

Prediction Assignment Writeup

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 dumbbell 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://web.archive.org/web/20161224072740/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>

The test data are available here:

<https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv>

The data for this project come from this source: <http://web.archive.org/web/20161224072740/http://groupware.les.inf.puc-rio.br/har>. 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.

Introduction and Loadind dataset

```
library(caret)

## Loading required package: lattice
## Loading required package: ggplot2
library(rpart)
library(rpart.plot)
library(RColorBrewer)
library(rattle)

## Loading required package: tibble
## Loading required package: bitops

## Rattle: A free graphical interface for data science with R.
## Version 5.4.0 Copyright (c) 2006-2020 Togaware Pty Ltd.
## Type 'rattle()' to shake, rattle, and roll your data.

library(randomForest)

## randomForest 4.6-14
## Type rfNews() to see new features/changes/bug fixes.
```

```

## 
## Attaching package: 'randomForest'
## The following object is masked from 'package:rattle':
## 
##     importance
## The following object is masked from 'package:ggplot2':
## 
##     margin
library(gbm)

## Loaded gbm 2.1.5
library(plyr)
dt_training <- read.csv("pml-training.csv", na.strings=c("NA","#DIV/0!",""))
dt_testing <- read.csv("pml-testing.csv", na.strings=c("NA","#DIV/0!",""))

```

Exploratory analysis and cleaning data

```
str(dt_testing)
```

```

## 'data.frame':    20 obs. of  160 variables:
## $ X                  : int  1 2 3 4 5 6 7 8 9 10 ...
## $ user_name          : Factor w/ 6 levels "adelmo","carlitos",...: 6 5 5 1 4 5 5 5 2 3 ...
## $ raw_timestamp_part_1 : int  1323095002 1322673067 1322673075 1322832789 1322489635 1322673149 ...
## $ raw_timestamp_part_2 : int  868349 778725 342967 560311 814776 510661 766645 54671 916313 3842 ...
## $ cvtd_timestamp      : Factor w/ 11 levels "02/12/2011 13:33",...: 5 10 10 1 6 11 11 10 3 2 ...
## $ new_window          : Factor w/ 1 level "no": 1 1 1 1 1 1 1 1 1 ...
## $ num_window          : int  74 431 439 194 235 504 485 440 323 664 ...
## $ roll_belt            : num  123 1.02 0.87 125 1.35 -5.92 1.2 0.43 0.93 114 ...
## $ pitch_belt           : num  27 4.87 1.82 -41.6 3.33 1.59 4.44 4.15 6.72 22.4 ...
## $ yaw_belt              : num  -4.75 -88.9 -88.5 162 -88.6 -87.7 -87.3 -88.5 -93.7 -13.1 ...
## $ total_accel_belt     : int  20 4 5 17 3 4 4 4 4 18 ...
## $ kurtosis_roll_belt   : logi NA NA NA NA NA NA ...
## $ kurtosis_pitch_belt  : logi NA NA NA NA NA NA ...
## $ kurtosis_yaw_belt    : logi NA NA NA NA NA NA ...
## $ skewness_roll_belt    : logi NA NA NA NA NA NA ...
## $ skewness_roll_belt.1  : logi NA NA NA NA NA NA ...
## $ skewness_yaw_belt     : logi NA NA NA NA NA NA ...
## $ max_roll_belt         : logi NA NA NA NA NA NA ...
## $ max_pitch_belt        : logi NA NA NA NA NA NA ...
## $ max_yaw_belt          : logi NA NA NA NA NA NA ...
## $ min_roll_belt         : logi NA NA NA NA NA NA ...
## $ min_pitch_belt        : logi NA NA NA NA NA NA ...
## $ min_yaw_belt          : logi NA NA NA NA NA NA ...
## $ amplitude_roll_belt   : logi NA NA NA NA NA NA ...
## $ amplitude_pitch_belt  : logi NA NA NA NA NA NA ...
## $ amplitude_yaw_belt    : logi NA NA NA NA NA NA ...
## $ var_total_accel_belt  : logi NA NA NA NA NA NA ...
## $ avg_roll_belt          : logi NA NA NA NA NA NA ...
## $ stddev_roll_belt       : logi NA NA NA NA NA NA ...
## $ var_roll_belt          : logi NA NA NA NA NA NA ...
## $ avg_pitch_belt         : logi NA NA NA NA NA NA ...
## $ stddev_pitch_belt      : logi NA NA NA NA NA NA ...

