

CourseProj

Viraj Bhalala

11/27/2017

Machine Learning Course Project 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://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://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.

What you should submit

The goal of your project is to predict the manner in which they did the exercise. This is the “classe” variable in the training set. You may use any of the other variables to predict with. You should create a report describing how you built your model, how you used cross validation, what you think the expected out of sample error is, and why you made the choices you did. You will also use your prediction model to predict 20 different test cases.

Peer Review Portion

Your submission for the Peer Review portion should consist of a link to a Github repo with your R markdown and compiled HTML file describing your analysis. Please constrain the text of the writeup to < 2000 words and the number of figures to be less than 5. It will make it easier for the graders if you submit a repo with a gh-pages branch so the HTML page can be viewed online (and you always want to make it easy on graders :-).

Approach

```
library(lattice);  
library(ggplot2);  
library(caret);  
library(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
library(rpart);
library(rpart.plot)
```

Including Plots

You can also embed plots, for example:

```
set.seed(1234)

# data load and clean up
trainingset <- read.csv("pml-training.csv", na.strings=c("NA", "#DIV/0!", ""))
testingset <- read.csv("pml-testing.csv", na.strings=c("NA", "#DIV/0!", ""))

head(trainingset)

##      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_pitch_belt kurtosis_yaw_belt
## 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_belt skewness_roll_belt.1 skewness_yaw_belt max_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
##      max_pitch_belt max_yaw_belt min_roll_belt min_pitch_belt min_yaw_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_roll_belt	amplitude_pitch_belt	amplitude_yaw_belt			
## 1	NA	NA	NA			
## 2	NA	NA	NA			
## 3	NA	NA	NA			
## 4	NA	NA	NA			
## 5	NA	NA	NA			
## 6	NA	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	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	
## 2	NA	NA	0.02	0.00	-0.02	
## 3	NA	NA	0.00	0.00	-0.02	
## 4	NA	NA	0.02	0.00	-0.03	
## 5	NA	NA	0.02	0.02	-0.02	
## 6	NA	NA	0.02	0.00	-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	
## 4	-22	3	21	-6	604	
## 5	-21	2	24	-6	600	
## 6	-21	4	21	0	603	
##	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
## 4	-310	-128	22.1	-161	34	NA
## 5	-302	-128	22.1	-161	34	NA
## 6	-312	-128	22.0	-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	
## 4	NA	NA	NA	NA	NA	
## 5	NA	NA	NA	NA	NA	
## 6	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	
## 4	NA	NA	NA	NA	0.02	

## 5	NA	NA	NA	NA	0.00	
## 6	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
## 4	-0.03	0.02	-289	111	-123	-372
## 5	-0.03	0.00	-289	111	-123	-374
