

Practical Machine Learning Prediction Project

Savio Sebastian

February 22, 2015

- 1 Background
- 2 Data
- 3 Set Up - Environment
 - 3.1 Required Packages
- 4 Getting and Cleaning Data
 - 4.1 Downloading the Data File
 - 4.2 Loading the Data
 - 4.3 Studying The Testing Data
- 5 Building the Prediction Model
 - 5.1 Random Forest
 - 5.2 SVM (Radial Kernel)
 - 5.3 KNN
 - 5.4 Investigate the Cross Validation Performance Accuracy
- 6 Prediction

1 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, the 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> (<http://groupware.les.inf.puc-rio.br/har>) (see the section on the Weight Lifting Exercise Dataset).

2 Data

The training data for this project are available here:

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

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

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

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

3 Set Up - Environment

A few packages are required to be used for this project.

The method `checkInstallPackages` is used to install / download packages as required.

```

checkInstallPackages <- function( packageName ) {
  if ( !require( packageName , character.only = T ) ) {
    message( paste( Sys.time() , "Install Package:" , packageName ) )
    install.packages(
      paste0( "\"" , packageName , "\"" ) ,      ## to pass package names as variable
s
      repos='http://cran.us.r-project.org' ,
      verbose = F , quiet = T )
    library(
      packageName ,
      character.only = T )      ## to pass package names as variables to library()
    if ( require( packageName , character.only = T ) ) {
      message( paste( Sys.time() , packageName , "Installed and loaded" ) )
    }
  }
}

```

3.1 Required Packages

```
checkInstallPackages( "caret" )
```

```

## Loading required package: caret
## Loading required package: lattice
## Loading required package: ggplot2

```

```
checkInstallPackages( "knitr" )
```

```
## Loading required package: knitr
```

```
checkInstallPackages( "randomForest" )
```

```

## Loading required package: randomForest
## randomForest 4.6-10
## Type rfNews() to see new features/changes/bug fixes.

```

```
checkInstallPackages( "doMC" )
```

```

## Loading required package: doMC
## Loading required package: foreach
## Loading required package: iterators
## Loading required package: parallel

```

```
checkInstallPackages( "e1071" )
```

```
## Loading required package: e1071
```

```
opts_chunk$set(cache=TRUE)          ## setting cache = true across the RMD
registerDoMC(cores = 4)
set.seed(140819)
```

4 Getting and Cleaning Data

4.1 Downloading the Data File

As mentioned before, a training and testing CSV File need to be downloaded from the links mentioned before.

The method, `downloadFile` accepts url and destination file name which can be reused for both the downloads

```
downloadFile <- function( urlString , fileName ) {
  download.file(
    url = urlString ,
    destfile = fileName ,
    method = "curl" , quiet = T )
}

downloadFile(
  "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv" ,
  "pml-training.csv" )
downloadFile(
  "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv" ,
  "pml-testing.csv" )
```

4.2 Loading the Data

```
dsTraining <- read.csv( "pml-training.csv" )
dsTesting  <- read.csv( "pml-testing.csv" )
```

4.3 Studying The Testing Data

```
dsSummary <- function( dataSet ) {
  print( paste(
    "No of Observations:" , dim( dataSet )[1] ,
    "      No of Variables:" , dim( dataSet )[2] ) )
  str( dataSet )
  summary( dataSet )
}

dsSummary( dsTesting )
```

```

## [1] "No of Observations: 20      No of Variables: 160"
## '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 132248963
  5 1322673149 1322673128 1322673076 1323084240 1322837822 ...
## $ raw_timestamp_part_2   : int  868349 778725 342967 560311 814776 510661 766645 5467
  1 916313 384285 ...
## $ 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 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 NA ...
## $ avg_roll_arm      : logi  NA NA NA NA NA NA NA ...
## $ stddev_roll_arm   : logi  NA NA NA NA NA NA NA ...
## $ var_roll_arm      : logi  NA NA NA NA NA NA NA ...
## $ avg_pitch_arm     : logi  NA NA NA NA NA NA NA ...
## $ stddev_pitch_arm  : logi  NA NA NA NA NA NA NA ...
## $ var_pitch_arm     : logi  NA NA NA NA NA NA NA ...
## $ avg_yaw_arm       : logi  NA NA NA NA NA NA NA ...
## $ stddev_yaw_arm    : logi  NA NA NA NA NA NA NA ...
## $ var_yaw_arm       : logi  NA 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 NA ...
## $ kurtosis_pitch_arm : logi  NA NA NA NA NA NA NA ...
## $ kurtosis_yaw_arm  : logi  NA NA NA NA NA NA NA ...
## $ skewness_roll_arm : logi  NA NA NA NA NA NA NA ...
## $ skewness_pitch_arm : logi  NA NA NA NA NA NA NA ...
## $ skewness_yaw_arm  : logi  NA NA NA NA NA NA NA ...
## $ max_roll_arm      : logi  NA NA NA NA NA NA NA ...
## $ max_pitch_arm     : logi  NA NA NA NA NA NA NA ...
## $ max_yaw_arm       : logi  NA NA NA NA NA NA NA ...
## $ min_roll_arm      : logi  NA NA NA NA NA NA NA ...
## $ min_pitch_arm     : logi  NA NA NA NA NA NA NA ...
## $ min_yaw_arm       : logi  NA NA NA NA NA NA NA ...
## $ amplitude_roll_arm : logi  NA NA NA NA NA NA NA ...
## $ amplitude_pitch_arm : logi  NA NA NA NA NA NA NA ...
## $ amplitude_yaw_arm  : logi  NA 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 NA ...
## $ kurtosis_pitch_dumbbell : logi  NA 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_picth_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]
```

```
##      X      user_name raw_timestamp_part_1 raw_timestamp_part_2
## Min.   : 1.00   adelmo   :1   Min.    :1.322e+09   Min.    : 36553
## 1st Qu.: 5.75   carlitos:3   1st Qu.:1.323e+09   1st Qu.:268655
## Median :10.50   charles  :1   Median :1.323e+09   Median :530706
## Mean   :10.50   eurico   :4   Mean   :1.323e+09   Mean   :512167
## 3rd Qu.:15.25   jeremy   :8   3rd Qu.:1.323e+09   3rd Qu.:787738
## Max.   :20.00   pedro    :3   Max.   :1.323e+09   Max.   :920315
##
##      cvtd_timestamp new_window  num_window  roll_belt
## 30/11/2011 17:11:4      no:20    Min.    : 48.0   Min.    : -5.9200
## 05/12/2011 11:24:3              1st Qu.:250.0   1st Qu.:  0.9075
## 30/11/2011 17:12:3              Median :384.5   Median :  1.1100
## 05/12/2011 14:23:2              Mean   :379.6   Mean    : 31.3055
## 28/11/2011 14:14:2              3rd Qu.:467.0   3rd Qu.: 32.5050
## 02/12/2011 13:33:1              Max.    :859.0   Max.    :129.0000
## (Other)           :5
##      pitch_belt      yaw_belt      total_accel_belt kurtosis_roll_belt
## Min.    : -41.600   Min.    : -93.70   Min.    : 2.00    Mode:logical
## 1st Qu.:  3.013   1st Qu.: -88.62   1st Qu.: 3.00    NA's:20
## Median :  4.655   Median : -87.85   Median : 4.00
## Mean    :  5.824   Mean    : -59.30   Mean    : 7.55
## 3rd Qu.:  6.135   3rd Qu.: -63.50   3rd Qu.: 8.00
## Max.    : 27.800   Max.    :162.00   Max.    :21.00
##
##      kurtosis_picth_belt kurtosis_yaw_belt skewness_roll_belt
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
##      skewness_roll_belt.1 skewness_yaw_belt max_roll_belt  max_picth_belt
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
```

```

##
##
##
## max_yaw_belt    min_roll_belt    min_pitch_belt min_yaw_belt
## Mode:logical    Mode:logical    Mode:logical    Mode:logical
## NA's:20         NA's:20         NA's:20         NA's:20
##
##
##
##
##
## amplitude_roll_belt amplitude_pitch_belt amplitude_yaw_belt
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
##
## var_total_accel_belt avg_roll_belt    stddev_roll_belt var_roll_belt
## Mode:logical      Mode:logical    Mode:logical      Mode:logical
## NA's:20           NA's:20         NA's:20           NA's:20
##
##
##
##
##
## avg_pitch_belt stddev_pitch_belt var_pitch_belt avg_yaw_belt
## Mode:logical    Mode:logical      Mode:logical      Mode:logical
## NA's:20         NA's:20         NA's:20           NA's:20
##
##
##
##
##
## stddev_yaw_belt var_yaw_belt    gyros_belt_x    gyros_belt_y
## Mode:logical    Mode:logical    Min.    :-0.500  Min.    :-0.050
## NA's:20         NA's:20         1st Qu.: -0.070  1st Qu.: -0.005
##                                     Median : 0.020    Median : 0.000
##                                     Mean    :-0.045    Mean    : 0.010
##                                     3rd Qu.: 0.070    3rd Qu.: 0.020
##                                     Max.    : 0.240    Max.    : 0.110
##
##
## gyros_belt_z    accel_belt_x    accel_belt_y    accel_belt_z
## Min.    :-0.4800  Min.    :-48.00  Min.    :-16.00  Min.    :-187.00
## 1st Qu.: -0.1375  1st Qu.: -19.00  1st Qu.:  2.00   1st Qu.: -24.00
## Median :-0.0250   Median :-13.00   Median :  4.50   Median :  27.00
## Mean   :-0.1005   Mean    :-13.50   Mean    : 18.35   Mean    : -17.60
## 3rd Qu.: 0.0000   3rd Qu.: -8.75   3rd Qu.: 25.50   3rd Qu.:  38.25
## Max.    : 0.0500   Max.     : 46.00   Max.     : 72.00   Max.     : 49.00