```

```

## $ var_pitch_belt      : logi  NA NA NA NA NA NA ...
## $ avg_yaw_belt        : logi  NA NA NA NA NA NA ...
## $ stddev_yaw_belt     : logi  NA NA NA NA NA NA ...
## $ var_yaw_belt        : logi  NA NA NA NA NA NA ...
## $ gyros_belt_x         : num   -0.5 -0.06 0.05 0.11 0.03 0.1 -0.06 -0.18 0.1 0.14 ...
## $ gyros_belt_y         : num   -0.02 -0.02 0.02 0.11 0.02 0.05 0 -0.02 0 0.11 ...
## $ gyros_belt_z         : num   -0.46 -0.07 0.03 -0.16 0 -0.13 0 -0.03 -0.02 -0.16 ...
## $ accel_belt_x         : int   -38 -13 1 46 -8 -11 -14 -10 -15 -25 ...
## $ accel_belt_y         : int   69 11 -1 45 4 -16 2 -2 1 63 ...
## $ accel_belt_z         : int   -179 39 49 -156 27 38 35 42 32 -158 ...
## $ magnet_belt_x        : int   -13 43 29 169 33 31 50 39 -6 10 ...
## $ magnet_belt_y        : int   581 636 631 608 566 638 622 635 600 601 ...
## $ magnet_belt_z        : int   -382 -309 -312 -304 -418 -291 -315 -305 -302 -330 ...
## $ roll_arm              : num   40.7 0 0 -109 76.1 0 0 0 -137 -82.4 ...
## $ pitch_arm             : num   -27.8 0 0 55 2.76 0 0 0 11.2 -63.8 ...
## $ yaw_arm               : num   178 0 0 -142 102 0 0 0 -167 -75.3 ...
## $ total_accel_arm       : int   10 38 44 25 29 14 15 22 34 32 ...
## $ var_accel_arm         : logi  NA NA NA NA NA NA ...
## $ avg_roll_arm          : logi  NA NA NA NA NA NA ...
## $ stddev_roll_arm        : logi  NA NA NA NA NA NA ...
## $ var_roll_arm           : logi  NA NA NA NA NA NA ...
## $ avg_pitch_arm          : logi  NA NA NA NA NA NA ...
## $ stddev_pitch_arm        : logi  NA NA NA NA NA NA ...
## $ var_pitch_arm          : logi  NA NA NA NA NA NA ...
## $ avg_yaw_arm            : logi  NA NA NA NA NA NA ...
## $ stddev_yaw_arm          : logi  NA NA NA NA NA NA ...
## $ var_yaw_arm             : logi  NA NA NA NA NA NA ...
## $ gyros_arm_x             : num   -1.65 -1.17 2.1 0.22 -1.96 0.02 2.36 -3.71 0.03 0.26 ...
## $ gyros_arm_y             : num   0.48 0.85 -1.36 -0.51 0.79 0.05 -1.01 1.85 -0.02 -0.5 ...
## $ gyros_arm_z             : num   -0.18 -0.43 1.13 0.92 -0.54 -0.07 0.89 -0.69 -0.02 0.79 ...
## $ accel_arm_x             : int   16 -290 -341 -238 -197 -26 99 -98 -287 -301 ...
## $ accel_arm_y             : int   38 215 245 -57 200 130 79 175 111 -42 ...
## $ accel_arm_z             : int   93 -90 -87 6 -30 -19 -67 -78 -122 -80 ...
## $ magnet_arm_x            : int   -326 -325 -264 -173 -170 396 702 535 -367 -420 ...
## $ magnet_arm_y            : int   385 447 474 257 275 176 15 215 335 294 ...
## $ magnet_arm_z            : int   481 434 413 633 617 516 217 385 520 493 ...
## $ kurtosis_roll_arm       : logi  NA NA NA NA NA NA ...
## $ kurtosis_pitch_arm      : logi  NA NA NA NA NA NA ...
## $ kurtosis_yaw_arm        : logi  NA NA NA NA NA NA ...
## $ skewness_roll_arm        : logi  NA NA NA NA NA NA ...
## $ skewness_pitch_arm       : logi  NA NA NA NA NA NA ...
## $ skewness_yaw_arm         : logi  NA NA NA NA NA NA ...
## $ max_roll_arm            : logi  NA NA NA NA NA NA ...
## $ max_pitch_arm           : logi  NA NA NA NA NA NA ...
## $ max_yaw_arm              : logi  NA NA NA NA NA NA ...
## $ min_roll_arm             : logi  NA NA NA NA NA NA ...
## $ min_pitch_arm            : logi  NA NA NA NA NA NA ...
## $ min_yaw_arm              : logi  NA NA NA NA NA NA ...
## $ amplitude_roll_arm       : logi  NA NA NA NA NA NA ...
## $ amplitude_pitch_arm      : logi  NA NA NA NA NA NA ...
## $ amplitude_yaw_arm        : logi  NA NA NA NA NA NA ...
## $ roll_dumbbell            : num   -17.7 54.5 57.1 43.1 -101.4 ...
## $ pitch_dumbbell           : num   25 -53.7 -51.4 -30 -53.4 ...
## $ yaw_dumbbell             : num   126.2 -75.5 -75.2 -103.3 -14.2 ...

```

```

## $ kurtosis_roll_dumbbell : logi NA NA NA NA NA NA ...
## $ kurtosis_pictch_dumbbell : logi NA NA NA NA NA NA ...
## $ kurtosis_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ skewness_roll_dumbbell : logi NA NA NA NA NA NA ...
## $ skewness_pitch_dumbbell : logi NA NA NA NA NA NA ...
## $ skewness_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ max_roll_dumbbell : logi NA NA NA NA NA NA ...
## $ max_pictch_dumbbell : logi NA NA NA NA NA NA ...
## $ max_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ min_roll_dumbbell : logi NA NA NA NA NA NA ...
## $ min_pitch_dumbbell : logi NA NA NA NA NA NA ...
## $ min_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ amplitude_roll_dumbbell : logi NA NA NA NA NA NA ...
## [list output truncated]

head(dt_testing)

##   X user_name raw_timestamp_part_1 raw_timestamp_part_2   cvtd_timestamp
## 1 1     pedro      1323095002           868349 05/12/2011 14:23
## 2 2    jeremy      1322673067          778725 30/11/2011 17:11
## 3 3    jeremy      1322673075          342967 30/11/2011 17:11
## 4 4    adelmo      1322832789          560311 02/12/2011 13:33
## 5 5    eurico      1322489635          814776 28/11/2011 14:13
## 6 6    jeremy      1322673149          510661 30/11/2011 17:12
##   new_window num_window roll_belt pitch_belt yaw_belt total_accel_belt
## 1       no        74    123.00     27.00    -4.75       20
## 2       no       431     1.02      4.87   -88.90        4
## 3       no       439     0.87      1.82   -88.50        5
## 4       no       194    125.00    -41.60    162.00       17
## 5       no       235     1.35      3.33   -88.60        3
## 6       no       504    -5.92      1.59   -87.70       4
##   kurtosis_roll_belt kurtosis_pictch_belt kurtosis_yaw_belt skewness_roll_belt
## 1             NA                 NA                 NA                 NA
## 2             NA                 NA                 NA                 NA
## 3             NA                 NA                 NA                 NA
## 4             NA                 NA                 NA                 NA
## 5             NA                 NA                 NA                 NA
## 6             NA                 NA                 NA                 NA
##   skewness_roll_belt.1 skewness_yaw_belt max_roll_belt max_pictch_belt
## 1             NA                 NA                 NA                 NA
## 2             NA                 NA                 NA                 NA
## 3             NA                 NA                 NA                 NA
## 4             NA                 NA                 NA                 NA
## 5             NA                 NA                 NA                 NA
## 6             NA                 NA                 NA                 NA
##   max_yaw_belt min_roll_belt min_pitch_belt min_yaw_belt amplitude_roll_belt
## 1             NA                 NA                 NA                 NA                 NA
## 2             NA                 NA                 NA                 NA                 NA
## 3             NA                 NA                 NA                 NA                 NA
## 4             NA                 NA                 NA                 NA                 NA
## 5             NA                 NA                 NA                 NA                 NA
## 6             NA                 NA                 NA                 NA                 NA
##   amplitude_pitch_belt amplitude_yaw_belt var_total_accel_belt avg_roll_belt
## 1             NA                 NA                 NA                 NA                 NA
## 2             NA                 NA                 NA                 NA                 NA