## 6	-0.03	0.00	-289	111	-122	-369
##	magnet_arm_y	magnet_arm_z	kurtosis_roll_arm	kurtosis_pitch_arm		
## 1	337	516	NA	NA		
## 2	337	513	NA	NA		
## 3	344	513	NA	NA		
## 4	344	512	NA	NA		
## 5	337	506	NA	NA		
## 6	342	513	NA	NA		
##	kurtosis_yaw_arm	skewness_roll_arm	skewness_pitch_arm	skewness_yaw_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		
##	max_roll_arm	max_pitch_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	
## 4	NA	NA	NA	NA	NA	
## 5	NA	NA	NA	NA	NA	
## 6	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	NA	NA	NA	NA		
##	roll_dumbbell	pitch_dumbbell	yaw_dumbbell	kurtosis_roll_dumbbell		
## 1	13.05217	-70.49400	-84.87394	NA		
## 2	13.13074	-70.63751	-84.71065	NA		
## 3	12.85075	-70.27812	-85.14078	NA		
## 4	13.43120	-70.39379	-84.87363	NA		
## 5	13.37872	-70.42856	-84.85306	NA		
## 6	13.38246	-70.81759	-84.46500	NA		
##	kurtosis_pitch_dumbbell	kurtosis_yaw_dumbbell	skewness_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			
##	skewness_pitch_dumbbell	skewness_yaw_dumbbell	max_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	
##	max_pitch_dumbbell	max_yaw_dumbbell	min_roll_dumbbell	min_pitch_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_yaw_dumbbell	amplitude_roll_dumbbell	amplitude_pitch_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_yaw_dumbbell	total_accel_dumbbell	var_accel_dumbbell	
## 1	NA	37	NA	
## 2	NA	37	NA	
## 3	NA	37	NA	
## 4	NA	37	NA	
## 5	NA	37	NA	
## 6	NA	37	NA	
##	avg_roll_dumbbell	stddev_roll_dumbbell	var_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	
##	avg_pitch_dumbbell	stddev_pitch_dumbbell	var_pitch_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	
##	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
## 4	NA	NA	NA	0
## 5	NA	NA	NA	0
## 6	NA	NA	NA	0
##	gyros_dumbbell_y	gyros_dumbbell_z	accel_dumbbell_x	accel_dumbbell_y
## 1	-0.02	0.00	-234	47
## 2	-0.02	0.00	-233	47
## 3	-0.02	0.00	-232	46
## 4	-0.02	-0.02	-232	48
## 5	-0.02	0.00	-233	48
## 6	-0.02	0.00	-234	48
##	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
## 4	-269	-552	303	-60
## 5	-270	-554	292	-68
## 6	-269	-558	294	-66
##	roll_forearm	pitch_forearm	yaw_forearm	kurtosis_roll_forearm
## 1	28.4	-63.9	-153	NA
## 2	28.3	-63.9	-153	NA
## 3	28.3	-63.9	-152	NA
## 4	28.1	-63.9	-152	NA
## 5	28.0	-63.9	-152	NA
## 6	27.9	-63.9	-152	NA
##	kurtosis_pitch_forearm	kurtosis_yaw_forearm	skewness_roll_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_pitch_forearm	skewness_yaw_forearm	max_roll_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	
##	max_pitch_forearm	max_yaw_forearm	min_roll_forearm	min_pitch_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_yaw_forearm	amplitude_roll_forearm	amplitude_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	
##	amplitude_yaw_forearm	total_accel_forearm	var_accel_forearm	
## 1	NA	36	NA	
## 2	NA	36	NA	
## 3	NA	36	NA	
## 4	NA	36	NA	
## 5	NA	36	NA	
## 6	NA	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
## 4	NA	NA	NA	NA
## 5	NA	NA	NA	NA

