practical machine learning course project

YH

January 5, 2017

#Background Introduction  
#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"  
  
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

library(rpart)  
library(rpart.plot)  
library(RColorBrewer)  
library(rattle)

## Rattle: A free graphical interface for data mining with R.  
## Version 4.1.0 Copyright (c) 2006-2015 Togaware Pty Ltd.  
## Type 'rattle()' to shake, rattle, and roll your data.

set.seed(123456)  
if (!file.exists("data")) { dir.create("data") }  
setwd("C:/Users/yhe/Desktop/data")  
  
#Downloading data  
trainUrl <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"  
testUrl <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"  
download.file(trainUrl, destfile = "./trainUrl")  
download.file(testUrl, destfile = "./testUrl")  
  
data\_train <- read.csv(trainUrl, na.strings=c("NA","#DIV/0!",""))  
data\_test <- read.csv(testUrl, na.strings=c("NA","#DIV/0!",""))  
  
#Partioning Training data set into two data sets, 70% for myTraining, 30% for myTesting:  
forTraining <- createDataPartition(y=data\_train$classe, p=0.7, list=FALSE)  
myTraining <- data\_train[forTraining, ]  
myTesting <- data\_train[-forTraining, ]  
dim(myTraining)

## [1] 13737 160

dim(myTesting)

## [1] 5885 160

head(myTraining)

