Prediction Assignment

Jeremy Eade

October 22, 2016

Prediction Assignment Writeup

1. Objectives

The goal of this project is to predict the manner (A,B,C,D or E) that participants performed an exercise using accelerometer data from the belt, forearm, arm and dumbell of the six particpants.

The report below describes how the model was built, use of cross-validation, estimated sample error, decisions made (eg model selection) and the results obtained (prediction of the twenty different test cases).

2. Read and Prepare Data

```
# For reproducibility
set.seed(62433)

# Read in the data and handle NAs in the source files
training <- read.csv('pml-training.csv',header=TRUE,stringsAsFactors=FALSE,na.strings = c('NA','','#DIV
testing <- read.csv('pml-testing.csv',header=TRUE,stringsAsFactors=FALSE,na.strings = c('NA','','#DIV/0

training$classe <- as.factor(training$classe)

# Look at the data
summary(training)
```

```
##        X          user_name         raw_timestamp_part_1
##  Min.   :    1   Length:19622       Min.   :1.322e+09
##  1st Qu.: 4906   Class :character   1st Qu.:1.323e+09
##  Median : 9812   Mode  :character   Median :1.323e+09
##  Mean   : 9812                      Mean   :1.323e+09
##  3rd Qu.:14717                      3rd Qu.:1.323e+09
##  Max.   :19622                      Max.   :1.323e+09
##
##  raw_timestamp_part_2 cvtd_timestamp      new_window
##  Min.   :   294       Length:19622       Length:19622
##  1st Qu.:252912       Class :character   Class :character
##  Median :496380       Mode  :character   Mode  :character
##  Mean   :500656
##  3rd Qu.:751891
##  Max.   :998801
##
##    num_window      roll_belt        pitch_belt          yaw_belt
##  Min.   :  1.0   Min.   :-28.90   Min.   :-55.8000   Min.   :-180.00
##  1st Qu.:222.0   1st Qu.:  1.10   1st Qu.:  1.7600   1st Qu.: -88.30
##  Median :424.0   Median :113.00   Median :  5.2800   Median : -13.00
##  Mean   :430.6   Mean   : 64.41   Mean   :  0.3053   Mean   : -11.21
```

```
##  3rd Qu.:644.0   3rd Qu.:123.00   3rd Qu.: 14.9000   3rd Qu.:  12.90
##  Max.   :864.0   Max.   :162.00   Max.   : 60.3000   Max.   : 179.00
##
##  total_accel_belt kurtosis_roll_belt kurtosis_picth_belt kurtosis_yaw_belt
##  Min.   : 0.00    Min.   :-2.121     Min.   :-2.190      Mode:logical
##  1st Qu.: 3.00    1st Qu.:-1.329     1st Qu.:-1.107      NA's:19622
##  Median :17.00    Median :-0.899     Median :-0.151
##  Mean   :11.31    Mean   :-0.220     Mean   : 4.334
##  3rd Qu.:18.00    3rd Qu.:-0.219     3rd Qu.: 3.178
##  Max.   :29.00    Max.   :33.000     Max.   :58.000
##                   NA's   :19226      NA's   :19248
##  skewness_roll_belt skewness_roll_belt.1 skewness_yaw_belt
##  Min.   :-5.745     Min.   :-7.616       Mode:logical
##  1st Qu.:-0.444     1st Qu.:-1.114       NA's:19622
##  Median : 0.000     Median :-0.068
##  Mean   :-0.026     Mean   :-0.296
##  3rd Qu.: 0.417     3rd Qu.: 0.661
##  Max.   : 3.595     Max.   : 7.348
##  NA's   :19225      NA's   :19248
##  max_roll_belt     max_picth_belt   max_yaw_belt   min_roll_belt
##  Min.   :-94.300   Min.   : 3.00   Min.   :-2.10   Min.   :-180.00
##  1st Qu.:-88.000   1st Qu.: 5.00   1st Qu.:-1.30   1st Qu.: -88.40
##  Median : -5.100   Median :18.00   Median :-0.90   Median :  -7.85
##  Mean   : -6.667   Mean   :12.92   Mean   :-0.22   Mean   : -10.44
##  3rd Qu.: 18.500   3rd Qu.:19.00   3rd Qu.:-0.20   3rd Qu.:   9.05
##  Max.   :180.000   Max.   :30.00   Max.   :33.00   Max.   : 173.00
##  NA's   :19216     NA's   :19216   NA's   :19226   NA's   :19216
##  min_pitch_belt   min_yaw_belt   amplitude_roll_belt amplitude_pitch_belt
##  Min.   : 0.00   Min.   :-2.10   Min.   :  0.000     Min.   : 0.000
##  1st Qu.: 3.00   1st Qu.:-1.30   1st Qu.:  0.300     1st Qu.: 1.000
##  Median :16.00   Median :-0.90   Median :  1.000     Median : 1.000
##  Mean   :10.76   Mean   :-0.22   Mean   :  3.769     Mean   : 2.167
##  3rd Qu.:17.00   3rd Qu.:-0.20   3rd Qu.:  2.083     3rd Qu.: 2.000
##  Max.   :23.00   Max.   :33.00   Max.   :360.000     Max.   :12.000
##  NA's   :19216   NA's   :19226   NA's   :19216       NA's   :19216
##  amplitude_yaw_belt var_total_accel_belt avg_roll_belt    stddev_roll_belt
##  Min.   :0          Min.   : 0.000       Min.   :-27.40   Min.   : 0.000
##  1st Qu.:0          1st Qu.: 0.100       1st Qu.:  1.10   1st Qu.: 0.200
##  Median :0          Median : 0.200       Median :116.35   Median : 0.400
##  Mean   :0          Mean   : 0.926       Mean   : 68.06   Mean   : 1.337
##  3rd Qu.:0          3rd Qu.: 0.300       3rd Qu.:123.38   3rd Qu.: 0.700
##  Max.   :0          Max.   :16.500       Max.   :157.40   Max.   :14.200
##  NA's   :19226      NA's   :19216        NA's   :19216    NA's   :19216
##  var_roll_belt     avg_pitch_belt    stddev_pitch_belt var_pitch_belt
##  Min.   :  0.000   Min.   :-51.400   Min.   :0.000     Min.   : 0.000
##  1st Qu.:  0.000   1st Qu.:  2.025   1st Qu.:0.200     1st Qu.: 0.000
##  Median :  0.100   Median :  5.200   Median :0.400     Median : 0.100
##  Mean   :  7.699   Mean   :  0.520   Mean   :0.603     Mean   : 0.766
##  3rd Qu.:  0.500   3rd Qu.: 15.775   3rd Qu.:0.700     3rd Qu.: 0.500
##  Max.   :200.700   Max.   : 59.700   Max.   :4.000     Max.   :16.200
##  NA's   :19216     NA's   :19216     NA's   :19216     NA's   :19216
##   avg_yaw_belt      stddev_yaw_belt    var_yaw_belt
##  Min.   :-138.300   Min.   :  0.000   Min.   :    0.000
##  1st Qu.: -88.175   1st Qu.:  0.100   1st Qu.:    0.010
```

```
##  Median :  -6.550   Median :  0.300   Median :    0.090
##  Mean   :  -8.831   Mean   :  1.341   Mean   :  107.487
##  3rd Qu.:  14.125   3rd Qu.:  0.700   3rd Qu.:    0.475
##  Max.   : 173.500   Max.   :176.600   Max.   :31183.240
##  NA's   :19216      NA's   :19216     NA's   :19216
##   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 var_accel_arm     avg_roll_arm
##  Min.   :-180.0000   Min.   : 1.00   Min.   :  0.00   Min.   :-166.67
##  1st Qu.: -43.1000   1st Qu.:17.00   1st Qu.:  9.03   1st Qu.: -38.37
##  Median :   0.0000   Median :27.00   Median : 40.61   Median :   0.00
##  Mean   :  -0.6188   Mean   :25.51   Mean   : 53.23   Mean   :  12.68
##  3rd Qu.:  45.8750   3rd Qu.:33.00   3rd Qu.: 75.62   3rd Qu.:  76.33
##  Max.   : 180.0000   Max.   :66.00   Max.   :331.70   Max.   : 163.33
##                                      NA's   :19216    NA's   :19216
##  stddev_roll_arm    var_roll_arm       avg_pitch_arm     stddev_pitch_arm
##  Min.   :  0.000   Min.   :    0.000   Min.   :-81.773   Min.   : 0.000
##  1st Qu.:  1.376   1st Qu.:    1.898   1st Qu.:-22.770   1st Qu.: 1.642
##  Median :  5.702   Median :   32.517   Median :  0.000   Median : 8.133
##  Mean   : 11.201   Mean   :  417.264   Mean   : -4.901   Mean   :10.383
##  3rd Qu.: 14.921   3rd Qu.:  222.647   3rd Qu.:  8.277   3rd Qu.:16.327
##  Max.   :161.964   Max.   :26232.208   Max.   : 75.659   Max.   :43.412
##  NA's   :19216     NA's   :19216       NA's   :19216     NA's   :19216
##  var_pitch_arm       avg_yaw_arm       stddev_yaw_arm
##  Min.   :   0.000   Min.   :-173.440   Min.   :  0.000
##  1st Qu.:   2.697   1st Qu.: -29.198   1st Qu.:  2.577
##  Median :  66.146   Median :   0.000   Median : 16.682
##  Mean   : 195.864   Mean   :   2.359   Mean   : 22.270
##  3rd Qu.: 266.576   3rd Qu.:  38.185   3rd Qu.: 35.984
##  Max.   :1884.565   Max.   : 152.000   Max.   :177.044
##  NA's   :19216      NA's   :19216      NA's   :19216
##   var_yaw_arm         gyros_arm_x        gyros_arm_y
```

```
##  Min.   :    0.000   Min.   :-6.37000   Min.   :-3.4400
##  1st Qu.:    6.642   1st Qu.:-1.33000   1st Qu.:-0.8000
##  Median :  278.309   Median : 0.08000   Median :-0.2400
##  Mean   : 1055.933   Mean   : 0.04277   Mean   :-0.2571
##  3rd Qu.: 1294.850   3rd Qu.: 1.57000   3rd Qu.: 0.1400
##  Max.   :31344.568   Max.   : 4.87000   Max.   : 2.8400
##  NA's   :19216
##   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    kurtosis_roll_arm
##  Min.   :-584.0   Min.   :-392.0   Min.   :-597.0   Min.   :-1.809
##  1st Qu.:-300.0   1st Qu.:  -9.0   1st Qu.: 131.2   1st Qu.:-1.345
##  Median : 289.0   Median : 202.0   Median : 444.0   Median :-0.894
##  Mean   : 191.7   Mean   : 156.6   Mean   : 306.5   Mean   :-0.366
##  3rd Qu.: 637.0   3rd Qu.: 323.0   3rd Qu.: 545.0   3rd Qu.:-0.038
##  Max.   : 782.0   Max.   : 583.0   Max.   : 694.0   Max.   :21.456
##                                                     NA's   :19294
##  kurtosis_picth_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm
##  Min.   :-2.084     Min.   :-2.103   Min.   :-2.541    Min.   :-4.565
##  1st Qu.:-1.280     1st Qu.:-1.220   1st Qu.:-0.561    1st Qu.:-0.618
##  Median :-1.010     Median :-0.733   Median : 0.040    Median :-0.035
##  Mean   :-0.542     Mean   : 0.406   Mean   : 0.068    Mean   :-0.065
##  3rd Qu.:-0.379     3rd Qu.: 0.115   3rd Qu.: 0.671    3rd Qu.: 0.454
##  Max.   :19.751     Max.   :56.000   Max.   : 4.394    Max.   : 3.043
##  NA's   :19296      NA's   :19227    NA's   :19293     NA's   :19296
##  skewness_yaw_arm  max_roll_arm     max_picth_arm       max_yaw_arm
##  Min.   :-6.708   Min.   :-73.100   Min.   :-173.000   Min.   : 4.00
##  1st Qu.:-0.743   1st Qu.: -0.175   1st Qu.:  -1.975   1st Qu.:29.00
##  Median :-0.133   Median :  4.950   Median :  23.250   Median :34.00
##  Mean   :-0.229   Mean   : 11.236   Mean   :  35.751   Mean   :35.46
##  3rd Qu.: 0.344   3rd Qu.: 26.775   3rd Qu.:  95.975   3rd Qu.:41.00
##  Max.   : 7.483   Max.   : 85.500   Max.   : 180.000   Max.   :65.00
##  NA's   :19227    NA's   :19216     NA's   :19216      NA's   :19216
##   min_roll_arm    min_pitch_arm      min_yaw_arm    amplitude_roll_arm
##  Min.   :-89.10   Min.   :-180.00   Min.   : 1.00   Min.   :  0.000
##  1st Qu.:-41.98   1st Qu.: -72.62   1st Qu.: 8.00   1st Qu.:  5.425
##  Median :-22.45   Median : -33.85   Median :13.00   Median : 28.450
##  Mean   :-21.22   Mean   : -33.92   Mean   :14.66   Mean   : 32.452
##  3rd Qu.:  0.00   3rd Qu.:   0.00   3rd Qu.:19.00   3rd Qu.: 50.960
##  Max.   : 66.40   Max.   : 152.00   Max.   :38.00   Max.   :119.500
##  NA's   :19216    NA's   :19216     NA's   :19216   NA's   :19216
##  amplitude_pitch_arm amplitude_yaw_arm roll_dumbbell     pitch_dumbbell
##  Min.   :  0.000     Min.   : 0.00     Min.   :-153.71   Min.   :-149.59
##  1st Qu.:  9.925     1st Qu.:13.00     1st Qu.: -18.49   1st Qu.: -40.89
##  Median : 54.900     Median :22.00     Median :  48.17   Median : -20.96
##  Mean   : 69.677     Mean   :20.79     Mean   :  23.84   Mean   : -10.78
##  3rd Qu.:115.175     3rd Qu.:28.75     3rd Qu.:  67.61   3rd Qu.:  17.50
##  Max.   :360.000     Max.   :52.00     Max.   : 153.55   Max.   : 149.40
```

```
##  NA's   :19216       NA's   :19216
##   yaw_dumbbell      kurtosis_roll_dumbbell kurtosis_picth_dumbbell
##  Min.   :-150.871   Min.   :-2.174         Min.   :-2.200
##  1st Qu.: -77.644   1st Qu.:-0.682         1st Qu.:-0.721
##  Median :  -3.324   Median :-0.033         Median :-0.133
##  Mean   :   1.674   Mean   : 0.452         Mean   : 0.286
##  3rd Qu.:  79.643   3rd Qu.: 0.940         3rd Qu.: 0.584
##  Max.   : 154.952   Max.   :54.998         Max.   :55.628
##                     NA's   :19221          NA's   :19218
##  kurtosis_yaw_dumbbell skewness_roll_dumbbell skewness_pitch_dumbbell
##  Mode:logical          Min.   :-7.384         Min.   :-7.447
##  NA's:19622            1st Qu.:-0.581         1st Qu.:-0.526
##                        Median :-0.076         Median :-0.091
##                        Mean   :-0.115         Mean   :-0.035
##                        3rd Qu.: 0.400         3rd Qu.: 0.505
##                        Max.   : 1.958         Max.   : 3.769
##                        NA's   :19220          NA's   :19217
##  skewness_yaw_dumbbell max_roll_dumbbell max_picth_dumbbell
##  Mode:logical          Min.   :-70.10    Min.   :-112.90
##  NA's:19622            1st Qu.:-27.15    1st Qu.: -66.70
##                        Median : 14.85    Median :  40.05
##                        Mean   : 13.76    Mean   :  32.75
##                        3rd Qu.: 50.58    3rd Qu.: 133.22
##                        Max.   :137.00    Max.   : 155.00
##                        NA's   :19216     NA's   :19216
##  max_yaw_dumbbell min_roll_dumbbell min_pitch_dumbbell min_yaw_dumbbell
##  Min.   :-2.20    Min.   :-149.60   Min.   :-147.00    Min.   :-2.20
##  1st Qu.:-0.70    1st Qu.: -59.67   1st Qu.: -91.80    1st Qu.:-0.70
##  Median : 0.00    Median : -43.55   Median : -66.15    Median : 0.00
##  Mean   : 0.45    Mean   : -41.24   Mean   : -33.18    Mean   : 0.45
##  3rd Qu.: 0.90    3rd Qu.: -25.20   3rd Qu.:  21.20    3rd Qu.: 0.90
##  Max.   :55.00    Max.   :  73.20   Max.   : 120.90    Max.   :55.00
##  NA's   :19221    NA's   :19216     NA's   :19216      NA's   :19221
##  amplitude_roll_dumbbell amplitude_pitch_dumbbell amplitude_yaw_dumbbell
##  Min.   :  0.00          Min.   :  0.00           Min.   :0
##  1st Qu.: 14.97          1st Qu.: 17.06           1st Qu.:0
##  Median : 35.05          Median : 41.73           Median :0
##  Mean   : 55.00          Mean   : 65.93           Mean   :0
##  3rd Qu.: 81.04          3rd Qu.: 99.55           3rd Qu.:0
##  Max.   :256.48          Max.   :273.59           Max.   :0
##  NA's   :19216           NA's   :19216            NA's   :19221
##  total_accel_dumbbell var_accel_dumbbell avg_roll_dumbbell
##  Min.   : 0.00        Min.   :  0.000    Min.   :-128.96
##  1st Qu.: 4.00        1st Qu.:  0.378    1st Qu.: -12.33
##  Median :10.00        Median :  1.000    Median :  48.23
##  Mean   :13.72        Mean   :  4.388    Mean   :  23.86
##  3rd Qu.:19.00        3rd Qu.:  3.434    3rd Qu.:  64.37
##  Max.   :58.00        Max.   :230.428    Max.   : 125.99
##                       NA's   :19216      NA's   :19216
##  stddev_roll_dumbbell var_roll_dumbbell  avg_pitch_dumbbell
##  Min.   :  0.000      Min.   :    0.00   Min.   :-70.73
##  1st Qu.:  4.639      1st Qu.:   21.52   1st Qu.:-42.00
##  Median : 12.204      Median :  148.95   Median :-19.91
##  Mean   : 20.761      Mean   : 1020.27   Mean   :-12.33
```

```
##  3rd Qu.: 26.356      3rd Qu.:  694.65   3rd Qu.: 13.21
##  Max.   :123.778      Max.   :15321.01   Max.   : 94.28
##  NA's   :19216        NA's   :19216      NA's   :19216
##  stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
##  Min.   : 0.000        Min.   :   0.00    Min.   :-117.950
##  1st Qu.: 3.482        1st Qu.:  12.12    1st Qu.: -76.696
##  Median : 8.089        Median :  65.44    Median :  -4.505
##  Mean   :13.147        Mean   : 350.31    Mean   :   0.202
##  3rd Qu.:19.238        3rd Qu.: 370.11    3rd Qu.:  71.234
##  Max.   :82.680        Max.   :6836.02    Max.   : 134.905
##  NA's   :19216         NA's   :19216      NA's   :19216
##  stddev_yaw_dumbbell var_yaw_dumbbell   gyros_dumbbell_x
##  Min.   :  0.000     Min.   :    0.00   Min.   :-204.0000
##  1st Qu.:  3.885     1st Qu.:   15.09   1st Qu.:  -0.0300
##  Median : 10.264     Median :  105.35   Median :   0.1300
##  Mean   : 16.647     Mean   :  589.84   Mean   :   0.1611
##  3rd Qu.: 24.674     3rd Qu.:  608.79   3rd Qu.:   0.3500
##  Max.   :107.088     Max.   :11467.91   Max.   :   2.2200
##  NA's   :19216       NA's   :19216
##  gyros_dumbbell_y   gyros_dumbbell_z  accel_dumbbell_x  accel_dumbbell_y
##  Min.   :-2.10000   Min.   : -2.380   Min.   :-419.00   Min.   :-189.00
##  1st Qu.:-0.14000   1st Qu.: -0.310   1st Qu.: -50.00   1st Qu.:  -8.00
##  Median : 0.03000   Median : -0.130   Median :  -8.00   Median :  41.50
##  Mean   : 0.04606   Mean   : -0.129   Mean   : -28.62   Mean   :  52.63
##  3rd Qu.: 0.21000   3rd Qu.:  0.030   3rd Qu.:  11.00   3rd Qu.: 111.00
##  Max.   :52.00000   Max.   :317.000   Max.   : 235.00   Max.   : 315.00
##
##  accel_dumbbell_z  magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z
##  Min.   :-334.00   Min.   :-643.0    Min.   :-3600     Min.   :-262.00
##  1st Qu.:-142.00   1st Qu.:-535.0    1st Qu.:  231     1st Qu.: -45.00
##  Median :  -1.00   Median :-479.0    Median :  311     Median :  13.00
##  Mean   : -38.32   Mean   :-328.5    Mean   :  221     Mean   :  46.05
##  3rd Qu.:  38.00   3rd Qu.:-304.0    3rd Qu.:  390     3rd Qu.:  95.00
##  Max.   : 318.00   Max.   : 592.0    Max.   :  633     Max.   : 452.00
##
##   roll_forearm       pitch_forearm     yaw_forearm
##  Min.   :-180.0000   Min.   :-72.50   Min.   :-180.00
##  1st Qu.:  -0.7375   1st Qu.:  0.00   1st Qu.: -68.60
##  Median :  21.7000   Median :  9.24   Median :   0.00
##  Mean   :  33.8265   Mean   : 10.71   Mean   :  19.21
##  3rd Qu.: 140.0000   3rd Qu.: 28.40   3rd Qu.: 110.00
##  Max.   : 180.0000   Max.   : 89.80   Max.   : 180.00
##
##  kurtosis_roll_forearm kurtosis_picth_forearm kurtosis_yaw_forearm
##  Min.   :-1.879        Min.   :-2.098         Mode:logical
##  1st Qu.:-1.398        1st Qu.:-1.376         NA's:19622
##  Median :-1.119        Median :-0.890
##  Mean   :-0.689        Mean   : 0.419
##  3rd Qu.:-0.618        3rd Qu.: 0.054
##  Max.   :40.060        Max.   :33.626
##  NA's   :19300         NA's   :19301
##  skewness_roll_forearm skewness_pitch_forearm skewness_yaw_forearm
##  Min.   :-2.297        Min.   :-5.241         Mode:logical
##  1st Qu.:-0.402        1st Qu.:-0.881         NA's:19622
```

```
##  Median : 0.003        Median :-0.156
##  Mean   :-0.009        Mean   :-0.223
##  3rd Qu.: 0.370        3rd Qu.: 0.514
##  Max.   : 5.856        Max.   : 4.464
##  NA's   :19299         NA's   :19301
##  max_roll_forearm max_picth_forearm max_yaw_forearm  min_roll_forearm
##  Min.   :-66.60   Min.   :-151.00   Min.   :-1.900   Min.   :-72.500
##  1st Qu.:  0.00   1st Qu.:   0.00   1st Qu.:-1.400   1st Qu.: -6.075
##  Median : 26.80   Median : 113.00   Median :-1.100   Median :  0.000
##  Mean   : 24.49   Mean   :  81.49   Mean   :-0.689   Mean   : -0.167
##  3rd Qu.: 45.95   3rd Qu.: 174.75   3rd Qu.:-0.600   3rd Qu.: 12.075
##  Max.   : 89.80   Max.   : 180.00   Max.   :40.100   Max.   : 62.100
##  NA's   :19216    NA's   :19216     NA's   :19300    NA's   :19216
##  min_pitch_forearm min_yaw_forearm  amplitude_roll_forearm
##  Min.   :-180.00   Min.   :-1.900   Min.   :  0.000
##  1st Qu.:-175.00   1st Qu.:-1.400   1st Qu.:  1.125
##  Median : -61.00   Median :-1.100   Median : 17.770
##  Mean   : -57.57   Mean   :-0.689   Mean   : 24.653
##  3rd Qu.:   0.00   3rd Qu.:-0.600   3rd Qu.: 39.875
##  Max.   : 167.00   Max.   :40.100   Max.   :126.000
##  NA's   :19216     NA's   :19300    NA's   :19216
##  amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
##  Min.   :  0.0           Min.   :0             Min.   :  0.00
##  1st Qu.:  2.0           1st Qu.:0             1st Qu.: 29.00
##  Median : 83.7           Median :0             Median : 36.00
##  Mean   :139.1           Mean   :0             Mean   : 34.72
##  3rd Qu.:350.0           3rd Qu.:0             3rd Qu.: 41.00
##  Max.   :360.0           Max.   :0             Max.   :108.00
##  NA's   :19216           NA's   :19300
##  var_accel_forearm avg_roll_forearm   stddev_roll_forearm
##  Min.   :  0.000   Min.   :-177.234   Min.   :  0.000
##  1st Qu.:  6.759   1st Qu.:  -0.909   1st Qu.:  0.428
##  Median : 21.165   Median :  11.172   Median :  8.030
##  Mean   : 33.502   Mean   :  33.165   Mean   : 41.986
##  3rd Qu.: 51.240   3rd Qu.: 107.132   3rd Qu.: 85.373
##  Max.   :172.606   Max.   : 177.256   Max.   :179.171
##  NA's   :19216     NA's   :19216      NA's   :19216
##  var_roll_forearm   avg_pitch_forearm stddev_pitch_forearm
##  Min.   :    0.00   Min.   :-68.17    Min.   : 0.000
##  1st Qu.:    0.18   1st Qu.:  0.00    1st Qu.: 0.336
##  Median :   64.48   Median : 12.02    Median : 5.516
##  Mean   : 5274.10   Mean   : 11.79    Mean   : 7.977
##  3rd Qu.: 7289.08   3rd Qu.: 28.48    3rd Qu.:12.866
##  Max.   :32102.24   Max.   : 72.09    Max.   :47.745
##  NA's   :19216      NA's   :19216     NA's   :19216
##  var_pitch_forearm  avg_yaw_forearm   stddev_yaw_forearm
##  Min.   :   0.000   Min.   :-155.06   Min.   :  0.000
##  1st Qu.:   0.113   1st Qu.: -26.26   1st Qu.:  0.524
##  Median :  30.425   Median :   0.00   Median : 24.743
##  Mean   : 139.593   Mean   :  18.00   Mean   : 44.854
##  3rd Qu.: 165.532   3rd Qu.:  85.79   3rd Qu.: 85.817
##  Max.   :2279.617   Max.   : 169.24   Max.   :197.508
##  NA's   :19216      NA's   :19216     NA's   :19216
##  var_yaw_forearm    gyros_forearm_x   gyros_forearm_y
```

```
##  Min.   :    0.00   Min.   :-22.000   Min.   : -7.02000
##  1st Qu.:    0.27   1st Qu.: -0.220   1st Qu.: -1.46000
##  Median :  612.21   Median :  0.050   Median :  0.03000
##  Mean   : 4639.85   Mean   :  0.158   Mean   :  0.07517
##  3rd Qu.: 7368.41   3rd Qu.:  0.560   3rd Qu.:  1.62000
##  Max.   :39009.33   Max.   :  3.970   Max.   :311.00000
##  NA's   :19216
##  gyros_forearm_z    accel_forearm_x   accel_forearm_y  accel_forearm_z
##  Min.   : -8.0900   Min.   :-498.00   Min.   :-632.0   Min.   :-446.00
##  1st Qu.: -0.1800   1st Qu.:-178.00   1st Qu.:  57.0   1st Qu.:-182.00
##  Median :  0.0800   Median : -57.00   Median : 201.0   Median : -39.00
##  Mean   :  0.1512   Mean   : -61.65   Mean   : 163.7   Mean   : -55.29
##  3rd Qu.:  0.4900   3rd Qu.:  76.00   3rd Qu.: 312.0   3rd Qu.:  26.00
##  Max.   :231.0000   Max.   : 477.00   Max.   : 923.0   Max.   : 291.00
##
##  magnet_forearm_x  magnet_forearm_y magnet_forearm_z classe
##  Min.   :-1280.0   Min.   :-896.0   Min.   :-973.0   A:5580
##  1st Qu.: -616.0   1st Qu.:   2.0   1st Qu.: 191.0   B:3797
##  Median : -378.0   Median : 591.0   Median : 511.0   C:3422
##  Mean   : -312.6   Mean   : 380.1   Mean   : 393.6   D:3216
##  3rd Qu.:  -73.0   3rd Qu.: 737.0   3rd Qu.: 653.0   E:3607
##  Max.   :  672.0   Max.   :1480.0   Max.   :1090.0
##
```

```
summary(testing)
```

```
##        X          user_name         raw_timestamp_part_1
##  Min.   : 1.00   Length:20          Min.   :1.322e+09
##  1st Qu.: 5.75   Class :character   1st Qu.:1.323e+09
##  Median :10.50   Mode  :character   Median :1.323e+09
##  Mean   :10.50                      Mean   :1.323e+09
##  3rd Qu.:15.25                      3rd Qu.:1.323e+09
##  Max.   :20.00                      Max.   :1.323e+09
##  raw_timestamp_part_2 cvtd_timestamp      new_window
##  Min.   : 36553       Length:20          Length:20
##  1st Qu.:268655       Class :character   Class :character
##  Median :530706       Mode  :character   Mode  :character
##  Mean   :512167
##  3rd Qu.:787738
##  Max.   :920315
##    num_window      roll_belt          pitch_belt         yaw_belt
##  Min.   : 48.0   Min.   : -5.9200   Min.   :-41.600   Min.   :-93.70
##  1st Qu.:250.0   1st Qu.:  0.9075   1st Qu.:  3.013   1st Qu.:-88.62
##  Median :384.5   Median :  1.1100   Median :  4.655   Median :-87.85
##  Mean   :379.6   Mean   : 31.3055   Mean   :  5.824   Mean   :-59.30
##  3rd Qu.:467.0   3rd Qu.: 32.5050   3rd Qu.:  6.135   3rd Qu.:-63.50
##  Max.   :859.0   Max.   :129.0000   Max.   : 27.800   Max.   :162.00
##  total_accel_belt kurtosis_roll_belt kurtosis_picth_belt kurtosis_yaw_belt
##  Min.   : 2.00    Mode:logical       Mode:logical        Mode:logical
##  1st Qu.: 3.00    NA's:20            NA's:20             NA's:20
##  Median : 4.00
##  Mean   : 7.55
##  3rd Qu.: 8.00
##  Max.   :21.00
```