```
## 6      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
## 4      NA      NA      NA
## 5      NA      NA      NA
## 6      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
```

```
head(testingset)
```

```
## 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_pitch_belt kurtosis_yaw_belt
## 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_belt skewness_roll_belt.1 skewness_yaw_belt max_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		
##	max_pitch_belt	max_yaw_belt	min_roll_belt	min_pitch_belt	min_yaw_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_roll_belt	amplitude_pitch_belt	amplitude_yaw_belt			
## 1	NA	NA	NA			
## 2	NA	NA	NA			
## 3	NA	NA	NA			
## 4	NA	NA	NA			
## 5	NA	NA	NA			
## 6	NA	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	NA	NA	NA	NA		
##	stddev_yaw_belt	var_yaw_belt	gyros_belt_x	gyros_belt_y	gyros_belt_z	
## 1	NA	NA	-0.50	-0.02	-0.46	
## 2	NA	NA	-0.06	-0.02	-0.07	
## 3	NA	NA	0.05	0.02	0.03	
## 4	NA	NA	0.11	0.11	-0.16	
## 5	NA	NA	0.03	0.02	0.00	
## 6	NA	NA	0.10	0.05	-0.13	
##	accel_belt_x	accel_belt_y	accel_belt_z	magnet_belt_x	magnet_belt_y	
## 1	-38	69	-179	-13	581	
## 2	-13	11	39	43	636	
## 3	1	-1	49	29	631	
## 4	46	45	-156	169	608	
## 5	-8	4	27	33	566	
## 6	-11	-16	38	31	638	
##	magnet_belt_z	roll_arm	pitch_arm	yaw_arm	total_accel_arm	var_accel_arm
## 1	-382	40.7	-27.80	178	10	NA
## 2	-309	0.0	0.00	0	38	NA
## 3	-312	0.0	0.00	0	44	NA
## 4	-304	-109.0	55.00	-142	25	NA
## 5	-418	76.1	2.76	102	29	NA
## 6	-291	0.0	0.00	0	14	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	
## 4	NA	NA	NA	NA	NA	
## 5	NA	NA	NA	NA	NA	
## 6	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	-1.65	
## 2	NA	NA	NA	NA	-1.17	
## 3	NA	NA	NA	NA	2.10	
## 4	NA	NA	NA	NA	0.22	
## 5	NA	NA	NA	NA	-1.96	
## 6	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	0.48	-0.18	16	38	93	-326
## 2	0.85	-0.43	-290	215	-90	-325
## 3	-1.36	1.13	-341	245	-87	-264
## 4	-0.51	0.92	-238	-57	6	-173
## 5	0.79	-0.54	-197	200	-30	-170
## 6	0.05	-0.07	-26	130	-19	396
##	magnet_arm_y	magnet_arm_z	kurtosis_roll_arm	kurtosis_pitch_arm		
## 1	385	481	NA	NA		
## 2	447	434	NA	NA		
## 3	474	413	NA	NA		
## 4	257	633	NA	NA		
## 5	275	617	NA	NA		
## 6	176	516	NA	NA		
##	kurtosis_yaw_arm	skewness_roll_arm	skewness_pitch_arm	skewness_yaw_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		
##	max_roll_arm	max_pitch_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	
## 4	NA	NA	NA	NA	NA	
## 5	NA	NA	NA	NA	NA	
## 6	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		
## 4	NA	NA	NA	NA		
## 5	NA	NA	NA	NA		
## 6	NA	NA	NA	NA		
##	roll_dumbbell	pitch_dumbbell	yaw_dumbbell	kurtosis_roll_dumbbell		
## 1	-17.73748	24.96085	126.23596	NA		
## 2	54.47761	-53.69758	-75.51480	NA		
## 3	57.07031	-51.37303	-75.20287	NA		
## 4	43.10927	-30.04885	-103.32003	NA		

## 5	-101.38396	-53.43952	-14.19542	NA
## 6	62.18750	-50.55595	-71.12063	NA
##	kurtosis_picth_dumbbell	kurtosis_yaw_dumbbell	skewness_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
##	skewness_pitch_dumbbell	skewness_yaw_dumbbell	max_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
##	max_picth_dumbbell	max_yaw_dumbbell	min_roll_dumbbell	min_pitch_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_yaw_dumbbell	amplitude_roll_dumbbell	amplitude_pitch_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_yaw_dumbbell	total_accel_dumbbell	var_accel_dumbbell	
## 1	NA	9	NA	NA
## 2	NA	31	NA	NA
## 3	NA	29	NA	NA
## 4	NA	18	NA	NA
## 5	NA	4	NA	NA
## 6	NA	29	NA	NA
##	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	
## 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_yaw_dumbbell	stddev_yaw_dumbbell	var_yaw_dumbbell	gyros_dumbbell_x
## 1	NA	NA	NA	0.64
## 2	NA	NA	NA	0.34