## X user\_name raw\_timestamp\_part\_1 raw\_timestamp\_part\_2 cvtd\_timestamp  
## 1 1 carlitos 1323084231 788290 05/12/2011 11:23  
## 3 3 carlitos 1323084231 820366 05/12/2011 11:23  
## 5 5 carlitos 1323084232 196328 05/12/2011 11:23  
## 7 7 carlitos 1323084232 368296 05/12/2011 11:23  
## 8 8 carlitos 1323084232 440390 05/12/2011 11:23  
## 10 10 carlitos 1323084232 484434 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  
## 3 no 11 1.42 8.07 -94.4 3  
## 5 no 12 1.48 8.07 -94.4 3  
## 7 no 12 1.42 8.09 -94.4 3  
## 8 no 12 1.42 8.13 -94.4 3  
## 10 no 12 1.45 8.17 -94.4 3  
## kurtosis\_roll\_belt kurtosis\_picth\_belt kurtosis\_yaw\_belt  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## skewness\_roll\_belt skewness\_roll\_belt.1 skewness\_yaw\_belt max\_roll\_belt  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## max\_picth\_belt max\_yaw\_belt min\_roll\_belt min\_pitch\_belt min\_yaw\_belt  
## 1 NA NA NA NA NA  
## 3 NA NA NA NA NA  
## 5 NA NA NA NA NA  
## 7 NA NA NA NA NA  
## 8 NA NA NA NA NA  
## 10 NA NA NA NA NA  
## amplitude\_roll\_belt amplitude\_pitch\_belt amplitude\_yaw\_belt  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## var\_total\_accel\_belt avg\_roll\_belt stddev\_roll\_belt var\_roll\_belt  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## avg\_pitch\_belt stddev\_pitch\_belt var\_pitch\_belt avg\_yaw\_belt  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## stddev\_yaw\_belt var\_yaw\_belt gyros\_belt\_x gyros\_belt\_y gyros\_belt\_z  
## 1 NA NA 0.00 0.00 -0.02  
## 3 NA NA 0.00 0.00 -0.02  
## 5 NA NA 0.02 0.02 -0.02  
## 7 NA NA 0.02 0.00 -0.02  
## 8 NA NA 0.02 0.00 -0.02  
## 10 NA NA 0.03 0.00 0.00  
## accel\_belt\_x accel\_belt\_y accel\_belt\_z magnet\_belt\_x magnet\_belt\_y  
## 1 -21 4 22 -3 599  
## 3 -20 5 23 -2 600  
## 5 -21 2 24 -6 600  
## 7 -22 3 21 -4 599  
## 8 -22 4 21 -2 603  
## 10 -21 4 22 -3 609  
## magnet\_belt\_z roll\_arm pitch\_arm yaw\_arm total\_accel\_arm var\_accel\_arm  
## 1 -313 -128 22.5 -161 34 NA  
## 3 -305 -128 22.5 -161 34 NA  
## 5 -302 -128 22.1 -161 34 NA  
## 7 -311 -128 21.9 -161 34 NA  
## 8 -313 -128 21.8 -161 34 NA  
## 10 -308 -128 21.6 -161 34 NA  
## avg\_roll\_arm stddev\_roll\_arm var\_roll\_arm avg\_pitch\_arm  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## stddev\_pitch\_arm var\_pitch\_arm avg\_yaw\_arm stddev\_yaw\_arm var\_yaw\_arm  
## 1 NA NA NA NA NA  
## 3 NA NA NA NA NA  
## 5 NA NA NA NA NA  
## 7 NA NA NA NA NA  
## 8 NA NA NA NA NA  
## 10 NA NA NA NA NA  
## gyros\_arm\_x gyros\_arm\_y gyros\_arm\_z accel\_arm\_x accel\_arm\_y accel\_arm\_z  
## 1 0.00 0.00 -0.02 -288 109 -123  
## 3 0.02 -0.02 -0.02 -289 110 -126  
## 5 0.00 -0.03 0.00 -289 111 -123  
## 7 0.00 -0.03 0.00 -289 111 -125  
## 8 0.02 -0.02 0.00 -289 111 -124  
## 10 0.02 -0.03 -0.02 -288 110 -124  
## magnet\_arm\_x magnet\_arm\_y magnet\_arm\_z kurtosis\_roll\_arm  
## 1 -368 337 516 NA  
## 3 -368 344 513 NA  
## 5 -374 337 506 NA  
## 7 -373 336 509 NA  
## 8 -372 338 510 NA  
## 10 -376 334 516 NA  
## kurtosis\_picth\_arm kurtosis\_yaw\_arm skewness\_roll\_arm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## skewness\_pitch\_arm skewness\_yaw\_arm max\_roll\_arm max\_picth\_arm  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## max\_yaw\_arm min\_roll\_arm min\_pitch\_arm min\_yaw\_arm amplitude\_roll\_arm  
## 1 NA NA NA NA NA  
## 3 NA NA NA NA NA  
## 5 NA NA NA NA NA  
## 7 NA NA NA NA NA  
## 8 NA NA NA NA NA  
## 10 NA NA NA NA NA  
## amplitude\_pitch\_arm amplitude\_yaw\_arm roll\_dumbbell pitch\_dumbbell  
## 1 NA NA 13.05217 -70.49400  
## 3 NA NA 12.85075 -70.27812  
## 5 NA NA 13.37872 -70.42856  
## 7 NA NA 13.12695 -70.24757  
## 