```

```

## 3 NA NA NA NA NA
## 4 NA NA NA NA NA
## 5 NA NA NA NA NA
## 6 NA NA NA NA NA
## stddev_roll_belt var_roll_belt avg_pitch_belt stddev_pitch_belt
## 1 NA NA NA NA NA
## 2 NA NA NA NA NA
## 3 NA NA NA NA NA
## 4 NA NA NA NA NA
## 5 NA NA NA NA NA
## 6 NA NA NA NA NA
## var_pitch_belt avg_yaw_belt stddev_yaw_belt var_yaw_belt gyros_belt_x
## 1 NA NA NA NA NA -0.50
## 2 NA NA NA NA NA -0.06
## 3 NA NA NA NA NA 0.05
## 4 NA NA NA NA NA 0.11
## 5 NA NA NA NA NA 0.03
## 6 NA NA NA NA NA 0.10
## gyros_belt_y gyros_belt_z accel_belt_x accel_belt_y accel_belt_z
## 1 -0.02 -0.46 -38 69 -179
## 2 -0.02 -0.07 -13 11 39
## 3 0.02 0.03 1 -1 49
## 4 0.11 -0.16 46 45 -156
## 5 0.02 0.00 -8 4 27
## 6 0.05 -0.13 -11 -16 38
## magnet_belt_x magnet_belt_y magnet_belt_z roll_arm pitch_arm yaw_arm
## 1 -13 581 -382 40.7 -27.80 178
## 2 43 636 -309 0.0 0.00 0
## 3 29 631 -312 0.0 0.00 0
## 4 169 608 -304 -109.0 55.00 -142
## 5 33 566 -418 76.1 2.76 102
## 6 31 638 -291 0.0 0.00 0
## total_accel_arm var_accel_arm avg_roll_arm stddev_roll_arm var_roll_arm
## 1 10 NA NA NA NA NA
## 2 38 NA NA NA NA NA
## 3 44 NA NA NA NA NA
## 4 25 NA NA NA NA NA
## 5 29 NA NA NA NA NA
## 6 14 NA NA NA NA NA
## avg_pitch_arm stddev_pitch_arm var_pitch_arm avg_yaw_arm stddev_yaw_arm
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA
## 4 NA NA NA NA NA NA
## 5 NA NA NA NA NA NA
## 6 NA NA NA NA NA NA
## var_yaw_arm gyros_arm_x gyros_arm_y gyros_arm_z accel_arm_x accel_arm_y
## 1 NA -1.65 0.48 -0.18 16 38
## 2 NA -1.17 0.85 -0.43 -290 215
## 3 NA 2.10 -1.36 1.13 -341 245
## 4 NA 0.22 -0.51 0.92 -238 -57
## 5 NA -1.96 0.79 -0.54 -197 200
## 6 NA 0.02 0.05 -0.07 -26 130
## accel_arm_z magnet_arm_x magnet_arm_y magnet_arm_z kurtosis_roll_arm

```

```

## 1      93     -326      385      481      NA
## 2     -90     -325      447      434      NA
## 3     -87     -264      474      413      NA
## 4      6     -173      257      633      NA
## 5     -30     -170      275      617      NA
## 6     -19      396      176      516      NA
##   kurtosis_pictch_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA
## 6             NA          NA          NA          NA      NA
##   skewness_yaw_arm max_roll_arm max_pictch_arm max_yaw_arm min_roll_arm
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA
## 6             NA          NA          NA          NA      NA
##   min_pitch_arm min_yaw_arm amplitude_roll_arm amplitude_pitch_arm
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA
## 6             NA          NA          NA          NA      NA
##   amplitude_yaw_arm roll_dumbbell pitch_dumbbell yaw_dumbbell
## 1             NA     -17.73748    24.96085   126.23596
## 2             NA      54.47761   -53.69758   -75.51480
## 3             NA      57.07031   -51.37303   -75.20287
## 4             NA      43.10927   -30.04885  -103.32003
## 5             NA     -101.38396  -53.43952  -14.19542
## 6             NA      62.18750   -50.55595  -71.12063
##   kurtosis_roll_dumbbell kurtosis_pictch_dumbbell kurtosis_yaw_dumbbell
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA
## 6             NA          NA          NA          NA      NA
##   skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA
## 6             NA          NA          NA          NA      NA
##   max_roll_dumbbell max_pictch_dumbbell max_yaw_dumbbell min_roll_dumbbell
## 1             NA          NA          NA          NA      NA
## 2             NA          NA          NA          NA      NA
## 3             NA          NA          NA          NA      NA
## 4             NA          NA          NA          NA      NA
## 5             NA          NA          NA          NA      NA

```

```

## 6          NA          NA          NA          NA
## min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
## amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
## 1          NA          NA          9
## 2          NA          NA          31
## 3          NA          NA          29
## 4          NA          NA          18
## 5          NA          NA          4
## 6          NA          NA          29
## var_accel_dumbbell avg_roll_dumbbell stddev_roll_dumbbell var_roll_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
## avg_pitch_dumbbell stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
## stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x gyros_dumbbell_y
## 1          NA          NA          0.64         0.06
## 2          NA          NA          0.34         0.05
## 3          NA          NA          0.39         0.14
## 4          NA          NA          0.10        -0.02
## 5          NA          NA          0.29        -0.47
## 6          NA          NA         -0.59         0.80
## gyros_dumbbell_z accel_dumbbell_x accel_dumbbell_y accel_dumbbell_z
## 1          -0.61         21         -15          81
## 2          -0.71        -153         155        -205
## 3          -0.34        -141         155        -196
## 4          0.05         -51          72        -148
## 5          -0.46        -18          -30          -5
## 6          1.10        -138         166        -186
## magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z roll_forearm
## 1          523         -528         -56          141
## 2          -502         388         -36          109
## 3          -506         349          41          131
## 4          -576         238          53           0
## 5          -424         252         312        -176
## 6          -543         262          96          150
## pitch_forearm yaw_forearm kurtosis_roll_forearm kurtosis_pictch_forearm
## 1          49.30       156.0          NA          NA
## 2          -17.60      106.0          NA          NA
## 3          -32.60      93.0          NA          NA

```

```

## 4      0.00      0.0      NA      NA
## 5     -2.16    -47.9      NA      NA
## 6      1.46     89.7      NA      NA
##   kurtosis_yaw_forearm skewness_roll_forearm skewness_pitch_forearm
## 1             NA          NA          NA
## 2             NA          NA          NA
## 3             NA          NA          NA
## 4             NA          NA          NA
## 5             NA          NA          NA
## 6             NA          NA          NA
##   skewness_yaw_forearm max_roll_forearm max_pictch_forearm max_yaw_forearm
## 1             NA          NA          NA          NA
## 2             NA          NA          NA          NA
## 3             NA          NA          NA          NA
## 4             NA          NA          NA          NA
## 5             NA          NA          NA          NA
## 6             NA          NA          NA          NA
##   min_roll_forearm min_pitch_forearm min_yaw_forearm amplitude_roll_forearm
## 1             NA          NA          NA          NA
## 2             NA          NA          NA          NA
## 3             NA          NA          NA          NA
## 4             NA          NA          NA          NA
## 5             NA          NA          NA          NA
## 6             NA          NA          NA          NA
##   amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
## 1             NA          NA          33
## 2             NA          NA          39
## 3             NA          NA          34
## 4             NA          NA          43
## 5             NA          NA          24
## 6             NA          NA          43
##   var_accel_forearm avg_roll_forearm stddev_roll_forearm var_roll_forearm
## 1             NA          NA          NA          NA
## 2             NA          NA          NA          NA
## 3             NA          NA          NA          NA
## 4             NA          NA          NA          NA
## 5             NA          NA          NA          NA
## 6             NA          NA          NA          NA
##   avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## 1             NA          NA          NA          NA
## 2             NA          NA          NA          NA
## 3             NA          NA          NA          NA
## 4             NA          NA          NA          NA
## 5             NA          NA          NA          NA
## 6             NA          NA          NA          NA
##   stddev_yaw_forearm var_yaw_forearm gyros_forearm_x gyros_forearm_y
## 1             NA          NA          0.74        -3.34
## 2             NA          NA          1.12        -2.78
## 3             NA          NA          0.18        -0.79
## 4             NA          NA          1.38         0.69
## 5             NA          NA          -0.75         3.10
## 6             NA          NA          -0.88         4.26
##   gyros_forearm_z accel_forearm_x accel_forearm_y accel_forearm_z
## 1            -0.59         -110         267        -149

