1 1 1 1 1 1 1 ...
##  $ amplitude_roll_dumbbell : num  NA NA NA NA NA NA NA NA NA NA ...
##   [list output truncated]
```

```
## NULL
```

```
names(training)
```

```
##   [1] "X"                        "user_name"
##   [3] "raw_timestamp_part_1"     "raw_timestamp_part_2"
##   [5] "cvtd_timestamp"           "new_window"
##   [7] "num_window"               "roll_belt"
##   [9] "pitch_belt"               "yaw_belt"
##  [11] "total_accel_belt"         "kurtosis_roll_belt"
##  [13] "kurtosis_picth_belt"      "kurtosis_yaw_belt"
##  [15] "skewness_roll_belt"       "skewness_roll_belt.1"
##  [17] "skewness_yaw_belt"        "max_roll_belt"
##  [19] "max_picth_belt"           "max_yaw_belt"
##  [21] "min_roll_belt"            "min_pitch_belt"
##  [23] "min_yaw_belt"             "amplitude_roll_belt"
##  [25] "amplitude_pitch_belt"     "amplitude_yaw_belt"
##  [27] "var_total_accel_belt"     "avg_roll_belt"
```

```
##  [29] "stddev_roll_belt"          "var_roll_belt"
##  [31] "avg_pitch_belt"            "stddev_pitch_belt"
##  [33] "var_pitch_belt"            "avg_yaw_belt"
##  [35] "stddev_yaw_belt"           "var_yaw_belt"
##  [37] "gyros_belt_x"              "gyros_belt_y"
##  [39] "gyros_belt_z"              "accel_belt_x"
##  [41] "accel_belt_y"              "accel_belt_z"
##  [43] "magnet_belt_x"             "magnet_belt_y"
##  [45] "magnet_belt_z"             "roll_arm"
##  [47] "pitch_arm"                 "yaw_arm"
##  [49] "total_accel_arm"           "var_accel_arm"
##  [51] "avg_roll_arm"              "stddev_roll_arm"
##  [53] "var_roll_arm"              "avg_pitch_arm"
##  [55] "stddev_pitch_arm"          "var_pitch_arm"
##  [57] "avg_yaw_arm"               "stddev_yaw_arm"
##  [59] "var_yaw_arm"               "gyros_arm_x"
##  [61] "gyros_arm_y"               "gyros_arm_z"
##  [63] "accel_arm_x"               "accel_arm_y"
##  [65] "accel_arm_z"               "magnet_arm_x"
##  [67] "magnet_arm_y"              "magnet_arm_z"
##  [69] "kurtosis_roll_arm"         "kurtosis_picth_arm"
##  [71] "kurtosis_yaw_arm"          "skewness_roll_arm"
##  [73] "skewness_pitch_arm"        "skewness_yaw_arm"
##  [75] "max_roll_arm"              "max_picth_arm"
##  [77] "max_yaw_arm"               "min_roll_arm"
##  [79] "min_pitch_arm"             "min_yaw_arm"
##  [81] "amplitude_roll_arm"        "amplitude_pitch_arm"
##  [83] "amplitude_yaw_arm"         "roll_dumbbell"