```

```

##
## magnet_belt_x      magnet_belt_y      magnet_belt_z      roll_arm
## Min.      :-13.00    Min.      :566.0    Min.      :-426.0    Min.      :-137.00
## 1st Qu.:   5.50     1st Qu.:578.5    1st Qu.: -398.5    1st Qu.:   0.00
## Median :  33.50     Median :600.5    Median : -313.5    Median :   0.00
## Mean   :  35.15     Mean   :601.5    Mean   : -346.9    Mean    :  16.42
## 3rd Qu.:  46.25     3rd Qu.:631.2    3rd Qu.: -305.0    3rd Qu.:  71.53
## Max.    : 169.00     Max.    :638.0    Max.    : -291.0    Max.    : 152.00
##
##      pitch_arm      yaw_arm      total_accel_arm var_accel_arm
## Min.      :-63.800    Min.      :-167.00    Min.      : 3.00    Mode:logical
## 1st Qu.:  -9.188     1st Qu.: -60.15     1st Qu.:20.25     NA's:20
## Median :   0.000     Median :   0.00     Median :29.50
## Mean   :  -3.950     Mean   :  -2.80     Mean   :26.40
## 3rd Qu.:   3.465     3rd Qu.:  25.50     3rd Qu.:33.25
## Max.    :  55.000     Max.    : 178.00     Max.    :44.00
##
## avg_roll_arm      stddev_roll_arm var_roll_arm      avg_pitch_arm
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
##
##
##      stddev_pitch_arm var_pitch_arm      avg_yaw_arm      stddev_yaw_arm
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
##
##
##      var_yaw_arm      gyros_arm_x      gyros_arm_y      gyros_arm_z
## Mode:logical      Min.      :-3.710    Min.      :-2.0900    Min.      :-0.6900
## NA's:20           1st Qu.: -0.645    1st Qu.: -0.6350    1st Qu.: -0.1800
##                   Median :  0.020    Median : -0.0400    Median : -0.0250
##                   Mean   :  0.077    Mean   : -0.1595    Mean   :  0.1205
##                   3rd Qu.:  1.248    3rd Qu.:  0.2175    3rd Qu.:  0.5650
##                   Max.    :  3.660    Max.    :  1.8500    Max.    :  1.1300
##
##
##      accel_arm_x      accel_arm_y      accel_arm_z      magnet_arm_x
## Min.      :-341.0    Min.      :-65.00    Min.      :-404.00    Min.      :-428.00
## 1st Qu.: -277.0    1st Qu.:  52.25     1st Qu.: -128.50     1st Qu.: -373.75
## Median : -194.5    Median :112.00     Median :  -83.50     Median : -265.00
## Mean   : -134.6    Mean   :103.10     Mean   :  -87.85     Mean   :  -38.95
## 3rd Qu.:   5.5     3rd Qu.:168.25     3rd Qu.:  -27.25     3rd Qu.: 250.50
## Max.    : 106.0    Max.    :245.00     Max.    :   93.00     Max.    : 750.00
##
##
##      magnet_arm_y      magnet_arm_z      kurtosis_roll_arm kurtosis_picth_arm