## 3	NA	NA	NA	0.39
## 4	NA	NA	NA	0.10
## 5	NA	NA	NA	0.29
## 6	NA	NA	NA	-0.59
##	gyros_dumbbell_y	gyros_dumbbell_z	accel_dumbbell_x	accel_dumbbell_y
## 1	0.06	-0.61	21	-15
## 2	0.05	-0.71	-153	155
## 3	0.14	-0.34	-141	155
## 4	-0.02	0.05	-51	72
## 5	-0.47	-0.46	-18	-30
## 6	0.80	1.10	-138	166
##	accel_dumbbell_z	magnet_dumbbell_x	magnet_dumbbell_y	magnet_dumbbell_z
## 1	81	523	-528	-56
## 2	-205	-502	388	-36
## 3	-196	-506	349	41
## 4	-148	-576	238	53
## 5	-5	-424	252	312
## 6	-186	-543	262	96
##	roll_forearm	pitch_forearm	yaw_forearm	kurtosis_roll_forearm
## 1	141	49.30	156.0	NA
## 2	109	-17.60	106.0	NA
## 3	131	-32.60	93.0	NA
## 4	0	0.00	0.0	NA
## 5	-176	-2.16	-47.9	NA
## 6	150	1.46	89.7	NA
##	kurtosis_pitch_forearm	kurtosis_yaw_forearm	skewness_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
##	skewness_pitch_forearm	skewness_yaw_forearm	max_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
##	max_pitch_forearm	max_yaw_forearm	min_roll_forearm	min_pitch_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_yaw_forearm	amplitude_roll_forearm	amplitude_pitch_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_yaw_forearm	total_accel_forearm	var_accel_forearm	

## 1	NA	33	NA	
## 2	NA	39	NA	
## 3	NA	34	NA	
## 4	NA	43	NA	
## 5	NA	24	NA	
## 6	NA	43	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
## 4	NA	NA	NA	NA
## 5	NA	NA	NA	NA
## 6	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	
## 4	NA	NA	NA	
## 5	NA	NA	NA	
## 6	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

Clean dataset

```

trainingset<-trainingset[,colSums(is.na(trainingset)) == 0]
testingset <-testingset[,colSums(is.na(testingset)) == 0]
names(trainingset)

```

## [1]	"X"	"user_name"	"raw_timestamp_part_1"
## [4]	"raw_timestamp_part_2"	"cvtd_timestamp"	"new_window"
## [7]	"num_window"	"roll_belt"	"pitch_belt"
## [10]	"yaw_belt"	"total_accel_belt"	"gyros_belt_x"
## [13]	"gyros_belt_y"	"gyros_belt_z"	"accel_belt_x"
## [16]	"accel_belt_y"	"accel_belt_z"	"magnet_belt_x"
## [19]	"magnet_belt_y"	"magnet_belt_z"	"roll_arm"

```
## [22] "pitch_arm"          "yaw_arm"          "total_accel_arm"
## [25] "gyros_arm_x"        "gyros_arm_y"      "gyros_arm_z"
## [28] "accel_arm_x"        "accel_arm_y"      "accel_arm_z"
## [31] "magnet_arm_x"       "magnet_arm_y"     "magnet_arm_z"
## [34] "roll_dumbbell"     "pitch_dumbbell"   "yaw_dumbbell"
## [37] "total_accel_dumbbell" "gyros_dumbbell_x" "gyros_dumbbell_y"
## [40] "gyros_dumbbell_z"   "accel_dumbbell_x" "accel_dumbbell_y"
## [43] "accel_dumbbell_z"   "magnet_dumbbell_x" "magnet_dumbbell_y"
## [46] "magnet_dumbbell_z"   "roll_forearm"     "pitch_forearm"
## [49] "yaw_forearm"        "total_accel_forearm" "gyros_forearm_x"
## [52] "gyros_forearm_y"    "gyros_forearm_z"   "accel_forearm_x"
## [55] "accel_forearm_y"    "accel_forearm_z"   "magnet_forearm_x"
## [58] "magnet_forearm_y"   "magnet_forearm_z"   "classe"
```

```
names(testingset)
```

```
## [1] "X"                  "user_name"        "raw_timestamp_part_1"
## [4] "raw_timestamp_part_2" "cvt_d_timestamp"   "new_window"
## [7] "num_window"        "roll_belt"        "pitch_belt"
## [10] "yaw_belt"          "total_accel_belt"  "gyros_belt_x"
## [13] "gyros_belt_y"       "gyros_belt_z"      "accel_belt_x"
## [16] "accel_belt_y"       "accel_belt_z"      "magnet_belt_x"
## [19] "magnet_belt_y"      "magnet_belt_z"     "roll_arm"
## [22] "pitch_arm"         "yaw_arm"          "total_accel_arm"
## [25] "gyros_arm_x"        "gyros_arm_y"      "gyros_arm_z"
## [28] "accel_arm_x"        "accel_arm_y"      "accel_arm_z"
## [31] "magnet_arm_x"       "magnet_arm_y"     "magnet_arm_z"
## [34] "roll_dumbbell"     "pitch_dumbbell"   "yaw_dumbbell"
## [37] "total_accel_dumbbell" "gyros_dumbbell_x" "gyros_dumbbell_y"
## [40] "gyros_dumbbell_z"   "accel_dumbbell_x" "accel_dumbbell_y"
## [43] "accel_dumbbell_z"   "magnet_dumbbell_x" "magnet_dumbbell_y"
## [46] "magnet_dumbbell_z"   "roll_forearm"     "pitch_forearm"
## [49] "yaw_forearm"        "total_accel_forearm" "gyros_forearm_x"
## [52] "gyros_forearm_y"    "gyros_forearm_z"   "accel_forearm_x"
## [55] "accel_forearm_y"    "accel_forearm_z"   "magnet_forearm_x"
## [58] "magnet_forearm_y"   "magnet_forearm_z"   "problem_id"
```

```
#delete irrelevant columns
```

```
trainingset <- trainingset[,-c(1:7)]
```

```
testingset <- testingset[,-c(1:7)]
```

```
# partition the data 70-30 train test split
```

```
traintrainset <- createDataPartition(y=trainingset$classe, p=0.70, list=FALSE)
```

```
TrainTrainingSet <- trainingset[traintrainset, ]
```

```
TestTrainingSet <- trainingset[-traintrainset, ]
```