8 NA NA 12.75083 -70.34768  
## 10 NA NA 13.33034 -70.85059  
## yaw\_dumbbell kurtosis\_roll\_dumbbell kurtosis\_picth\_dumbbell  
## 1 -84.87394 NA NA  
## 3 -85.14078 NA NA  
## 5 -84.85306 NA NA  
## 7 -85.09961 NA NA  
## 8 -85.09708 NA NA  
## 10 -84.44602 NA NA  
## kurtosis\_yaw\_dumbbell skewness\_roll\_dumbbell skewness\_pitch\_dumbbell  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## skewness\_yaw\_dumbbell max\_roll\_dumbbell max\_picth\_dumbbell  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## max\_yaw\_dumbbell min\_roll\_dumbbell min\_pitch\_dumbbell min\_yaw\_dumbbell  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## amplitude\_roll\_dumbbell amplitude\_pitch\_dumbbell amplitude\_yaw\_dumbbell  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## total\_accel\_dumbbell var\_accel\_dumbbell avg\_roll\_dumbbell  
## 1 37 NA NA  
## 3 37 NA NA  
## 5 37 NA NA  
## 7 37 NA NA  
## 8 37 NA NA  
## 10 37 NA NA  
## stddev\_roll\_dumbbell var\_roll\_dumbbell avg\_pitch\_dumbbell  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## stddev\_pitch\_dumbbell var\_pitch\_dumbbell avg\_yaw\_dumbbell  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## stddev\_yaw\_dumbbell var\_yaw\_dumbbell gyros\_dumbbell\_x gyros\_dumbbell\_y  
## 1 NA NA 0 -0.02  
## 3 NA NA 0 -0.02  
## 5 NA NA 0 -0.02  
## 7 NA NA 0 -0.02  
## 8 NA NA 0 -0.02  
## 10 NA NA 0 -0.02  
## gyros\_dumbbell\_z accel\_dumbbell\_x accel\_dumbbell\_y accel\_dumbbell\_z  
## 1 0 -234 47 -271  
## 3 0 -232 46 -270  
## 5 0 -233 48 -270  
## 7 0 -232 47 -270  
## 8 0 -234 46 -272  
## 10 0 -235 48 -270  
## magnet\_dumbbell\_x magnet\_dumbbell\_y magnet\_dumbbell\_z roll\_forearm  
## 1 -559 293 -65 28.4  
## 3 -561 298 -63 28.3  
## 5 -554 292 -68 28.0  
## 7 -551 295 -70 27.9  
## 8 -555 300 -74 27.8  
## 10 -558 291 -69 27.7  
## pitch\_forearm yaw\_forearm kurtosis\_roll\_forearm kurtosis\_picth\_forearm  
## 1 -63.9 -153 NA NA  
## 3 -63.9 -152 NA NA  
## 5 -63.9 -152 NA NA  
## 7 -63.9 -152 NA NA  
## 8 -63.8 -152 NA NA  
## 10 -63.8 -152 NA NA  
## kurtosis\_yaw\_forearm skewness\_roll\_forearm skewness\_pitch\_forearm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## skewness\_yaw\_forearm max\_roll\_forearm max\_picth\_forearm max\_yaw\_forearm  
## 1 NA NA NA NA  
## 3 NA NA NA NA  
## 5 NA NA NA NA  
## 7 NA NA NA NA  
## 8 NA NA NA NA  
## 10 NA NA NA NA  
## min\_roll\_forearm min\_pitch\_forearm min\_yaw\_forearm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## amplitude\_roll\_forearm amplitude\_pitch\_forearm amplitude\_yaw\_forearm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## total\_accel\_forearm var\_accel\_forearm avg\_roll\_forearm  
## 1 36 NA NA  
## 3 36 NA NA  
## 5 36 NA NA  
## 7 36 NA NA  
## 8 36 NA NA  
## 10 36 NA NA  
## stddev\_roll\_forearm var\_roll\_forearm avg\_pitch\_forearm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## stddev\_pitch\_forearm var\_pitch\_forearm avg\_yaw\_forearm  
## 1 NA NA NA  
## 3 NA NA NA  
## 5 NA NA NA  
## 7 NA NA NA  
## 8 NA NA NA  
## 10 NA NA NA  
## stddev\_yaw\_forearm var\_yaw\_forearm gyros\_forearm\_x gyros\_forearm\_y  
## 1 NA NA 0.03 0.00  
## 3 NA NA 0.03 -0.02  
## 5 NA NA 0.02 0.00  
## 7 NA NA 0.02 0.00  
## 8 NA NA 0.02 -0.02  
## 10 NA NA 0.02 0.00  
## gyros\_forearm\_z accel\_forearm\_x accel\_forearm\_y accel\_forearm\_z  
## 1 -0.02 192 203 -215  
## 3 0.00 196 204 -213  
## 5 -0.02 189 206 -214  
## 7 -0.02 195 205 -215  
## 8 0.00 193 205 -213  
## 10 -0.02 190 205 -215  
## magnet\_forearm\_x magnet\_forearm\_y magnet\_forearm\_z classe  
## 1 -17 654 476 A  
## 3 -18 658 469 A  
## 5 -17 655 473 A  
## 7 -18 659 470 A  
## 8 -9 660 474 A  
## 10 -22 656 473 A