```



```

## Min.      :-307.0    Min.      :-499.0    Mode:logical    Mode:logical
## 1st Qu.: 205.2    1st Qu.: 403.0    NA's:20        NA's:20
## Median : 291.0    Median : 476.5
## Mean    : 239.4    Mean    : 369.8
## 3rd Qu.: 358.8    3rd Qu.: 517.0
## Max.    : 474.0    Max.    : 633.0
##
## kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm skewness_yaw_arm
## Mode:logical    Mode:logical    Mode:logical    Mode:logical
## NA's:20        NA's:20        NA's:20        NA's:20
##
##
##
##
##
## max_roll_arm    max_picth_arm    max_yaw_arm    min_roll_arm
## Mode:logical    Mode:logical    Mode:logical    Mode:logical
## NA's:20        NA's:20        NA's:20        NA's:20
##
##
##
##
##
## min_pitch_arm    min_yaw_arm    amplitude_roll_arm amplitude_pitch_arm
## Mode:logical    Mode:logical    Mode:logical    Mode:logical
## NA's:20        NA's:20        NA's:20        NA's:20
##
##
##
##
##
## amplitude_yaw_arm roll_dumbbell    pitch_dumbbell    yaw_dumbbell
## Mode:logical    Min.      :-111.118    Min.      :-54.97    Min.      :-103.3200
## NA's:20        1st Qu.:   7.494    1st Qu.: -51.89    1st Qu.: -75.2809
##                Median : 50.403    Median : -40.81    Median :  -8.2863
##                Mean    : 33.760    Mean    : -19.47    Mean    :  -0.9385
##                3rd Qu.: 58.129    3rd Qu.: 16.12    3rd Qu.: 55.8335
##                Max.    : 123.984    Max.    : 96.87    Max.    : 132.2337
##
## kurtosis_roll_dumbbell kurtosis_picth_dumbbell kurtosis_yaw_dumbbell
## Mode:logical    Mode:logical    Mode:logical
## NA's:20        NA's:20        NA's:20
##
##
##
##
##
## skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
## Mode:logical    Mode:logical    Mode:logical
## NA's:20        NA's:20        NA's:20

```

```
##
##
##
##
##
## max_roll_dumbbell max_picth_dumbbell max_yaw_dumbbell min_roll_dumbbell
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
##
##
##
## min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
##
## amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
## Mode:logical      Mode:logical      Min.    : 1.0
## NA's:20           NA's:20           1st Qu.: 7.0
##                                     Median :15.5
##                                     Mean   :17.2
##                                     3rd Qu.:29.0
##                                     Max.   :31.0
##
##
## var_accel_dumbbell avg_roll_dumbbell stddev_roll_dumbbell
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
##
## var_roll_dumbbell avg_pitch_dumbbell stddev_pitch_dumbbell
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
##
##
## var_pitch_dumbbell avg_yaw_dumbbell stddev_yaw_dumbbell var_yaw_dumbbell
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
```

```

##
##
##
## gyros_dumbbell_x gyros_dumbbell_y gyros_dumbbell_z accel_dumbbell_x
## Min.      :-1.0300 Min.      :-1.1100 Min.      :-1.180 Min.      :-159.00
## 1st Qu.: 0.1600 1st Qu.: -0.2100 1st Qu.: -0.485 1st Qu.: -140.25
## Median : 0.3600 Median : 0.0150 Median : -0.280 Median : -19.00
## Mean    : 0.2690 Mean    : 0.0605 Mean    : -0.266 Mean    : -47.60
## 3rd Qu.: 0.4625 3rd Qu.: 0.1450 3rd Qu.: -0.165 3rd Qu.: 15.75
## Max.    : 1.0600 Max.    : 1.9100 Max.    : 1.100 Max.    : 185.00
##
## accel_dumbbell_y accel_dumbbell_z magnet_dumbbell_x magnet_dumbbell_y
## Min.      :-30.00 Min.      :-221.0 Min.      :-576.0 Min.      :-558.0
## 1st Qu.: 5.75 1st Qu.: -192.2 1st Qu.: -528.0 1st Qu.: 259.5
## Median : 71.50 Median : -3.0 Median : -508.5 Median : 316.0
## Mean    : 70.55 Mean    : -60.0 Mean    : -304.2 Mean    : 189.3
## 3rd Qu.:151.25 3rd Qu.: 76.5 3rd Qu.: -317.0 3rd Qu.: 348.2
## Max.    :166.00 Max.    : 100.0 Max.    : 523.0 Max.    : 403.0
##
## magnet_dumbbell_z roll_forearm pitch_forearm yaw_forearm
## Min.      :-164.00 Min.      :-176.00 Min.      :-63.500 Min.      :-168.000
## 1st Qu.: -33.00 1st Qu.: -40.25 1st Qu.: -11.457 1st Qu.: -93.375
## Median : 49.50 Median : 94.20 Median : 8.830 Median : -19.250
## Mean    : 71.40 Mean    : 38.66 Mean    : 7.099 Mean    : 2.195
## 3rd Qu.: 96.25 3rd Qu.: 143.25 3rd Qu.: 28.500 3rd Qu.: 104.500
## Max.    : 368.00 Max.    : 176.00 Max.    : 59.300 Max.    : 159.000
##
## kurtosis_roll_forearm kurtosis_pitch_forearm kurtosis_yaw_forearm
## Mode:logical Mode:logical Mode:logical
## NA's:20 NA's:20 NA's:20
##
##
##
##
## skewness_roll_forearm skewness_pitch_forearm skewness_yaw_forearm
## Mode:logical Mode:logical Mode:logical
## NA's:20 NA's:20 NA's:20
##
##
##
##
##
## max_roll_forearm max_pitch_forearm max_yaw_forearm min_roll_forearm
## Mode:logical Mode:logical Mode:logical Mode:logical
## NA's:20 NA's:20 NA's:20 NA's:20
##
##
##
##

```

```
## min_pitch_forearm min_yaw_forearm amplitude_roll_forearm
## Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20
##
##
##
##
## amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
## Mode:logical      Mode:logical      Min.      :21.00
## NA's:20           NA's:20           1st Qu.:24.00
##                                     Median   :32.50
##                                     Mean      :32.05
##                                     3rd Qu.:36.75
##                                     Max.      :47.00
##
## var_accel_forearm avg_roll_forearm stddev_roll_forearm var_roll_forearm
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
##
##
## avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## Mode:logical      Mode:logical      Mode:logical      Mode:logical
## NA's:20           NA's:20           NA's:20           NA's:20
##
##
##
##
##
## stddev_yaw_forearm var_yaw_forearm gyros_forearm_x  gyros_forearm_y
## Mode:logical      Mode:logical      Min.      : -1.0600  Min.      : -5.9700
## NA's:20           NA's:20           1st Qu.: -0.5850  1st Qu.: -1.2875
##                                     Median   : 0.0200  Median   : 0.0350
##                                     Mean      : -0.0200  Mean      : -0.0415
##                                     3rd Qu.: 0.2925  3rd Qu.: 2.0475
##                                     Max.      : 1.3800  Max.      : 4.2600
##
## gyros_forearm_z accel_forearm_x accel_forearm_y accel_forearm_z
## Min.      : -1.2600  Min.      : -212.0  Min.      : -331.0  Min.      : -282.0
## 1st Qu.: -0.0975  1st Qu.: -114.8  1st Qu.:    8.5  1st Qu.: -199.0
## Median   : 0.2300  Median   :  86.0  Median   : 138.0  Median   : -148.5
## Mean      : 0.2610  Mean      :  38.8  Mean      : 125.3  Mean      : -93.7
## 3rd Qu.: 0.7625  3rd Qu.: 166.2  3rd Qu.: 268.0  3rd Qu.: -31.0
## Max.      : 1.8000  Max.      : 232.0  Max.      : 406.0  Max.      : 179.0
##
## magnet_forearm_x magnet_forearm_y magnet_forearm_z  problem_id
```

```
## Min.      :-714.0    Min.      :-787.0    Min.      :-32.0    Min.      : 1.00
## 1st Qu.: -427.2    1st Qu.: -328.8    1st Qu.: 275.2    1st Qu.: 5.75
## Median : -189.5    Median : 487.0    Median : 491.5    Median : 10.50
## Mean    : -159.2    Mean     : 191.8    Mean     : 460.2    Mean     : 10.50
## 3rd Qu.:  41.5     3rd Qu.: 720.8    3rd Qu.: 661.5    3rd Qu.: 15.25
## Max.     : 532.0    Max.      : 800.0    Max.      : 884.0    Max.      : 20.00
##
```