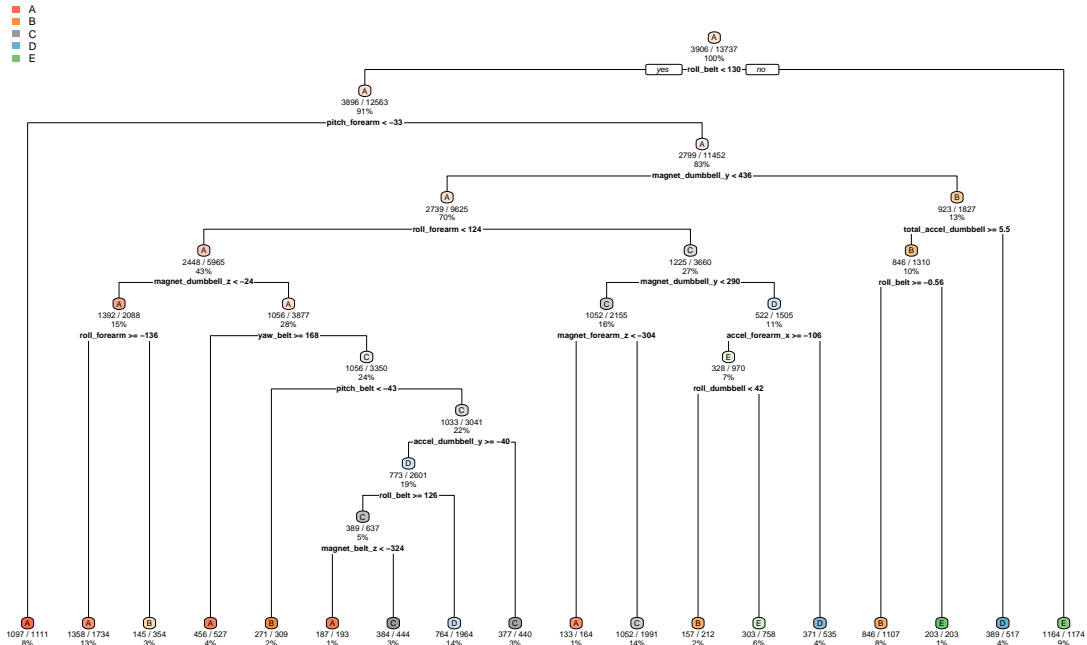
Using Precision Trees

```
modelDT <- rpart(classe ~ ., data=TrainTrainingSet, method="class")
```

```
predictionDT <- predict(modelDT, TestTrainingSet, type = "class")
```

```
rpart.plot(modelDT, main="Classification Tree", extra=102, under=TRUE, facLen=0)
```