```

```

## 2      -0.18      212      297     -118
## 3       0.28      154      271     -129
## 4       1.80      -92      406     -39
## 5       0.80      131     -93      172
## 6       1.35      230      322     -144
##   magnet_forearm_x magnet_forearm_y magnet_forearm_z problem_id
## 1      -714       419      617       1
## 2      -237       791      873       2
## 3       -51       698      783       3
## 4      -233       783      521       4
## 5       375      -787       91       5
## 6      -300       800      884       6

str(dt_training)

## '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 1323084232 1323084232 ...
## $ raw_timestamp_part_2  : int  788290 808298 820366 120339 196328 304277 368296 440390 484323 484 ...
## $ 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  : num  NA NA NA NA NA NA NA NA NA ...
## $ kurtosis_pitch_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ kurtosis_yaw_belt  : logi  NA NA NA NA NA NA ...
## $ skewness_roll_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ skewness_roll_belt.1  : num  NA NA NA NA NA NA NA NA NA ...
## $ skewness_yaw_belt  : logi  NA NA NA NA NA ...
## $ max_roll_belt    : num  NA NA NA NA NA NA NA NA NA ...
## $ max_pitch_belt   : int  NA NA NA NA NA NA NA NA NA ...
## $ max_yaw_belt     : num  NA NA NA NA NA NA NA NA NA ...
## $ min_roll_belt    : num  NA NA NA NA NA NA NA NA NA ...
## $ min_pitch_belt   : int  NA NA NA NA NA NA NA NA NA ...
## $ min_yaw_belt     : num  NA NA NA NA NA NA NA NA NA ...
## $ amplitude_roll_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ amplitude_pitch_belt  : int  NA NA NA NA NA NA NA NA NA ...
## $ amplitude_yaw_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ var_total_accel_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ avg_roll_belt    : num  NA NA NA NA NA NA NA NA NA ...
## $ stddev_roll_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ var_roll_belt     : num  NA NA NA NA NA NA NA NA NA ...
## $ avg_pitch_belt   : num  NA NA NA NA NA NA NA NA NA ...
## $ stddev_pitch_belt  : num  NA NA NA NA NA NA NA NA NA ...
## $ var_pitch_belt    : num  NA NA NA NA NA NA NA NA NA ...
## $ avg_yaw_belt     : num  NA NA NA NA NA NA NA NA NA ...
## $ stddev_yaw_belt   : num  NA NA NA NA NA NA NA NA NA ...
## $ var_yaw_belt     : num  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 0.03 ...
## $ 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 ...
## $ avg_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ stddev_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ var_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ avg_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ stddev_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ var_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ avg_yaw_arm : num NA NA NA NA NA NA NA NA NA ...
## $ stddev_yaw_arm : num NA NA NA NA NA NA NA NA NA ...
## $ var_yaw_arm : num NA NA NA NA NA NA NA NA NA ...
## $ gyros_arm_x : num 0 0.02 0.02 0.02 0 0.02 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 ...
## $ gyros_arm_z : num -0.02 -0.02 -0.02 0.02 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 -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 : num NA NA NA NA NA NA NA NA NA ...
## $ kurtosis_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ kurtosis_yaw_arm : num NA NA NA NA NA NA NA NA NA ...
## $ skewness_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ skewness_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ skewness_yaw_arm : num NA NA NA NA NA NA NA NA NA ...
## $ max_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ max_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ max_yaw_arm : int NA NA NA NA NA NA NA NA NA ...
## $ min_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ min_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ min_yaw_arm : int NA NA NA NA NA NA NA NA NA ...
## $ amplitude_roll_arm : num NA NA NA NA NA NA NA NA NA ...
## $ amplitude_pitch_arm : num NA NA NA NA NA NA NA NA NA ...
## $ amplitude_yaw_arm : int 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 : num NA NA NA NA NA NA NA NA ...
## $ kurtosis_pitch_dumbbell : num NA NA NA NA NA NA NA NA ...
## $ kurtosis_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ skewness_roll_dumbbell : num NA NA NA NA NA NA NA NA ...
## $ skewness_pitch_dumbbell : num NA NA NA NA NA NA NA NA ...
## $ skewness_yaw_dumbbell : logi NA NA NA NA NA NA ...
## $ max_roll_dumbbell : num NA NA NA NA NA NA NA NA ...

```

```

## $ max_pictch_dumbbell      : num  NA ...
## $ max_yaw_dumbbell         : num  NA ...
## $ min_roll_dumbbell        : num  NA ...
## $ min_pitch_dumbbell       : num  NA ...
## $ min_yaw_dumbbell         : num  NA ...
## $ amplitude_roll_dumbbell  : num  NA ...
##   [list output truncated]

head(dt_training)

##   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
## 4 4 carlitos     1323084232           120339 05/12/2011 11:23
## 5 5 carlitos     1323084232           196328 05/12/2011 11:23
## 6 6 carlitos     1323084232           304277 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
## 4       no          12     1.48     8.05    -94.4         3
## 5       no          12     1.48     8.07    -94.4         3
## 6       no          12     1.45     8.06    -94.4         3
##   kurtosis_roll_belt kurtosis_pictch_belt kurtosis_yaw_belt skewness_roll_belt
## 1             NA                 NA                 NA                NA
## 2             NA                 NA                 NA                NA
## 3             NA                 NA                 NA                NA
## 4             NA                 NA                 NA                NA
## 5             NA                 NA                 NA                NA
## 6             NA                 NA                 NA                NA
##   skewness_roll_belt.1 skewness_yaw_belt max_roll_belt max_pictch_belt
## 1             NA                 NA                 NA                NA
## 2             NA                 NA                 NA                NA
## 3             NA                 NA                 NA                NA
## 4             NA                 NA                 NA                NA
## 5             NA                 NA                 NA                NA
## 6             NA                 NA                 NA                NA
##   max_yaw_belt min_roll_belt min_pitch_belt min_yaw_belt amplitude_roll_belt
## 1             NA                 NA                 NA                NA
## 2             NA                 NA                 NA                NA
## 3             NA                 NA                 NA                NA
## 4             NA                 NA                 NA                NA
## 5             NA                 NA                 NA                NA
## 6             NA                 NA                 NA                NA
##   amplitude_pitch_belt amplitude_yaw_belt var_total_accel_belt avg_roll_belt
## 1             NA                 NA                 NA                NA
## 2             NA                 NA                 NA                NA
## 3             NA                 NA                 NA                NA
## 4             NA                 NA                 NA                NA
## 5             NA                 NA                 NA                NA
## 6             NA                 NA                 NA                NA
##   stddev_roll_belt var_roll_belt avg_pitch_belt stddev_pitch_belt
## 1             NA                 NA                 NA                NA
## 2             NA                 NA                 NA                NA