It is evident from the data that certain columns are not important for Testing & Training Data Sets since they only contain `NA` values. These columns are removed.

The function `filterData` will do this.

```
## Remove NA Columns
## Function to filter the features
filterData <- function(dataSet) {
  # Since we have lots of variables, remove any with NA's
  # or have empty strings
  tmpDataSet.keep <- !sapply(dataSet, function(x) any(is.na(x)))
  dataSet <- dataSet[, tmpDataSet.keep]
  tmpDataSet.keep <- !sapply(dataSet, function(x) any(x==""))
  dataSet <- dataSet[, tmpDataSet.keep]

  # Remove the columns that aren't the predictor variables
  col.rm <- c("X", "user_name", "raw_timestamp_part_1", "raw_timestamp_part_2",
             "cvtd_timestamp", "new_window", "num_window")
  tmpDataSet.rm <- which(colnames(dataSet) %in% col.rm)
  dataSet <- dataSet[, -tmpDataSet.rm]

  return(dataSet)
}
```

```
dsTrainingCleaned <- filterData(dsTraining)
dsTestingCleaned <- filterData(dsTesting)

## setting factor on classe
dsTrainingCleaned$classe <- factor( dsTrainingCleaned$classe )

dsSummary( dsTrainingCleaned )
```

```
## [1] "No of Observations: 19622      No of Variables: 53"
## 'data.frame':    19622 obs. of  53 variables:
## $ 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 ...
## $ gyros_belt_x    : num   0 0.02 0 0.02 0.02 0.02 0.02 0.02 0.02 0.03 ...
## $ gyros_belt_y    : num   0 0 0 0 0.02 0 0 0 0 0 ...
## $ gyros_belt_z    : num  -0.02 -0.02 -0.02 -0.03 -0.02 -0.02 -0.02 -0.02 -0.02 0 .
```

```

..
## $ accel_belt_x      : int  -21 -22 -20 -22 -21 -21 -22 -22 -20 -21 ...
## $ accel_belt_y      : int   4 4 5 3 2 4 3 4 2 4 ...
## $ accel_belt_z      : int  22 22 23 21 24 21 21 21 24 22 ...
## $ magnet_belt_x     : int   -3 -7 -2 -6 -6 0 -4 -2 1 -3 ...
## $ magnet_belt_y     : int  599 608 600 604 600 603 599 603 602 609 ...
## $ magnet_belt_z     : int -313 -311 -305 -310 -302 -312 -311 -313 -312 -308 ...
## $ roll_arm          : num -128 -128 -128 -128 -128 -128 -128 -128 -128 -128 ...
## $ pitch_arm         : num  22.5 22.5 22.5 22.1 22.1 22 21.9 21.8 21.7 21.6 ...
## $ yaw_arm           : num -161 -161 -161 -161 -161 -161 -161 -161 -161 -161 ...
## $ total_accel_arm    : int   34 34 34 34 34 34 34 34 34 34 ...
## $ gyros_arm_x        : num   0 0.02 0.02 0.02 0 0.02 0 0.02 0.02 0.02 ...
## $ gyros_arm_y        : num   0 -0.02 -0.02 -0.03 -0.03 -0.03 -0.03 -0.02 -0.03 -0.03 .
..
## $ gyros_arm_z        : num -0.02 -0.02 -0.02 0.02 0 0 0 0 -0.02 -0.02 ...
## $ accel_arm_x        : int -288 -290 -289 -289 -289 -289 -289 -289 -288 -288 ...
## $ accel_arm_y        : int  109 110 110 111 111 111 111 111 109 110 ...
## $ accel_arm_z        : int -123 -125 -126 -123 -123 -122 -125 -124 -122 -124 ...
## $ magnet_arm_x       : int -368 -369 -368 -372 -374 -369 -373 -372 -369 -376 ...
## $ magnet_arm_y       : int  337 337 344 344 337 342 336 338 341 334 ...
## $ magnet_arm_z       : int  516 513 513 512 506 513 509 510 518 516 ...
## $ 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 ...
## $ total_accel_dumbbell : int  37 37 37 37 37 37 37 37 37 37 ...
## $ gyros_dumbbell_x    : num   0 0 0 0 0 0 0 0 0 0 ...
## $ gyros_dumbbell_y    : num -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.
02 ...
## $ gyros_dumbbell_z    : num   0 0 0 -0.02 0 0 0 0 0 0 ...
## $ accel_dumbbell_x    : int -234 -233 -232 -232 -233 -234 -232 -234 -232 -235 ...
## $ accel_dumbbell_y    : int  47 47 46 48 48 48 47 46 47 48 ...
## $ accel_dumbbell_z    : int -271 -269 -270 -269 -270 -269 -270 -272 -269 -270 ...
## $ magnet_dumbbell_x   : int -559 -555 -561 -552 -554 -558 -551 -555 -549 -558 ...
## $ magnet_dumbbell_y   : int  293 296 298 303 292 294 295 300 292 291 ...
## $ magnet_dumbbell_z   : num  -65 -64 -63 -60 -68 -66 -70 -74 -65 -69 ...
## $ roll_forearm        : num  28.4 28.3 28.3 28.1 28 27.9 27.9 27.8 27.7 27.7 ...
## $ pitch_forearm       : num -63.9 -63.9 -63.9 -63.9 -63.9 -63.9 -63.9 -63.8 -63.8 -63
.8 ...
## $ yaw_forearm         : num -153 -153 -152 -152 -152 -152 -152 -152 -152 -152 ...
## $ total_accel_forearm : int   36 36 36 36 36 36 36 36 36 36 ...
## $ gyros_forearm_x     : num  0.03 0.02 0.03 0.02 0.02 0.02 0.02 0.02 0.02 0.03 0.02 ...
## $ gyros_forearm_y     : num   0 0 -0.02 -0.02 0 -0.02 0 -0.02 0 0 ...
## $ gyros_forearm_z     : num -0.02 -0.02 0 0 -0.02 -0.03 -0.02 0 -0.02 -0.02 ...
## $ accel_forearm_x     : int  192 192 196 189 189 193 195 193 193 190 ...
## $ accel_forearm_y     : int  203 203 204 206 206 203 205 205 204 205 ...
## $ accel_forearm_z     : int -215 -216 -213 -214 -214 -215 -215 -213 -214 -215 ...
## $ magnet_forearm_x    : int  -17 -18 -18 -16 -17 -9 -18 -9 -16 -22 ...
## $ magnet_forearm_y    : num  654 661 658 658 655 660 659 660 653 656 ...
## $ magnet_forearm_z    : num  476 473 469 469 473 478 470 474 476 473 ...
## $ classe              : Factor w/ 5 levels "A","B","C","D",...: 1 1 1 1 1 1 1 1 1 1 ...