Classification Tree



```
confusionMatrix(predictionDT, TestTrainingSet$classe)
```

Confusion Matrix and Statistics

```
##
##           Reference
## Prediction   A    B    C    D    E
##           A 1364  169   24   48   16
##           B   60  581   46   79   74
##           C   52  137  765  129  145
##           D  183  194  125  650  159
##           E   15   58   66   58  688
```

Overall Statistics

```
##           Accuracy : 0.6879
##           95% CI : (0.6758, 0.6997)
##           No Information Rate : 0.2845
##           P-Value [Acc > NIR] : < 2.2e-16
```

```
##           Kappa : 0.6066
##           McNemar's Test P-Value : < 2.2e-16
```

Statistics by Class:

```
##           Class: A Class: B Class: C Class: D Class: E
## Sensitivity      0.8148 0.51010 0.7456 0.6743 0.6359
## Specificity      0.9390 0.94543 0.9047 0.8657 0.9590
## Pos Pred Value   0.8415 0.69167 0.6230 0.4958 0.7774
## Neg Pred Value   0.9273 0.88940 0.9440 0.9314 0.9212
## Prevalence       0.2845 0.19354 0.1743 0.1638 0.1839
## Detection Rate   0.2318 0.09873 0.1300 0.1105 0.1169
```

```
## Detection Prevalence    0.2754  0.14274   0.2087   0.2228   0.1504
## Balanced Accuracy      0.8769  0.72776   0.8252   0.7700   0.7974
```

Using Random Forest

```
modelRF <- randomForest(classe ~. , data=TrainTrainingSet, method="class")

predictionRF <- predict(modelRF, TestTrainingSet, type = "class")
confusionMatrix(predictionRF, TestTrainingSet$class)
```

```
## Confusion Matrix and Statistics
```

```
##
##              Reference
## Prediction    A    B    C    D    E
##      A 1674     5     0     0     0
##      B     0 1133     6     0     0
##      C     0     1 1020     4     0
##      D     0     0     0  959     1
##      E     0     0     0     1 1081
##
```

```
## Overall Statistics
```

```
##
##              Accuracy : 0.9969
##              95% CI : (0.9952, 0.9982)
##      No Information Rate : 0.2845
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.9961
##      McNemar's Test P-Value : NA
##
```

```
## Statistics by Class:
```

```
##
##              Class: A Class: B Class: C Class: D Class: E
## Sensitivity          1.0000   0.9947   0.9942   0.9948   0.9991
## Specificity          0.9988   0.9987   0.9990   0.9998   0.9998
## Pos Pred Value       0.9970   0.9947   0.9951   0.9990   0.9991
## Neg Pred Value       1.0000   0.9987   0.9988   0.9990   0.9998
## Prevalence           0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate       0.2845   0.1925   0.1733   0.1630   0.1837
## Detection Prevalence 0.2853   0.1935   0.1742   0.1631   0.1839
## Balanced Accuracy    0.9994   0.9967   0.9966   0.9973   0.9994
```

Conclusion

Random forest performed better than decision trees with 0.9913 vs 0.6882 test training set accuracy

```
predictfinal <- predict(modelRF, testingset, type="class")
predictfinal
```

```
##  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
```