head(myTesting)

## X user\_name raw\_timestamp\_part\_1 raw\_timestamp\_part\_2 cvtd\_timestamp  
## 2 2 carlitos 1323084231 808298 05/12/2011 11:23  
## 4 4 carlitos 1323084232 120339 05/12/2011 11:23  
## 6 6 carlitos 1323084232 304277 05/12/2011 11:23  
## 9 9 carlitos 1323084232 484323 05/12/2011 11:23  
## 12 12 carlitos 1323084232 528316 05/12/2011 11:23  
## 15 15 carlitos 1323084232 604281 05/12/2011 11:23  
## new\_window num\_window roll\_belt pitch\_belt yaw\_belt total\_accel\_belt  
## 2 no 11 1.41 8.07 -94.4 3  
## 4 no 12 1.48 8.05 -94.4 3  
## 6 no 12 1.45 8.06 -94.4 3  
## 9 no 12 1.43 8.16 -94.4 3  
## 12 no 12 1.43 8.18 -94.4 3  
## 15 no 12 1.45 8.20 -94.4 3  
## kurtosis\_roll\_belt kurtosis\_picth\_belt kurtosis\_yaw\_belt  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## skewness\_roll\_belt skewness\_roll\_belt.1 skewness\_yaw\_belt max\_roll\_belt  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## max\_picth\_belt max\_yaw\_belt min\_roll\_belt min\_pitch\_belt min\_yaw\_belt  
## 2 NA NA NA NA NA  
## 4 NA NA NA NA NA  
## 6 NA NA NA NA NA  
## 9 NA NA NA NA NA  
## 12 NA NA NA NA NA  
## 15 NA NA NA NA NA  
## amplitude\_roll\_belt amplitude\_pitch\_belt amplitude\_yaw\_belt  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## var\_total\_accel\_belt avg\_roll\_belt stddev\_roll\_belt var\_roll\_belt  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## avg\_pitch\_belt stddev\_pitch\_belt var\_pitch\_belt avg\_yaw\_belt  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## stddev\_yaw\_belt var\_yaw\_belt gyros\_belt\_x gyros\_belt\_y gyros\_belt\_z  
## 2 NA NA 0.02 0 -0.02  
## 4 NA NA 0.02 0 -0.03  
## 6 NA NA 0.02 0 -0.02  
## 9 NA NA 0.02 0 -0.02  
## 12 NA NA 0.02 0 -0.02  
## 15 NA NA 0.00 0 0.00  
## accel\_belt\_x accel\_belt\_y accel\_belt\_z magnet\_belt\_x magnet\_belt\_y  
## 2 -22 4 22 -7 608  
## 4 -22 3 21 -6 604  
## 6 -21 4 21 0 603  
## 9 -20 2 24 1 602  
## 12 -22 2 23 -2 602  
## 15 -21 2 22 -1 597  
## magnet\_belt\_z roll\_arm pitch\_arm yaw\_arm total\_accel\_arm var\_accel\_arm  
## 2 -311 -128 22.5 -161 34 NA  
## 4 -310 -128 22.1 -161 34 NA  
## 6 -312 -128 22.0 -161 34 NA  
## 9 -312 -128 21.7 -161 34 NA  
## 12 -319 -128 21.5 -161 34 NA  
## 15 -310 -129 21.4 -161 34 NA  
## avg\_roll\_arm stddev\_roll\_arm var\_roll\_arm avg\_pitch\_arm  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## stddev\_pitch\_arm var\_pitch\_arm avg\_yaw\_arm stddev\_yaw\_arm var\_yaw\_arm  
## 2 NA NA NA NA NA  
## 4 NA NA NA