```

```

##      roll_belt      pitch_belt      yaw_belt      total_accel_belt
## Min.      :-28.90    Min.      :-55.8000    Min.      :-180.00    Min.      : 0.00
## 1st Qu.:   1.10    1st Qu.:   1.7600    1st Qu.:  -88.30    1st Qu.:  3.00
## Median : 113.00    Median :   5.2800    Median :  -13.00    Median : 17.00
## Mean      : 64.41    Mean      :  0.3053    Mean      : -11.21    Mean      : 11.31
## 3rd Qu.: 123.00    3rd Qu.:  14.9000    3rd Qu.:   12.90    3rd Qu.: 18.00
## Max.      :162.00    Max.      : 60.3000    Max.      : 179.00    Max.      :29.00
##      gyros_belt_x      gyros_belt_y      gyros_belt_z
## Min.      :-1.040000    Min.      :-0.64000    Min.      :-1.4600
## 1st Qu.: -0.030000    1st Qu.:  0.00000    1st Qu.: -0.2000
## Median :  0.030000    Median :  0.02000    Median : -0.1000
## Mean      :-0.005592    Mean      : 0.03959    Mean      :-0.1305
## 3rd Qu.:  0.110000    3rd Qu.:  0.11000    3rd Qu.: -0.0200
## Max.      :  2.220000    Max.      :  0.64000    Max.      :  1.6200
##      accel_belt_x      accel_belt_y      accel_belt_z      magnet_belt_x
## Min.      :-120.000    Min.      :-69.00    Min.      :-275.00    Min.      :-52.0
## 1st Qu.:  -21.000    1st Qu.:   3.00    1st Qu.: -162.00    1st Qu.:   9.0
## Median :  -15.000    Median :  35.00    Median : -152.00    Median :  35.0
## Mean      :  -5.595    Mean      : 30.15    Mean      : -72.59    Mean      : 55.6
## 3rd Qu.:  -5.000    3rd Qu.:  61.00    3rd Qu.:   27.00    3rd Qu.:  59.0
## Max.      :   85.000    Max.      :164.00    Max.      : 105.00    Max.      :485.0
##      magnet_belt_y      magnet_belt_z      roll_arm      pitch_arm
## Min.      :354.0    Min.      :-623.0    Min.      :-180.00    Min.      :-88.800
## 1st Qu.:581.0    1st Qu.: -375.0    1st Qu.:  -31.77    1st Qu.: -25.900
## Median :601.0    Median : -320.0    Median :   0.00    Median :   0.000
## Mean      :593.7    Mean      :-345.5    Mean      :  17.83    Mean      : -4.612
## 3rd Qu.:610.0    3rd Qu.: -306.0    3rd Qu.:   77.30    3rd Qu.:  11.200
## Max.      :673.0    Max.      : 293.0    Max.      : 180.00    Max.      : 88.500
##      yaw_arm      total_accel_arm      gyros_arm_x      gyros_arm_y
## Min.      :-180.0000    Min.      :  1.00    Min.      :-6.37000    Min.      :-3.4400
## 1st Qu.:  -43.1000    1st Qu.:17.00    1st Qu.: -1.33000    1st Qu.: -0.8000
## Median :   0.0000    Median :27.00    Median :  0.08000    Median : -0.2400
## Mean      :  -0.6188    Mean      :25.51    Mean      :  0.04277    Mean      :-0.2571
## 3rd Qu.:   45.8750    3rd Qu.:33.00    3rd Qu.:  1.57000    3rd Qu.:  0.1400
## Max.      : 180.0000    Max.      :66.00    Max.      : 4.87000    Max.      : 2.8400
##      gyros_arm_z      accel_arm_x      accel_arm_y      accel_arm_z
## Min.      :-2.3300    Min.      :-404.00    Min.      :-318.0    Min.      :-636.00
## 1st Qu.: -0.0700    1st Qu.: -242.00    1st Qu.:  -54.0    1st Qu.: -143.00
## Median :  0.2300    Median :  -44.00    Median :   14.0    Median :  -47.00
## Mean      :  0.2695    Mean      : -60.24    Mean      :   32.6    Mean      : -71.25
## 3rd Qu.:  0.7200    3rd Qu.:   84.00    3rd Qu.:  139.0    3rd Qu.:   23.00
## Max.      :  3.0200    Max.      : 437.00    Max.      : 308.0    Max.      : 292.00
##      magnet_arm_x      magnet_arm_y      magnet_arm_z      roll_dumbbell
## Min.      :-584.0    Min.      :-392.0    Min.      :-597.0    Min.      :-153.71
## 1st Qu.: -300.0    1st Qu.:  -9.0    1st Qu.: 131.2    1st Qu.: -18.49
## Median : 289.0    Median : 202.0    Median : 444.0    Median :  48.17
## Mean      : 191.7    Mean      : 156.6    Mean      : 306.5    Mean      :  23.84
## 3rd Qu.: 637.0    3rd Qu.: 323.0    3rd Qu.: 545.0    3rd Qu.:  67.61
## Max.      : 782.0    Max.      : 583.0    Max.      : 694.0    Max.      : 153.55
##      pitch_dumbbell      yaw_dumbbell      total_accel_dumbbell