```

```

## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   var_pitch_belt avg_yaw_belt stddev_yaw_belt var_yaw_belt gyros_belt_x
## 1          NA          NA          NA          NA      0.00
## 2          NA          NA          NA          NA      0.02
## 3          NA          NA          NA          NA      0.00
## 4          NA          NA          NA          NA      0.02
## 5          NA          NA          NA          NA      0.02
## 6          NA          NA          NA          NA      0.02
##   gyros_belt_y gyros_belt_z accel_belt_x accel_belt_y accel_belt_z
## 1      0.00     -0.02      -21         4       22
## 2      0.00     -0.02      -22         4       22
## 3      0.00     -0.02      -20         5       23
## 4      0.00     -0.03      -22         3       21
## 5      0.02     -0.02      -21         2       24
## 6      0.00     -0.02      -21         4       21
##   magnet_belt_x magnet_belt_y magnet_belt_z roll_arm pitch_arm yaw_arm
## 1      -3       599     -313     -128    22.5   -161
## 2      -7       608     -311     -128    22.5   -161
## 3      -2       600     -305     -128    22.5   -161
## 4      -6       604     -310     -128    22.1   -161
## 5      -6       600     -302     -128    22.1   -161
## 6       0       603     -312     -128    22.0   -161
##   total_accel_arm var_accel_arm avg_roll_arm stddev_roll_arm var_roll_arm
## 1       34          NA          NA          NA          NA
## 2       34          NA          NA          NA          NA
## 3       34          NA          NA          NA          NA
## 4       34          NA          NA          NA          NA
## 5       34          NA          NA          NA          NA
## 6       34          NA          NA          NA          NA
##   avg_pitch_arm stddev_pitch_arm var_pitch_arm avg_yaw_arm stddev_yaw_arm
## 1          NA          NA          NA          NA          NA
## 2          NA          NA          NA          NA          NA
## 3          NA          NA          NA          NA          NA
## 4          NA          NA          NA          NA          NA
## 5          NA          NA          NA          NA          NA
## 6          NA          NA          NA          NA          NA
##   var_yaw_arm gyros_arm_x gyros_arm_y gyros_arm_z accel_arm_x accel_arm_y
## 1          NA      0.00      0.00     -0.02     -288     109
## 2          NA      0.02     -0.02     -0.02     -290     110
## 3          NA      0.02     -0.02     -0.02     -289     110
## 4          NA      0.02     -0.03      0.02     -289     111
## 5          NA      0.00     -0.03      0.00     -289     111
## 6          NA      0.02     -0.03      0.00     -289     111
##   accel_arm_z magnet_arm_x magnet_arm_y magnet_arm_z kurtosis_roll_arm
## 1      -123     -368      337      516          NA
## 2      -125     -369      337      513          NA
## 3      -126     -368      344      513          NA
## 4      -123     -372      344      512          NA
## 5      -123     -374      337      506          NA
## 6      -122     -369      342      513          NA
##   kurtosis_pitch_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm

```

```

## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   skewness_yaw_arm max_roll_arm max_pictch_arm max_yaw_arm min_roll_arm
## 1          NA          NA          NA          NA          NA
## 2          NA          NA          NA          NA          NA
## 3          NA          NA          NA          NA          NA
## 4          NA          NA          NA          NA          NA
## 5          NA          NA          NA          NA          NA
## 6          NA          NA          NA          NA          NA
##   min_pitch_arm min_yaw_arm amplitude_roll_arm amplitude_pitch_arm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   amplitude_yaw_arm roll_dumbbell pitch_dumbbell yaw_dumbbell
## 1          NA    13.05217   -70.49400   -84.87394
## 2          NA    13.13074   -70.63751   -84.71065
## 3          NA    12.85075   -70.27812   -85.14078
## 4          NA    13.43120   -70.39379   -84.87363
## 5          NA    13.37872   -70.42856   -84.85306
## 6          NA    13.38246   -70.81759   -84.46500
##   kurtosis_roll_dumbbell kurtosis_pictch_dumbbell kurtosis_yaw_dumbbell
## 1          NA          NA          NA
## 2          NA          NA          NA
## 3          NA          NA          NA
## 4          NA          NA          NA
## 5          NA          NA          NA
## 6          NA          NA          NA
##   skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
## 1          NA          NA          NA
## 2          NA          NA          NA
## 3          NA          NA          NA
## 4          NA          NA          NA
## 5          NA          NA          NA
## 6          NA          NA          NA
##   max_roll_dumbbell max_pictch_dumbbell max_yaw_dumbbell min_roll_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
## 1          NA          NA          NA
## 2          NA          NA          NA
## 3          NA          NA          NA
## 4          NA          NA          NA
## 5          NA          NA          NA

```

```

## 6          NA          NA          NA
##   amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
## 1          NA          NA          37
## 2          NA          NA          37
## 3          NA          NA          37
## 4          NA          NA          37
## 5          NA          NA          37
## 6          NA          NA          37
##   var_accel_dumbbell avg_roll_dumbbell stddev_roll_dumbbell var_roll_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   avg_pitch_dumbbell stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x gyros_dumbbell_y
## 1          NA          NA          0         -0.02
## 2          NA          NA          0         -0.02
## 3          NA          NA          0         -0.02
## 4          NA          NA          0         -0.02
## 5          NA          NA          0         -0.02
## 6          NA          NA          0         -0.02
##   gyros_dumbbell_z accel_dumbbell_x accel_dumbbell_y accel_dumbbell_z
## 1          0.00        -234         47        -271
## 2          0.00        -233         47        -269
## 3          0.00        -232         46        -270
## 4         -0.02        -232         48        -269
## 5          0.00        -233         48        -270
## 6          0.00        -234         48        -269
##   magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z roll_forearm
## 1          -559         293        -65        28.4
## 2          -555         296        -64        28.3
## 3          -561         298        -63        28.3
## 4          -552         303        -60        28.1
## 5          -554         292        -68        28.0
## 6          -558         294        -66        27.9
##   pitch_forearm yaw_forearm kurtosis_roll_forearm kurtosis_pictch_forearm
## 1          -63.9       -153          NA          NA
## 2          -63.9       -153          NA          NA
## 3          -63.9       -152          NA          NA
## 4          -63.9       -152          NA          NA
## 5          -63.9       -152          NA          NA
## 6          -63.9       -152          NA          NA
##   kurtosis_yaw_forearm skewness_roll_forearm skewness_pitch_forearm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA

```

```

## 4          NA          NA          NA
## 5          NA          NA          NA
## 6          NA          NA          NA
##   skewness_yaw_forearm max_roll_forearm max_pictch_forearm max_yaw_forearm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   min_roll_forearm min_pitch_forearm min_yaw_forearm amplitude_roll_forearm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
## 1          NA          NA          36
## 2          NA          NA          36
## 3          NA          NA          36
## 4          NA          NA          36
## 5          NA          NA          36
## 6          NA          NA          36
##   var_accel_forearm avg_roll_forearm stddev_roll_forearm var_roll_forearm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
##   avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## 1          NA          NA          NA          NA
## 2          NA          NA          NA          NA
## 3          NA          NA          NA          NA
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          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
## 4          NA          NA          0.02        -0.02
## 5          NA          NA          0.02         0.00
## 6          NA          NA          0.02        -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
## 4           0.00        189         206        -214
## 5          -0.02        189         206        -214
## 6          -0.03        193         203        -215
##   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
## 4      -16      658      469      A
## 5      -17      655      473      A
## 6       -9      660      478      A

```

In this section, we will remove all columns that contains NA and remove features that are not in the testing dataset. The features containing NA are the variance, mean and standard devition (SD) within each window for each feature. Since the testing dataset has no time-dependence, these values are useless and can be disregarded. We will also remove the first 7 features since they are related to the time-series or are not numeric.

```

features <- names(dt_testing[, colSums(is.na(dt_testing)) == 0])[8:59]

# Only use features used in testing cases.
dt_training <- dt_training[, c(features, "classe")]
dt_testing <- dt_testing[, c(features, "problem_id")]

dim(dt_training); dim(dt_testing);

## [1] 19622    53
## [1] 20 53

```

Partitioning the Dataset

As recomended in course: we will split our data into a training data set (60% of the total cases) and a testing data set (40% of the total cases; the latter should not be confused with the data in the pml-testing.csv file). This will allow us to estimate the out of sample error of our predictor.

```

set.seed(1234)

inTrain <- createDataPartition(dt_training$classe, p=0.6, list=FALSE)
training <- dt_training[inTrain,]
testing <- dt_training[-inTrain,]

dim(training); dim(testing);

## [1] 11776    53
## [1] 7846     53

```

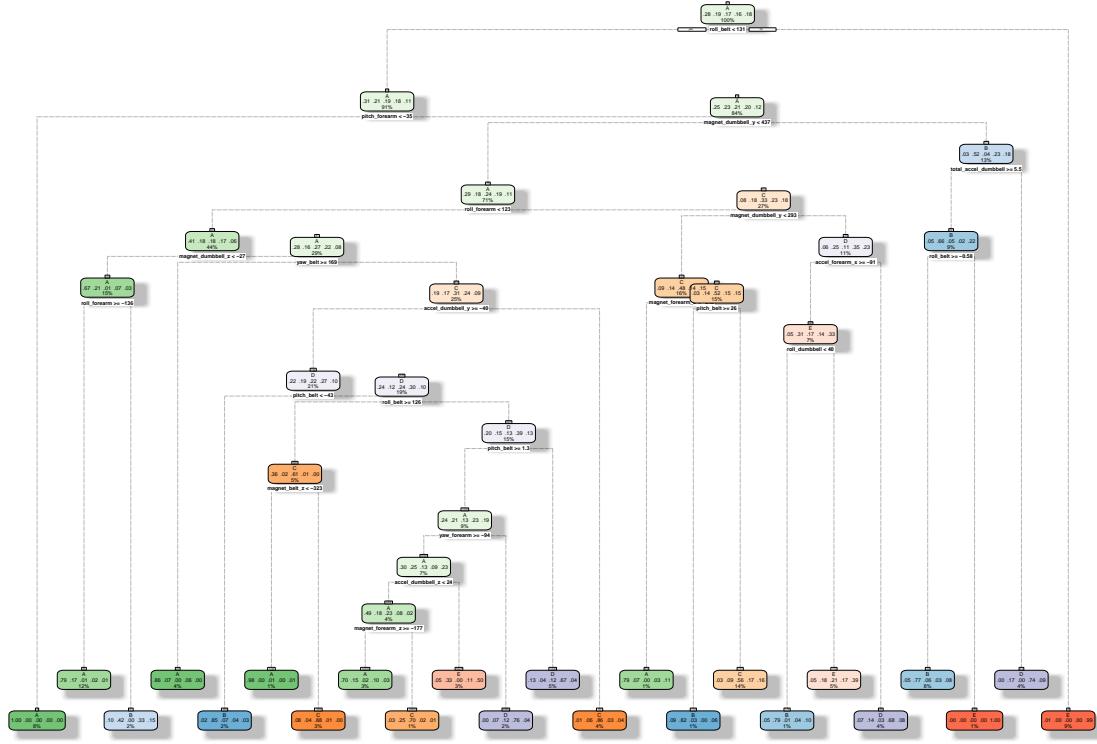
Building the Decision Tree Model

```

set.seed(1234)
modFitDT <- rpart(classe ~ ., data = training, method="class", control = rpart.control(method = "cv", n
fancyRpartPlot(modFitDT , sub = ("Decision Tree Model 14-Jul-2020"))

## Warning: labs do not fit even at cex 0.15, there may be some overplotting