NA NA  
## 6 NA NA NA NA NA  
## 9 NA NA NA NA NA  
## 12 NA NA NA NA NA  
## 15 NA NA NA NA NA  
## gyros\_arm\_x gyros\_arm\_y gyros\_arm\_z accel\_arm\_x accel\_arm\_y accel\_arm\_z  
## 2 0.02 -0.02 -0.02 -290 110 -125  
## 4 0.02 -0.03 0.02 -289 111 -123  
## 6 0.02 -0.03 0.00 -289 111 -122  
## 9 0.02 -0.03 -0.02 -288 109 -122  
## 12 0.02 -0.03 0.00 -288 111 -123  
## 15 0.02 0.00 -0.03 -289 111 -124  
## magnet\_arm\_x magnet\_arm\_y magnet\_arm\_z kurtosis\_roll\_arm  
## 2 -369 337 513 NA  
## 4 -372 344 512 NA  
## 6 -369 342 513 NA  
## 9 -369 341 518 NA  
## 12 -363 343 520 NA  
## 15 -374 342 510 NA  
## kurtosis\_picth\_arm kurtosis\_yaw\_arm skewness\_roll\_arm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## skewness\_pitch\_arm skewness\_yaw\_arm max\_roll\_arm max\_picth\_arm  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## max\_yaw\_arm min\_roll\_arm min\_pitch\_arm min\_yaw\_arm amplitude\_roll\_arm  
## 2 NA NA NA NA NA  
## 4 NA NA NA NA NA  
## 6 NA NA NA NA NA  
## 9 NA NA NA NA NA  
## 12 NA NA NA NA NA  
## 15 NA NA NA NA NA  
## amplitude\_pitch\_arm amplitude\_yaw\_arm roll\_dumbbell pitch\_dumbbell  
## 2 NA NA 13.13074 -70.63751  
## 4 NA NA 13.43120 -70.39379  
## 6 NA NA 13.38246 -70.81759  
## 9 NA NA 13.15463 -70.42520  
## 12 NA NA 13.10321 -70.45975  
## 15 NA NA 13.07949 -70.67116  
## yaw\_dumbbell kurtosis\_roll\_dumbbell kurtosis\_picth\_dumbbell  
## 2 -84.71065 NA NA  
## 4 -84.87363 NA NA  
## 6 -84.46500 NA NA  
## 9 -84.91563 NA NA  
## 12 -84.89472 NA NA  
## 15 -84.69053 NA NA  
## kurtosis\_yaw\_dumbbell skewness\_roll\_dumbbell skewness\_pitch\_dumbbell  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## skewness\_yaw\_dumbbell max\_roll\_dumbbell max\_picth\_dumbbell  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## max\_yaw\_dumbbell min\_roll\_dumbbell min\_pitch\_dumbbell min\_yaw\_dumbbell  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## amplitude\_roll\_dumbbell amplitude\_pitch\_dumbbell amplitude\_yaw\_dumbbell  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## total\_accel\_dumbbell var\_accel\_dumbbell avg\_roll\_dumbbell  
## 2 37 NA NA  
## 4 37 NA NA  
## 6 37 NA NA  
## 9 37 NA NA  
## 12 37 NA NA  
## 15 37 NA NA  
## stddev\_roll\_dumbbell var\_roll\_dumbbell avg\_pitch\_dumbbell  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## stddev\_pitch\_dumbbell var\_pitch\_dumbbell avg\_yaw\_dumbbell  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## stddev\_yaw\_dumbbell var\_yaw\_dumbbell