```

```

## Min.      :-149.59   Min.      :-150.871   Min.      : 0.00
## 1st Qu.: -40.89   1st Qu.: -77.644   1st Qu.: 4.00
## Median : -20.96   Median :  -3.324   Median :10.00
## Mean    : -10.78   Mean     :  1.674   Mean     :13.72
## 3rd Qu.:  17.50   3rd Qu.:  79.643   3rd Qu.:19.00
## Max.     : 149.40   Max.      : 154.952   Max.      :58.00
## gyros_dumbbell_x   gyros_dumbbell_y   gyros_dumbbell_z
## Min.      :-204.0000   Min.      :-2.10000   Min.      : -2.380
## 1st Qu.:  -0.0300   1st Qu.: -0.14000   1st Qu.: -0.310
## Median :    0.1300   Median :  0.03000   Median : -0.130
## Mean     :    0.1611   Mean      : 0.04606   Mean      : -0.129
## 3rd Qu.:    0.3500   3rd Qu.:  0.21000   3rd Qu.:  0.030
## Max.      :    2.2200   Max.       :52.00000   Max.       :317.000
## accel_dumbbell_x   accel_dumbbell_y   accel_dumbbell_z   magnet_dumbbell_x
## Min.      :-419.00   Min.      :-189.00   Min.      :-334.00   Min.      :-643.0
## 1st Qu.: -50.00   1st Qu.:  -8.00   1st Qu.: -142.00   1st Qu.: -535.0
## Median :  -8.00   Median :  41.50   Median :  -1.00   Median : -479.0
## Mean     : -28.62   Mean      : 52.63   Mean      : -38.32   Mean      : -328.5
## 3rd Qu.:  11.00   3rd Qu.: 111.00   3rd Qu.:  38.00   3rd Qu.: -304.0
## Max.      : 235.00   Max.      : 315.00   Max.      : 318.00   Max.      : 592.0
## magnet_dumbbell_y   magnet_dumbbell_z   roll_forearm       pitch_forearm
## Min.      :-3600    Min.      :-262.00   Min.      :-180.0000   Min.      :-72.50
## 1st Qu.:  231      1st Qu.: -45.00   1st Qu.:  -0.7375   1st Qu.:  0.00
## Median :  311      Median :  13.00   Median :  21.7000   Median :   9.24
## Mean     :  221      Mean      : 46.05   Mean      : 33.8265   Mean      : 10.71
## 3rd Qu.:  390      3rd Qu.:  95.00   3rd Qu.: 140.0000   3rd Qu.:  28.40
## Max.      :  633      Max.      : 452.00   Max.      : 180.0000   Max.      : 89.80
## yaw_forearm         total_accel_forearm gyros_forearm_x
## Min.      :-180.00   Min.      :  0.00    Min.      :-22.000
## 1st Qu.: -68.60   1st Qu.: 29.00    1st Qu.: -0.220
## Median :    0.00   Median : 36.00    Median :  0.050
## Mean     :  19.21   Mean      : 34.72    Mean      :  0.158
## 3rd Qu.: 110.00   3rd Qu.: 41.00    3rd Qu.:  0.560
## Max.      : 180.00   Max.      :108.00    Max.      :  3.970
## gyros_forearm_y     gyros_forearm_z     accel_forearm_x     accel_forearm_y
## Min.      : -7.02000   Min.      : -8.0900   Min.      : -498.00   Min.      : -632.0
## 1st Qu.: -1.46000   1st Qu.: -0.1800   1st Qu.: -178.00   1st Qu.:  57.0
## Median :  0.03000   Median :  0.0800   Median :  -57.00   Median : 201.0
## Mean     :  0.07517   Mean      :  0.1512   Mean      : -61.65   Mean      : 163.7
## 3rd Qu.:  1.62000   3rd Qu.:  0.4900   3rd Qu.:  76.00   3rd Qu.: 312.0
## Max.      :311.00000   Max.      :231.0000   Max.      : 477.00   Max.      : 923.0
## accel_forearm_z     magnet_forearm_x     magnet_forearm_y     magnet_forearm_z
## Min.      :-446.00   Min.      :-1280.0   Min.      :-896.0   Min.      :-973.0
## 1st Qu.: -182.00   1st Qu.: -616.0   1st Qu.:  2.0   1st Qu.: 191.0
## Median : -39.00   Median : -378.0   Median : 591.0   Median : 511.0
## Mean     : -55.29   Mean      : -312.6   Mean      : 380.1   Mean      : 393.6
## 3rd Qu.:  26.00   3rd Qu.:  -73.0   3rd Qu.: 737.0   3rd Qu.: 653.0
## Max.      : 291.00   Max.      :  672.0   Max.      :1480.0   Max.      :1090.0
## classe
## A:5580

```



```
## B:3797
## C:3422
## D:3216
## E:3607
##
```

```
dsSummary( dsTestingCleaned )
```

```
## [1] "No of Observations: 20      No of Variables: 53"
## 'data.frame':    20 obs. of  53 variables:
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ total_accel_dumbbell: int   9 31 29 18 4 29 29 29 3 2 ...
## $ gyros_dumbbell_x : num   0.64 0.34 0.39 0.1 0.29 -0.59 0.34 0.37 0.03 0.42 ...
## $ gyros_dumbbell_y : num   0.06 0.05 0.14 -0.02 -0.47 0.8 0.16 0.14 -0.21 0.51 ...
## $ gyros_dumbbell_z : num  -0.61 -0.71 -0.34 0.05 -0.46 1.1 -0.23 -0.39 -0.21 -0.03
## ...
## $ accel_dumbbell_x : int   21 -153 -141 -51 -18 -138 -145 -140 0 -7 ...
## $ accel_dumbbell_y : int  -15 155 155 72 -30 166 150 159 25 -20 ...
## $ accel_dumbbell_z : int   81 -205 -196 -148 -5 -186 -190 -191 9 7 ...
```

```

## $ magnet_dumbbell_x      : int  523 -502 -506 -576 -424 -543 -484 -515 -519 -531 ...
## $ magnet_dumbbell_y      : int  -528 388 349 238 252 262 354 350 348 321 ...
## $ magnet_dumbbell_z      : int  -56 -36 41 53 312 96 97 53 -32 -164 ...
## $ roll_forearm           : num  141 109 131 0 -176 150 155 -161 15.5 13.2 ...
## $ pitch_forearm          : num  49.3 -17.6 -32.6 0 -2.16 1.46 34.5 43.6 -63.5 19.4 ...
## $ yaw_forearm            : num  156 106 93 0 -47.9 89.7 152 -89.5 -139 -105 ...
## $ total_accel_forearm    : int   33 39 34 43 24 43 32 47 36 24 ...
## $ gyros_forearm_x        : num   0.74 1.12 0.18 1.38 -0.75 -0.88 -0.53 0.63 0.03 0.02 ...
## $ gyros_forearm_y        : num  -3.34 -2.78 -0.79 0.69 3.1 4.26 1.8 -0.74 0.02 0.13 ...
## $ gyros_forearm_z        : num  -0.59 -0.18 0.28 1.8 0.8 1.35 0.75 0.49 -0.02 -0.07 ...
## $ accel_forearm_x        : int  -110 212 154 -92 131 230 -192 -151 195 -212 ...
## $ accel_forearm_y        : int   267 297 271 406 -93 322 170 -331 204 98 ...
## $ accel_forearm_z        : int  -149 -118 -129 -39 172 -144 -175 -282 -217 -7 ...
## $ magnet_forearm_x       : int  -714 -237 -51 -233 375 -300 -678 -109 0 -403 ...
## $ magnet_forearm_y       : int   419 791 698 783 -787 800 284 -619 652 723 ...
## $ magnet_forearm_z       : int   617 873 783 521 91 884 585 -32 469 512 ...
## $ problem_id             : int    1 2 3 4 5 6 7 8 9 10 ...