```



Decision Tree Model 14–Jul–2020

Predicting with the Decision Tree Model

```
set.seed(1234)

prediction <- predict(modFitDT, testing, type = "class")
confusionMatrix(prediction, testing$classe)
```

```

## Confusion Matrix and Statistics
##
##             Reference
## Prediction    A     B     C     D     E
##           A 1980   212    21    72   31
##           B   85   862    72    90   98
##           C   56   153  1086   209  175
##           D   71   101   110   823   89
##           E   40   190    79    92 1049
##
## Overall Statistics
##
##                               Accuracy : 0.7392
##                               95% CI : (0.7294, 0.75)
## No Information Rate : 0.2845
## P-Value [Acc > NIR] : < 2.2e-16
##
##                               Kappa : 0.6699
##
```

```

##  Mcnemar's Test P-Value : < 2.2e-16
##
## Statistics by Class:
##
##          Class: A Class: B Class: C Class: D Class: E
## Sensitivity      0.8871   0.5679   0.7939   0.6400   0.7275
## Specificity     0.9401   0.9455   0.9085   0.9434   0.9374
## Pos Pred Value  0.8549   0.7142   0.6468   0.6893   0.7234
## Neg Pred Value  0.9544   0.9012   0.9543   0.9304   0.9386
## Prevalence       0.2845   0.1935   0.1744   0.1639   0.1838
## Detection Rate  0.2524   0.1099   0.1384   0.1049   0.1337
## Detection Prevalence 0.2952   0.1538   0.2140   0.1522   0.1848
## Balanced Accuracy 0.9136   0.7567   0.8512   0.7917   0.8324

```

Building the Random Forest Model

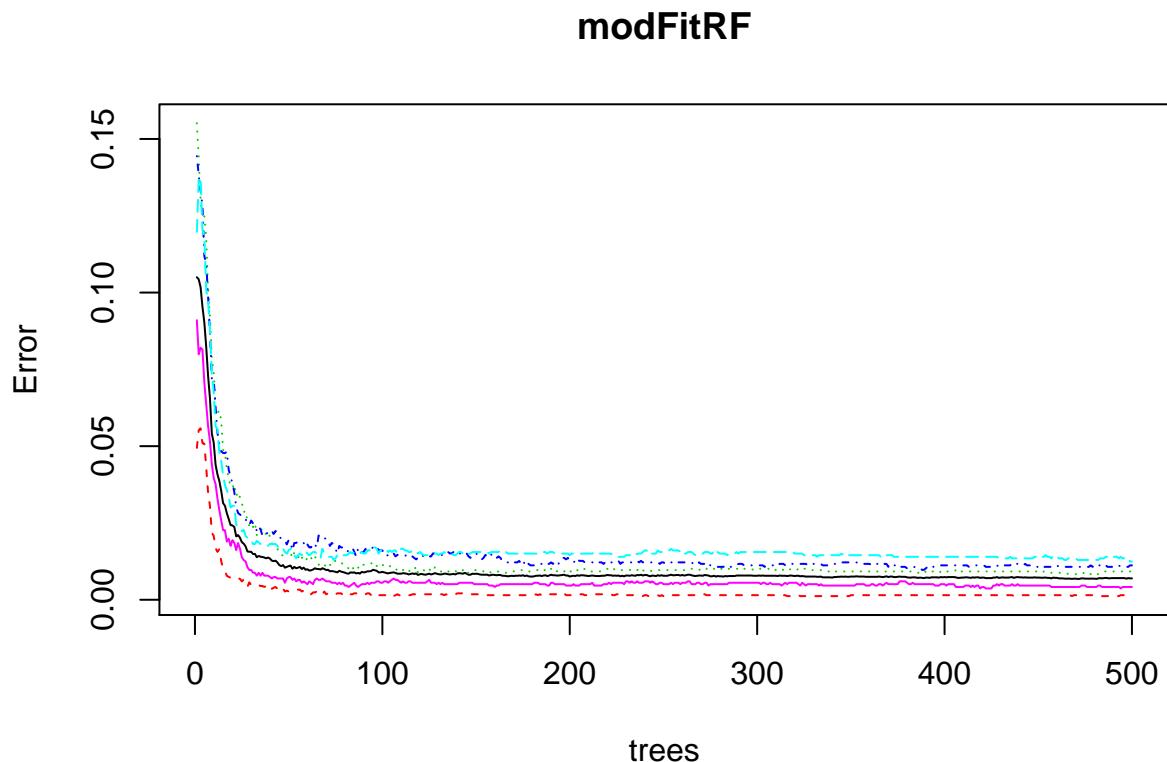
```

set.seed(1234)

modFitRF <- randomForest(classe ~ ., data = training, method = "rf", importance = T, trControl = trainControl)

plot(modFitRF)

```



Building the Boosting Model

```

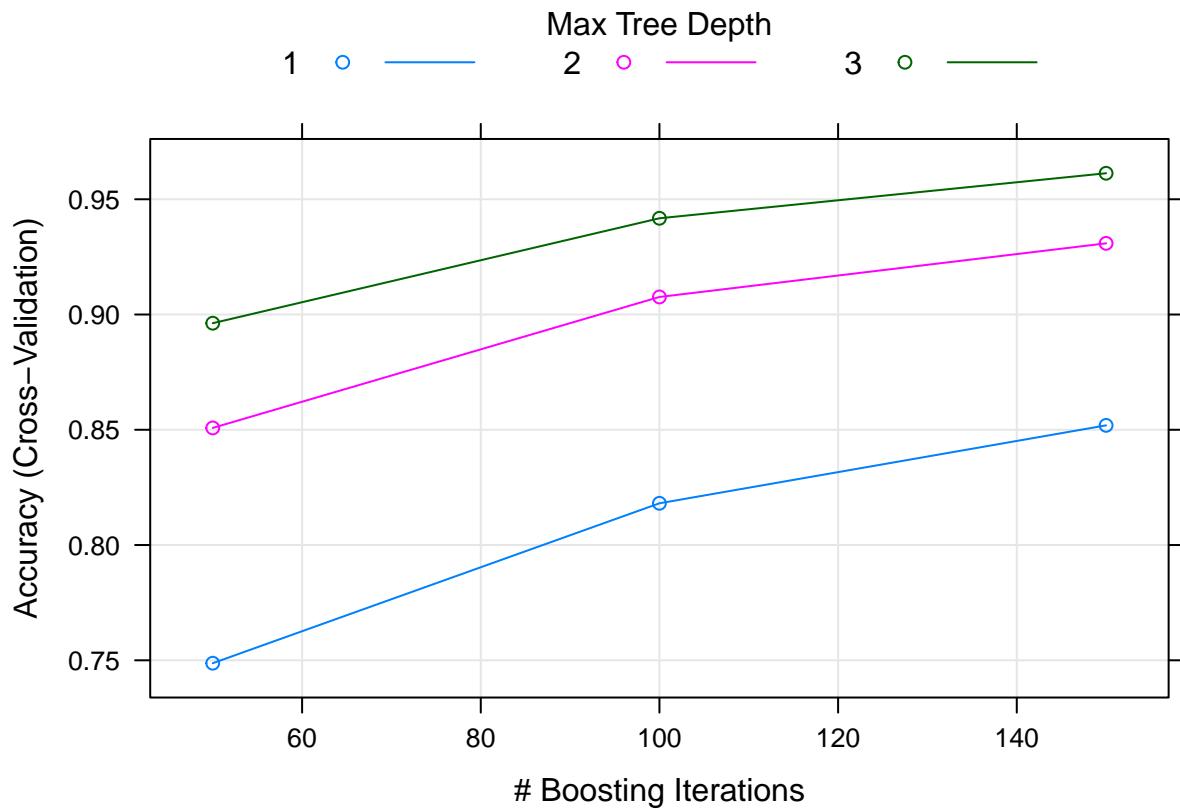
set.seed(1234)
modFitBoost <- train(classe ~ ., method = "gbm", data = training,
                      verbose = F,
                      trControl = trainControl(method = "cv", number = 10))

modFitBoost

## Stochastic Gradient Boosting
##
## 11776 samples
##      52 predictor
##      5 classes: 'A', 'B', 'C', 'D', 'E'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 10599, 10598, 10598, 10597, 10598, 10598, ...
## Resampling results across tuning parameters:
##
##   interaction.depth  n.trees  Accuracy  Kappa
##   1                  50       0.7487281  0.6815375
##   1                  100      0.8181076  0.7698721
##   1                  150      0.8519049  0.8126774
##   2                  50       0.8508001  0.8110092
##   2                  100      0.9076132  0.8830963
##   2                  150      0.9308794  0.9125384
##   3                  50       0.8962319  0.8686673
##   3                  100      0.9417474  0.9262954
##   3                  150      0.9612798  0.9510161
##
## Tuning parameter 'shrinkage' was held constant at a value of 0.1
##
## Tuning parameter 'n.minobsinnode' was held constant at a value of 10
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were n.trees = 150, interaction.depth =
## 3, shrinkage = 0.1 and n.minobsinnode = 10.

plot(modFitBoost)