##  [85] "pitch_dumbbell"            "yaw_dumbbell"
##  [87] "kurtosis_roll_dumbbell"    "kurtosis_picth_dumbbell"
##  [89] "kurtosis_yaw_dumbbell"     "skewness_roll_dumbbell"
##  [91] "skewness_pitch_dumbbell"   "skewness_yaw_dumbbell"
##  [93] "max_roll_dumbbell"         "max_picth_dumbbell"
##  [95] "max_yaw_dumbbell"          "min_roll_dumbbell"
##  [97] "min_pitch_dumbbell"        "min_yaw_dumbbell"
##  [99] "amplitude_roll_dumbbell"   "amplitude_pitch_dumbbell"
## [101] "amplitude_yaw_dumbbell"    "total_accel_dumbbell"
## [103] "var_accel_dumbbell"        "avg_roll_dumbbell"
## [105] "stddev_roll_dumbbell"      "var_roll_dumbbell"
## [107] "avg_pitch_dumbbell"        "stddev_pitch_dumbbell"
## [109] "var_pitch_dumbbell"        "avg_yaw_dumbbell"
## [111] "stddev_yaw_dumbbell"       "var_yaw_dumbbell"
## [113] "gyros_dumbbell_x"          "gyros_dumbbell_y"
## [115] "gyros_dumbbell_z"          "accel_dumbbell_x"
## [117] "accel_dumbbell_y"          "accel_dumbbell_z"
## [119] "magnet_dumbbell_x"         "magnet_dumbbell_y"
## [121] "magnet_dumbbell_z"         "roll_forearm"
## [123] "pitch_forearm"             "yaw_forearm"
## [125] "kurtosis_roll_forearm"     "kurtosis_picth_forearm"
## [127] "kurtosis_yaw_forearm"      "skewness_roll_forearm"
## [129] "skewness_pitch_forearm"    "skewness_yaw_forearm"
## [131] "max_roll_forearm"          "max_picth_forearm"
## [133] "max_yaw_forearm"           "min_roll_forearm"
## [135] "min_pitch_forearm"         "min_yaw_forearm"
```

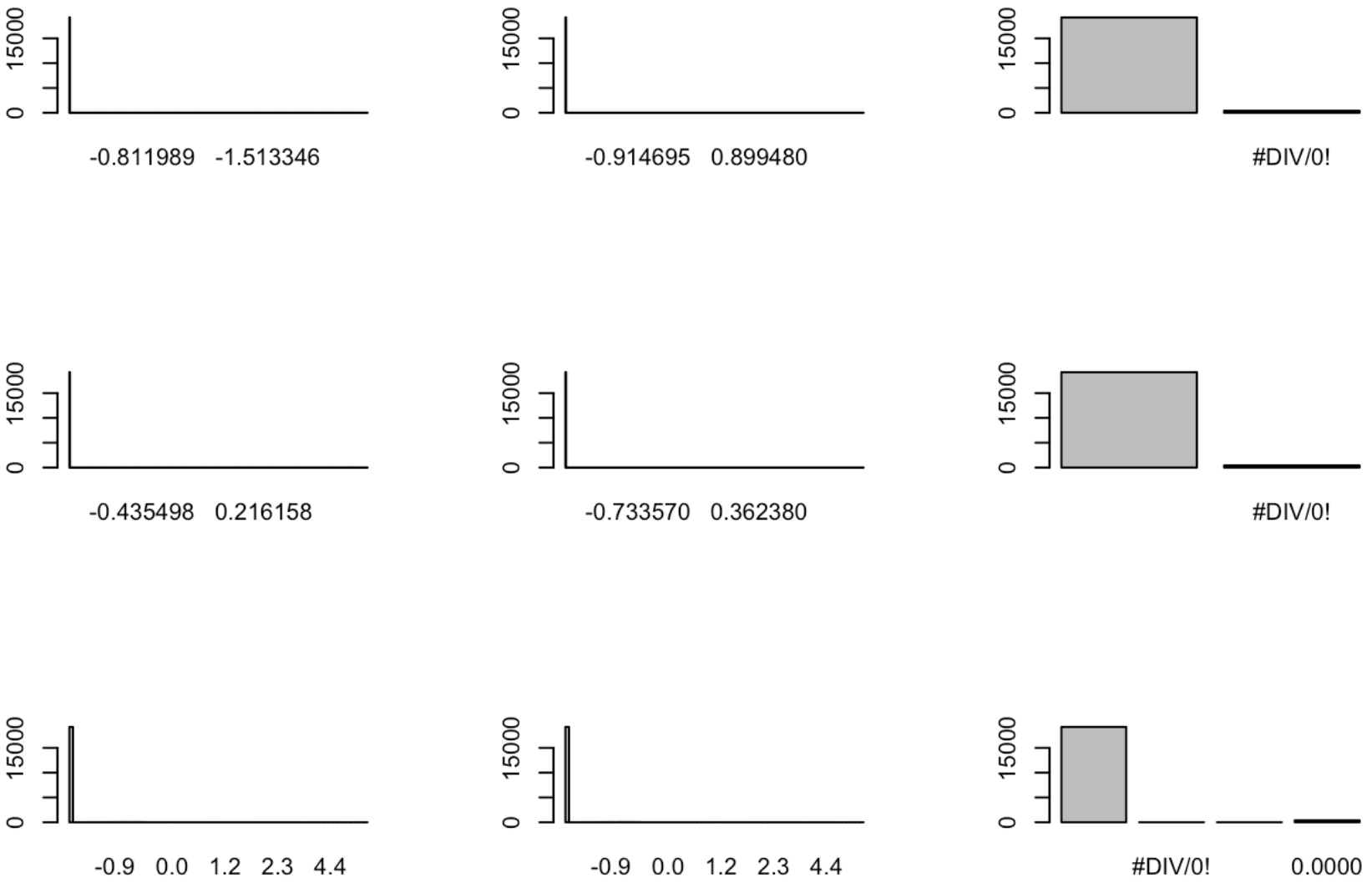
```
## [137] "amplitude_roll_forearm"   "amplitude_pitch_forearm"
## [139] "amplitude_yaw_forearm"    "total_accel_forearm"
## [141] "var_accel_forearm"        "avg_roll_forearm"
## [143] "stddev_roll_forearm"      "var_roll_forearm"
## [145] "avg_pitch_forearm"        "stddev_pitch_forearm"
## [147] "var_pitch_forearm"        "avg_yaw_forearm"
## [149] "stddev_yaw_forearm"       "var_yaw_forearm"
## [151] "gyros_forearm_x"          "gyros_forearm_y"
## [153] "gyros_forearm_z"          "accel_forearm_x"
## [155] "accel_forearm_y"          "accel_forearm_z"
## [157] "magnet_forearm_x"         "magnet_forearm_y"
## [159] "magnet_forearm_z"         "classe"
```

In summary there are 160 variables, including the “classe” as the dependent variable. Further investigations reveal that first 7 variables are not quantifiable and can be taken out.