```

```

##      roll_belt      pitch_belt      yaw_belt      total_accel_belt
## Min.   : -5.9200   Min.   : -41.600   Min.   : -93.70   Min.   : 2.00
## 1st Qu.:  0.9075   1st Qu.:  3.013   1st Qu.: -88.62   1st Qu.: 3.00
## Median :  1.1100   Median :  4.655   Median : -87.85   Median : 4.00
## Mean   : 31.3055   Mean   :  5.824   Mean   : -59.30   Mean   : 7.55
## 3rd Qu.: 32.5050   3rd Qu.:  6.135   3rd Qu.: -63.50   3rd Qu.: 8.00
## Max.   :129.0000   Max.   : 27.800   Max.   :162.00   Max.   :21.00
##      gyros_belt_x      gyros_belt_y      gyros_belt_z      accel_belt_x
## Min.   : -0.500   Min.   : -0.050   Min.   : -0.4800   Min.   : -48.00
## 1st Qu.: -0.070   1st Qu.: -0.005   1st Qu.: -0.1375   1st Qu.: -19.00
## Median :  0.020   Median :  0.000   Median : -0.0250   Median : -13.00
## Mean   : -0.045   Mean   :  0.010   Mean   : -0.1005   Mean   : -13.50
## 3rd Qu.:  0.070   3rd Qu.:  0.020   3rd Qu.:  0.0000   3rd Qu.:  -8.75
## Max.   :  0.240   Max.   :  0.110   Max.   :  0.0500   Max.   : 46.00
##      accel_belt_y      accel_belt_z      magnet_belt_x      magnet_belt_y
## Min.   : -16.00   Min.   : -187.00   Min.   : -13.00   Min.   : 566.0
## 1st Qu.:  2.00    1st Qu.: -24.00   1st Qu.:  5.50    1st Qu.: 578.5
## Median :  4.50    Median :  27.00   Median : 33.50    Median : 600.5
## Mean   : 18.35    Mean   : -17.60   Mean   : 35.15    Mean   : 601.5
## 3rd Qu.: 25.50    3rd Qu.: 38.25   3rd Qu.: 46.25    3rd Qu.: 631.2
## Max.   : 72.00    Max.   : 49.00    Max.   :169.00    Max.   : 638.0
##      magnet_belt_z      roll_arm      pitch_arm      yaw_arm
## Min.   : -426.0   Min.   : -137.00   Min.   : -63.800   Min.   : -167.00
## 1st Qu.: -398.5   1st Qu.:  0.00    1st Qu.: -9.188   1st Qu.: -60.15
## Median : -313.5   Median :  0.00    Median :  0.000   Median :  0.00
## Mean   : -346.9   Mean   : 16.42    Mean   : -3.950   Mean   : -2.80
## 3rd Qu.: -305.0   3rd Qu.: 71.53    3rd Qu.:  3.465   3rd Qu.: 25.50
## Max.   : -291.0   Max.   :152.00    Max.   :55.000   Max.   :178.00
##      total_accel_arm      gyros_arm_x      gyros_arm_y      gyros_arm_z
## Min.   : 3.00    Min.   : -3.710   Min.   : -2.0900   Min.   : -0.6900
## 1st Qu.:20.25    1st Qu.: -0.645   1st Qu.: -0.6350   1st Qu.: -0.1800
## Median :29.50    Median :  0.020   Median : -0.0400   Median : -0.0250

```

```

## Mean      :26.40      Mean      : 0.077      Mean      :-0.1595      Mean      : 0.1205
## 3rd Qu.:33.25      3rd Qu.: 1.248      3rd Qu.: 0.2175      3rd Qu.: 0.5650
## Max.       :44.00      Max.       : 3.660      Max.       : 1.8500      Max.       : 1.1300
## accel_arm_x      accel_arm_y      accel_arm_z      magnet_arm_x
## Min.       :-341.0      Min.       :-65.00      Min.       :-404.00      Min.       :-428.00
## 1st Qu.: -277.0      1st Qu.: 52.25      1st Qu.: -128.50      1st Qu.: -373.75
## Median : -194.5      Median :112.00      Median : -83.50      Median : -265.00
## Mean       :-134.6      Mean       :103.10      Mean       : -87.85      Mean       : -38.95
## 3rd Qu.:   5.5      3rd Qu.:168.25      3rd Qu.: -27.25      3rd Qu.: 250.50
## Max.       : 106.0      Max.       :245.00      Max.       : 93.00      Max.       : 750.00
## magnet_arm_y      magnet_arm_z      roll_dumbbell      pitch_dumbbell
## Min.       :-307.0      Min.       :-499.0      Min.       :-111.118      Min.       :-54.97
## 1st Qu.: 205.2      1st Qu.: 403.0      1st Qu.: 7.494      1st Qu.: -51.89
## Median : 291.0      Median : 476.5      Median : 50.403      Median : -40.81
## Mean       : 239.4      Mean       : 369.8      Mean       : 33.760      Mean       : -19.47
## 3rd Qu.: 358.8      3rd Qu.: 517.0      3rd Qu.: 58.129      3rd Qu.: 16.12
## Max.       : 474.0      Max.       : 633.0      Max.       : 123.984      Max.       : 96.87
## yaw_dumbbell      total_accel_dumbbell gyros_dumbbell_x
## Min.       :-103.3200      Min.       : 1.0      Min.       :-1.0300
## 1st Qu.: -75.2809      1st Qu.: 7.0      1st Qu.: 0.1600
## Median : -8.2863      Median :15.5      Median : 0.3600
## Mean       : -0.9385      Mean       :17.2      Mean       : 0.2690
## 3rd Qu.: 55.8335      3rd Qu.:29.0      3rd Qu.: 0.4625
## Max.       : 132.2337      Max.       :31.0      Max.       : 1.0600
## gyros_dumbbell_y      gyros_dumbbell_z      accel_dumbbell_x      accel_dumbbell_y
## Min.       :-1.1100      Min.       :-1.180      Min.       :-159.00      Min.       :-30.00
## 1st Qu.: -0.2100      1st Qu.: -0.485      1st Qu.: -140.25      1st Qu.: 5.75
## Median : 0.0150      Median : -0.280      Median : -19.00      Median : 71.50
## Mean       : 0.0605      Mean       : -0.266      Mean       : -47.60      Mean       : 70.55
## 3rd Qu.: 0.1450      3rd Qu.: -0.165      3rd Qu.: 15.75      3rd Qu.:151.25
## Max.       : 1.9100      Max.       : 1.100      Max.       : 185.00      Max.       :166.00
## accel_dumbbell_z      magnet_dumbbell_x      magnet_dumbbell_y      magnet_dumbbell_z
## Min.       :-221.0      Min.       :-576.0      Min.       :-558.0      Min.       :-164.00
## 1st Qu.: -192.2      1st Qu.: -528.0      1st Qu.: 259.5      1st Qu.: -33.00
## Median : -3.0      Median : -508.5      Median : 316.0      Median : 49.50
## Mean       : -60.0      Mean       : -304.2      Mean       : 189.3      Mean       : 71.40
## 3rd Qu.: 76.5      3rd Qu.: -317.0      3rd Qu.: 348.2      3rd Qu.: 96.25
## Max.       : 100.0      Max.       : 523.0      Max.       : 403.0      Max.       : 368.00
## roll_forearm      pitch_forearm      yaw_forearm
## Min.       :-176.00      Min.       :-63.500      Min.       :-168.000
## 1st Qu.: -40.25      1st Qu.: -11.457      1st Qu.: -93.375
## Median : 94.20      Median : 8.830      Median : -19.250
## Mean       : 38.66      Mean       : 7.099      Mean       : 2.195
## 3rd Qu.: 143.25      3rd Qu.: 28.500      3rd Qu.: 104.500
## Max.       : 176.00      Max.       : 59.300      Max.       : 159.000
## total_accel_forearm      gyros_forearm_x      gyros_forearm_y      gyros_forearm_z
## Min.       :21.00      Min.       :-1.0600      Min.       :-5.9700      Min.       :-1.2600
## 1st Qu.:24.00      1st Qu.: -0.5850      1st Qu.: -1.2875      1st Qu.: -0.0975
## Median :32.50      Median : 0.0200      Median : 0.0350      Median : 0.2300
## Mean       :32.05      Mean       : -0.0200      Mean       : -0.0415      Mean       : 0.2610