```



Predicting with the Random Forest Model

```

prediction <- predict(modFitRF, testing, type = "class")
confusionMatrix(prediction, testing$classe)

## Confusion Matrix and Statistics
##
##          Reference
## Prediction   A    B    C    D    E
##           A 2232   11    0    0    0
##           B    0 1502    9    0    0
##           C    0     5 1356   24    2
##           D    0     0    3 1260    2
##           E    0     0    0    2 1438
##
## Overall Statistics
##
##          Accuracy : 0.9926
##             95% CI : (0.9905, 0.9944)
##    No Information Rate : 0.2845
##    P-Value [Acc > NIR] : < 2.2e-16
##
##          Kappa : 0.9906
##
## McNemar's Test P-Value : NA
##

```

```

## Statistics by Class:
##
##          Class: A Class: B Class: C Class: D Class: E
## Sensitivity      1.0000  0.9895  0.9912  0.9798  0.9972
## Specificity     0.9980  0.9986  0.9952  0.9992  0.9997
## Pos Pred Value   0.9951  0.9940  0.9776  0.9960  0.9986
## Neg Pred Value   1.0000  0.9975  0.9981  0.9960  0.9994
## Prevalence       0.2845  0.1935  0.1744  0.1639  0.1838
## Detection Rate   0.2845  0.1914  0.1728  0.1606  0.1833
## Detection Prevalence 0.2859  0.1926  0.1768  0.1612  0.1835
## Balanced Accuracy 0.9990  0.9940  0.9932  0.9895  0.9985

```

Predicting with the Boosting Model

```

prediction <- predict(modFitBoost, testing)
confusionMatrix(prediction, testing$classe)

```

```

## Confusion Matrix and Statistics
##
##          Reference
## Prediction    A     B     C     D     E
##           A 2203    41     0     0     1
##           B   16 1433    49     5    17
##           C    7   42 1292    43     8
##           D    3    1   23 1231    20
##           E    3    1     4     7 1396
##
##          Overall Statistics
##
##                Accuracy : 0.9629
##                95% CI : (0.9585, 0.967)
##                No Information Rate : 0.2845
##                P-Value [Acc > NIR] : < 2.2e-16
##
##                Kappa : 0.9531
##
##    Mcnemar's Test P-Value : 7.303e-08
##
## Statistics by Class:
##
##          Class: A Class: B Class: C Class: D Class: E
## Sensitivity      0.9870  0.9440  0.9444  0.9572  0.9681
## Specificity     0.9925  0.9863  0.9846  0.9928  0.9977
## Pos Pred Value   0.9813  0.9428  0.9282  0.9632  0.9894
## Neg Pred Value   0.9948  0.9866  0.9882  0.9916  0.9929
## Prevalence       0.2845  0.1935  0.1744  0.1639  0.1838
## Detection Rate   0.2808  0.1826  0.1647  0.1569  0.1779
## Detection Prevalence 0.2861  0.1937  0.1774  0.1629  0.1798
## Balanced Accuracy 0.9898  0.9651  0.9645  0.9750  0.9829

```

Predicting with the Testing Data (pml-testing.csv)

Decision Tree Prediction

```
predictionDT <- predict(modFitDT, dt_testing)
predictionDT

##          A         B         C         D         E
## 1  0.08547009 0.820512821 0.034188034 0.000000000 0.059829060
## 2  0.79082696 0.168867269 0.014593468 0.01945796 0.006254343
## 3  0.05414013 0.184713376 0.205414013 0.16719745 0.388535032
## 4  0.13125000 0.043750000 0.120312500 0.66875000 0.035937500
## 5  0.69827586 0.149425287 0.022988506 0.10344828 0.025862069
## 6  0.02722772 0.091584158 0.560024752 0.16584158 0.155321782
## 7  0.07243461 0.140845070 0.026156942 0.67806841 0.082494970
## 8  0.69827586 0.149425287 0.022988506 0.10344828 0.025862069
## 9  0.99782372 0.002176279 0.000000000 0.000000000 0.000000000
## 10 0.79082696 0.168867269 0.014593468 0.01945796 0.006254343
## 11 0.02722772 0.091584158 0.560024752 0.16584158 0.155321782
## 12 0.05414013 0.184713376 0.205414013 0.16719745 0.388535032
## 13 0.02722772 0.091584158 0.560024752 0.16584158 0.155321782
## 14 0.99782372 0.002176279 0.000000000 0.000000000 0.000000000
## 15 0.05414013 0.184713376 0.205414013 0.16719745 0.388535032
## 16 0.05039788 0.334217507 0.000000000 0.11140584 0.503978780
## 17 0.98170732 0.000000000 0.012195122 0.000000000 0.006097561
## 18 0.09556314 0.419795222 0.003412969 0.33447099 0.146757679
## 19 0.09556314 0.419795222 0.003412969 0.33447099 0.146757679
## 20 0.04878049 0.792682927 0.012195122 0.04268293 0.103658537
```

Random Forest Prediction

```
predictionRF <- predict(modFitRF, dt_testing)
predictionRF

##  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
```

Boosting Prediction

```
predictionBoost <- predict(modFitBoost, dt_testing)
predictionBoost

## [1] 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
```