```
training1 <- training[,-(1:7)]
```

For the remaining variables, quite a number of them are factor variable. A quick plot is made onto some of them to check their behavior.

```
par(mfrow = c(3,3))
for (i in c(5:10, 13, 16, 19)) plot(training1[,i])
```



As such we further clean up the data by removing all factor variables. Also in order to make use of the Random Forrect, we remove variables with any NA data.

```
na_sum <- lapply(1:152, function(x) sum(is.na(training1[,x])))
training1 <- training1[,na_sum==0]
is_factor <- lapply(1:ncol(training1), function(x) is.factor(training1[,x]))
training1 <- training1[,is_factor==FALSE]
```

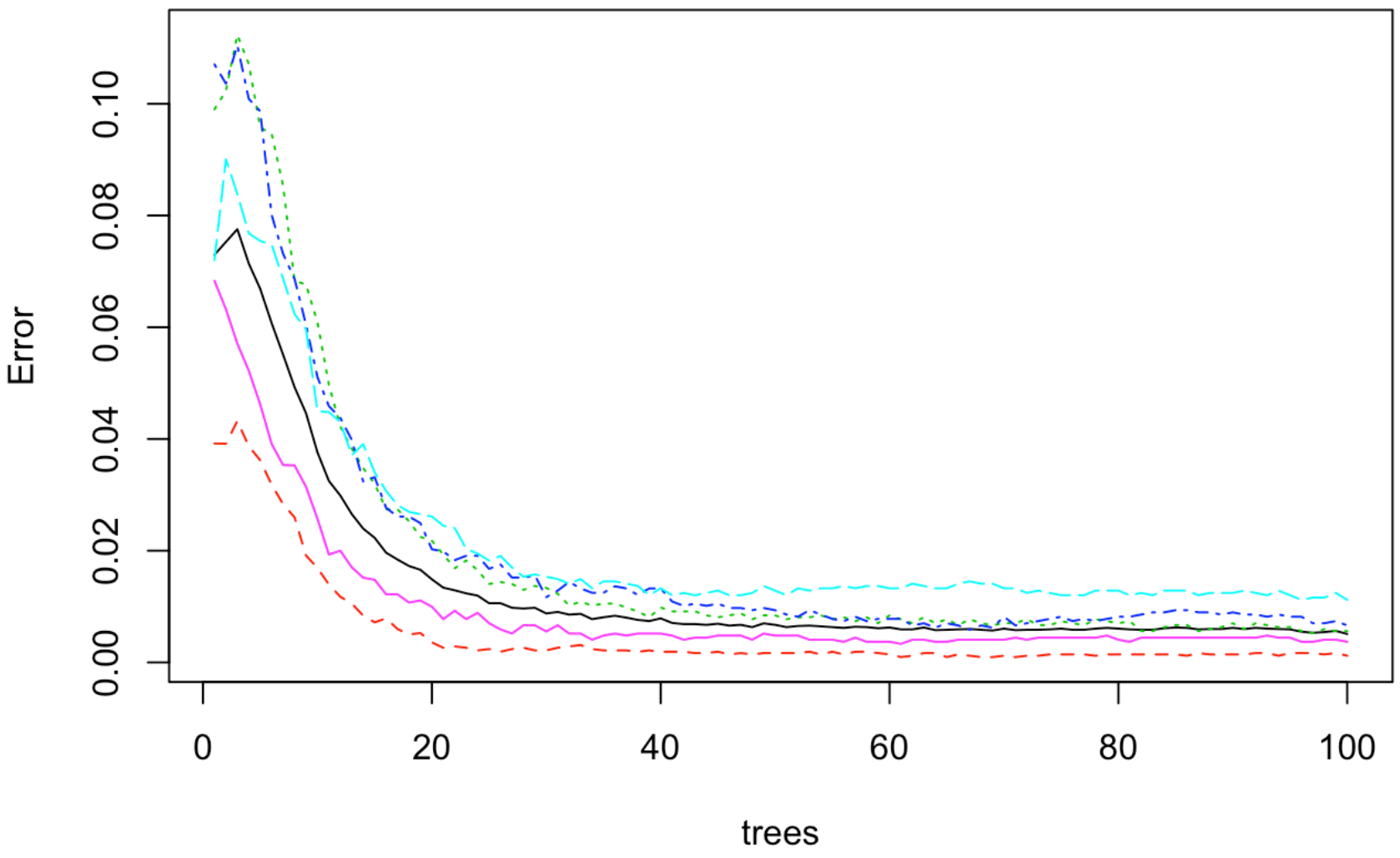
Model Fitting

We dissect the data into two portions. Three quarters will be used for model learning and the remaining 1/4 are taken aside for “probing” of the model.

We make use of Random Forest process, setting *ntree* to be 100. The plot reveals that the error is quite stable well before 100 trees.