gyros\_dumbbell\_x gyros\_dumbbell\_y  
## 2 NA NA 0 -0.02  
## 4 NA NA 0 -0.02  
## 6 NA NA 0 -0.02  
## 9 NA NA 0 -0.02  
## 12 NA NA 0 -0.02  
## 15 NA NA 0 -0.02  
## gyros\_dumbbell\_z accel\_dumbbell\_x accel\_dumbbell\_y accel\_dumbbell\_z  
## 2 0.00 -233 47 -269  
## 4 -0.02 -232 48 -269  
## 6 0.00 -234 48 -269  
## 9 0.00 -232 47 -269  
## 12 0.00 -233 47 -270  
## 15 0.00 -234 47 -270  
## magnet\_dumbbell\_x magnet\_dumbbell\_y magnet\_dumbbell\_z roll\_forearm  
## 2 -555 296 -64 28.3  
## 4 -552 303 -60 28.1  
## 6 -558 294 -66 27.9  
## 9 -549 292 -65 27.7  
## 12 -554 291 -65 27.5  
## 15 -554 294 -63 27.2  
## pitch\_forearm yaw\_forearm kurtosis\_roll\_forearm kurtosis\_picth\_forearm  
## 2 -63.9 -153 NA NA  
## 4 -63.9 -152 NA NA  
## 6 -63.9 -152 NA NA  
## 9 -63.8 -152 NA NA  
## 12 -63.8 -152 NA NA  
## 15 -63.9 -151 NA NA  
## kurtosis\_yaw\_forearm skewness\_roll\_forearm skewness\_pitch\_forearm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## skewness\_yaw\_forearm max\_roll\_forearm max\_picth\_forearm max\_yaw\_forearm  
## 2 NA NA NA NA  
## 4 NA NA NA NA  
## 6 NA NA NA NA  
## 9 NA NA NA NA  
## 12 NA NA NA NA  
## 15 NA NA NA NA  
## min\_roll\_forearm min\_pitch\_forearm min\_yaw\_forearm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## amplitude\_roll\_forearm amplitude\_pitch\_forearm amplitude\_yaw\_forearm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## total\_accel\_forearm var\_accel\_forearm avg\_roll\_forearm  
## 2 36 NA NA  
## 4 36 NA NA  
## 6 36 NA NA  
## 9 36 NA NA  
## 12 36 NA NA  
## 15 36 NA NA  
## stddev\_roll\_forearm var\_roll\_forearm avg\_pitch\_forearm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## stddev\_pitch\_forearm var\_pitch\_forearm avg\_yaw\_forearm  
## 2 NA NA NA  
## 4 NA NA NA  
## 6 NA NA NA  
## 9 NA NA NA  
## 12 NA NA NA  
## 15 NA NA NA  
## stddev\_yaw\_forearm var\_yaw\_forearm gyros\_forearm\_x gyros\_forearm\_y  
## 2 NA NA 0.02 0.00  
## 4 NA NA 0.02 -0.02  
## 6 NA NA 0.02 -0.02  
## 9 NA NA 0.03 0.00  
## 12 NA NA 0.02 0.02  
## 15 NA NA 0.00 -0.02  
## gyros\_forearm\_z accel\_forearm\_x accel\_forearm\_y accel\_forearm\_z  
## 2 -0.02 192 203 -216  
## 4 0.00 189 206 -214  
## 6 -0.03 193 203 -215  
## 9 -0.02 193 204 -214  
## 12 -0.03 191 203 -215  
## 15 -0.02 192 201 -214  
## magnet\_forearm\_x magnet\_forearm\_y magnet\_forearm\_z classe  
## 2 -18 661 473 A  
## 4 -16 658 469 A  
## 6 -9 660 478 A  
## 9 -16 653 476 A  
## 12 -11 657 478 A  
## 15 -16 656 472 A