```

```
## 3rd Qu.:36.75      3rd Qu.: 0.2925    3rd Qu.: 2.0475    3rd Qu.: 0.7625
## Max.      :47.00      Max.      : 1.3800    Max.      : 4.2600    Max.      : 1.8000
## accel_forearm_x accel_forearm_y accel_forearm_z magnet_forearm_x
## Min.      : -212.0    Min.      : -331.0    Min.      : -282.0    Min.      : -714.0
## 1st Qu.: -114.8    1st Qu.: 8.5      1st Qu.: -199.0    1st Qu.: -427.2
## Median : 86.0      Median : 138.0    Median : -148.5    Median : -189.5
## Mean : 38.8      Mean : 125.3    Mean : -93.7      Mean : -159.2
## 3rd Qu.: 166.2    3rd Qu.: 268.0    3rd Qu.: -31.0    3rd Qu.: 41.5
## Max. : 232.0      Max. : 406.0      Max. : 179.0      Max. : 532.0
## magnet_forearm_y magnet_forearm_z problem_id
## Min.      : -787.0    Min.      : -32.0      Min.      : 1.00
## 1st Qu.: -328.8    1st Qu.: 275.2      1st Qu.: 5.75
## Median : 487.0      Median : 491.5      Median : 10.50
## Mean : 191.8      Mean : 460.2      Mean : 10.50
## 3rd Qu.: 720.8      3rd Qu.: 661.5      3rd Qu.: 15.25
## Max. : 800.0      Max. : 884.0      Max. : 20.00
```

5 Building the Prediction Model

Three Models are built:

- Random Forest
- SVM (Radial Kernel)
- KNN

Parameters will be tuned via 5-fold cross validation.

5.1 Random Forest

```
cvCtrl <- trainControl(
  method = "cv",
  number = 5,
  allowParallel = TRUE,
  verboseIter = TRUE )

m1 <- train(
  classe ~ .,
  data = dsTrainingCleaned,
  method = "rf",
  trControl = cvCtrl)
```

```
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 2 on full training set
```

5.2 SVM (Radial Kernel)

```
m2 <- train(
  classe ~ .,
  data = dsTrainingCleaned,
  method = "svmRadial",
  trControl = cvCtrl)
```

```
## Loading required package: kernlab
```

```
## Aggregating results
## Selecting tuning parameters
## Fitting sigma = 0.0136, C = 1 on full training set
```

5.3 KNN

```
m3 <- train(
  classe ~ .,
  data = dsTrainingCleaned,
  method = "knn",
  trControl = cvCtrl)
```

```
## Aggregating results
## Selecting tuning parameters
## Fitting k = 5 on full training set
```

5.4 Investigate the Cross Validation Performance Accuracy

```
acc.tab <- data.frame(
  Model=c ( "Random Forest" , "SVM (radial)" , "KNN" ) ,
  Accuracy=c(
    round(max(head(m1$results)$Accuracy), 3 ) ,
    round(max(head(m2$results)$Accuracy), 3 ) ,
    round(max(head(m3$results)$Accuracy), 3 ) )
)
```

```
kable(acc.tab)
```

Model	Accuracy
Random Forest	0.995
SVM (radial)	0.932
KNN	0.921

Random Forest model appears to have the highest cross-validation accuracy, with the SVM and KNN slightly lower.

6 Prediction

```
# Do the predictions
test.pred.1 <- predict(m1, dsTestingCleaned)
test.pred.2 <- predict(m2, dsTestingCleaned)
test.pred.3 <- predict(m3, dsTestingCleaned)
```

```
# Make a table and check if they all agree
pred.df <- data.frame(
  rf.pred = test.pred.1,
  svm.pred = test.pred.2,
  knn.pred = test.pred.3 )
pred.df$agree <- with( pred.df , rf.pred == svm.pred && rf.pred == knn.pred )
all.agree <- all(pred.df$agree)
```

Here are the classifications predictions for the 3 models:

```
colnames(pred.df) <- c("Random Forest", "SVM", "KNN", "All Agree?")
kable(pred.df)
```

Random Forest	SVM	KNN	All Agree?
B	B	B	TRUE
A	A	A	TRUE
B	B	B	TRUE
A	A	A	TRUE
A	A	A	TRUE
E	E	E	TRUE
D	D	D	TRUE
B	B	B	TRUE
A	A	A	TRUE
A	A	A	TRUE
B	B	B	TRUE
C	C	C	TRUE
B	B	B	TRUE
A	A	A	TRUE
E	E	E	TRUE
E	E	E	TRUE
A	A	A	TRUE

B	B	B	TRUE
B	B	B	TRUE
B	B	B	TRUE