```
set.seed(12345)
inTrain <- createDataPartition(y=training$classe,p=0.75, list=FALSE)
hartrain <- training1[inTrain,]
harprobe <- training1[-inTrain,]

modFit <- randomForest(y = training$classe[inTrain], x = hartrain,
prox=TRUE, ntree=100)
par(mfrow = c(1,1))
plot(modFit)
```



We then apply the model to the set aside 1/4 data, following by checking with the *classe* data in the original dataset.

```
pred <- predict(modFit, harprobe)
sum(pred==training$classe[-inTrain])
```

```
## [1] 4876
```

The correctness is:

```
sum(pred==training$classe[-inTrain]) / nrow(harprobe)
```

```
## [1] 0.9942904
```

Error is incurred as we only use only 52 variables of the data set. Including more variables would improve the accuracy, however, with the obtained result the upside would be limited.

Testing Result

We next moved on to fit the model to predict the 20 test cases:

```
testing1 <- testing[,-(1:7)]
na_sum <- lapply(1:152, function(x) sum(is.na(testing1[,x])))
testing1 <- testing1[,na_sum==0]
is_factor <- lapply(1:ncol(testing1), function(x) is.factor(testing1[,x]))
testing1 <- testing1[,is_factor==FALSE]
testpred <- predict(modFit, testing1)
testpred
```

```
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
##  B  A  B  A  A  E  D  B  A  A  B  C  B  A  E  E  A  B  B  B
## Levels: A B C D E
```

Conclusion

This is a typical case that we can make apply tree prediction. The *caret* package in *R* is a very effective tool to build the model. The test results have been obtained accordingly.

Reference

Ugulino, W.; Cardador, D.; Vega, K.; Velloso, E.; Milidiu, R.; Fuks, H. Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements. Proceedings of 21st Brazilian Symposium on Artificial Intelligence. Advances in Artificial Intelligence - SBIA 2012. In: Lecture Notes in Computer Science. , pp. 52-61. Curitiba, PR: Springer Berlin / Heidelberg, 2012. ISBN 978-3-642-34458-9. DOI: 10.1007/978-3-642-34459-6_6