#Cleaning the data---Remove NearZeroVariance variables and NAs:  
nzv <- nearZeroVar(myTraining, saveMetrics=TRUE)  
myTraining <- myTraining[,nzv$nzv==FALSE]  
  
nzv<- nearZeroVar(myTesting,saveMetrics=TRUE)  
myTesting <- myTesting[,nzv$nzv==FALSE]  
  
myTraining <- myTraining[c(-1)]  
  
  
myTraining\_final <- myTraining   
for(i in 1:length(myTraining)){  
 if( sum( is.na( myTraining[, i] ) ) /nrow(myTraining) >= .7 ) {  
 for(j in 1:length(myTraining\_final)) {  
 if( length( grep(names(myTraining[i]), names(myTraining\_final)[j]) ) ==1) {  
 myTraining\_final <- myTraining\_final[ , -j]  
 }  
 }  
 }  
}  
dim(myTraining\_final)

## [1] 13737 58

myTraining <- myTraining\_final  
rm(myTraining\_final)  
  
myTesting\_final <- myTesting  
for(i in 1:length(myTesting)){  
 if( sum( is.na( myTesting[, i] ) ) /nrow(myTesting) >= .7 ) {  
 for(j in 1:length(myTesting\_final)) {  
 if( length( grep(names(myTesting[i]), names(myTesting\_final)[j]) ) ==1) {  
 myTesting\_final <- myTesting\_final[ , -j]  
 }  
 }  
 }  
}  
dim(myTesting\_final)

## [1] 5885 59

myTesting <- myTesting\_final  
rm(myTesting\_final)  
  
#Train Candidate Model and Cross Validate---Random Forest model  
fitControl <- trainControl(method="cv", number=3, verboseIter=F)  
fit <- train(classe ~ ., data=myTraining, method="rf", trControl=fitControl)

## Loading required package: randomForest

## randomForest 4.6-12

## Type rfNews() to see new features/changes/bug fixes.

##   
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':  
##   
## margin

fit$finalModel

##   
## Call:  
## randomForest(x = x, y = y, mtry = param$mtry)   
## Type of random forest: classification  
## Number of trees: 500  
## No. of variables tried at each split: 79  
##   
## OOB estimate of error rate: 0.14%  
## Confusion matrix:  
## A B C D E class.error  
## A 3906 0 0 0 0 0.000000000  
## B 5 2649 4 0 0 0.003386005  
## C 0 3 2392 1 0 0.001669449  
## D 0 0 4 2246 2 0.002664298  
## E 0 0 0 0 2525 0.000000000

# use model to predict classe in validation set (myTesting)  
preds <- predict(fit, newdata=myTesting)  
  
# show confusion matrix to get estimate of out-of-sample error  
confusionMatrix(myTesting$classe, preds)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction A B C D E  
## A 1674 0 0 0 0  
## B 1 1138 0 0 0  
## C 0 3 1023 0 0  
## D 0 0 2 962 0  
## E 0 0 0 0 1082  
##   
## Overall Statistics  
##   
## Accuracy : 0.999   
## 95% CI : (0.9978, 0.9996)  
## No Information Rate : 0.2846   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 0.9987   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: A Class: B Class: C Class: D Class: E  
## Sensitivity 0.9994 0.9974 0.9980 1.0000 1.0000  
## Specificity 1.0000 0.9998 0.9994 0.9996 1.0000  
## Pos Pred Value 1.0000 0.9991 0.9971 0.9979 1.0000  
## Neg Pred Value 0.9998 0.9994 0.9996 1.0000 1.0000  
## Prevalence 0.2846 0.1939 0.1742 0.1635 0.1839  
## Detection Rate 0.2845 0.1934 0.1738 0.1635 0.1839  
## Detection Prevalence 0.2845 0.1935 0.1743 0.1638 0.1839  
## Balanced Accuracy 0.9997 0.9986 0.9987 0.9998 1.0000

predTesting <- predict(fit, data\_test)  
predTesting

## [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

pml\_write\_files = function(x){  
 n = length(x)  
 for(i in 1:n){  
 filename = paste0("problem\_id\_",i,".txt")  
 write.table(x[i],file=filename,quote=FALSE,row.names=FALSE,col.names=FALSE)  
 }  
}  
  
pml\_write\_files(predTesting)  
  
#Conclusion---The model predicted the 20 test cases with 100% accuracy.

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.