```
##  skewness_roll_belt skewness_roll_belt.1 skewness_yaw_belt max_roll_belt
##  Mode:logical       Mode:logical         Mode:logical      Mode:logical
##  NA's:20            NA:s:20              NA:s:20           NA:s:20
##
##
##
##
##  max_picth_belt max_yaw_belt   min_roll_belt  min_pitch_belt
##  Mode:logical   Mode:logical   Mode:logical   Mode:logical
##  NA's:20        NA:s:20        NA:s:20        NA:s:20
##
##
##
##
##  min_yaw_belt   amplitude_roll_belt amplitude_pitch_belt
##  Mode:logical   Mode:logical        Mode:logical
##  NA's:20        NA:s:20             NA:s:20
##
##
##
##
##  amplitude_yaw_belt var_total_accel_belt avg_roll_belt  stddev_roll_belt
##  Mode:logical       Mode:logical         Mode:logical   Mode:logical
##  NA's:20            NA:s:20              NA:s:20        NA:s:20
##
##
##
##
##  var_roll_belt  avg_pitch_belt stddev_pitch_belt var_pitch_belt
##  Mode:logical   Mode:logical   Mode:logical      Mode:logical
##  NA's:20        NA:s:20        NA:s:20           NA:s:20
##
##
##
##
##  avg_yaw_belt   stddev_yaw_belt var_yaw_belt    gyros_belt_x
##  Mode:logical   Mode:logical    Mode:logical   Min.   :-0.500
##  NA's:20        NA:s:20         NA:s:20        1st Qu.:-0.070
##                                                Median : 0.020
##                                                Mean   :-0.045
##                                                3rd Qu.: 0.070
##                                                Max.   : 0.240
##   gyros_belt_y     gyros_belt_z      accel_belt_x     accel_belt_y
##  Min.   :-0.050   Min.   :-0.4800   Min.   :-48.00   Min.   :-16.00
##  1st Qu.:-0.005   1st Qu.:-0.1375   1st Qu.:-19.00   1st Qu.:  2.00
##  Median : 0.000   Median :-0.0250   Median :-13.00   Median :  4.50
##  Mean   : 0.010   Mean   :-0.1005   Mean   :-13.50   Mean   : 18.35
##  3rd Qu.: 0.020   3rd Qu.: 0.0000   3rd Qu.: -8.75   3rd Qu.: 25.50
##  Max.   : 0.110   Max.   : 0.0500   Max.   : 46.00   Max.   : 72.00
##   accel_belt_z     magnet_belt_x    magnet_belt_y   magnet_belt_z
##  Min.   :-187.00   Min.   :-13.00   Min.   :566.0   Min.   :-426.0
##  1st Qu.: -24.00   1st Qu.:  5.50   1st Qu.:578.5   1st Qu.:-398.5
##  Median :  27.00   Median : 33.50   Median :600.5   Median :-313.5
##  Mean   : -17.60   Mean   : 35.15   Mean   :601.5   Mean   :-346.9
```

```
##  3rd Qu.:  38.25   3rd Qu.: 46.25   3rd Qu.:631.2   3rd Qu.:-305.0
##  Max.   :  49.00   Max.   :169.00   Max.   :638.0   Max.   :-291.0
##     roll_arm         pitch_arm          yaw_arm        total_accel_arm
##  Min.   :-137.00   Min.   :-63.800   Min.   :-167.00   Min.   : 3.00
##  1st Qu.:   0.00   1st Qu.: -9.188   1st Qu.: -60.15   1st Qu.:20.25
##  Median :   0.00   Median :  0.000   Median :   0.00   Median :29.50
##  Mean   :  16.42   Mean   : -3.950   Mean   :  -2.80   Mean   :26.40
##  3rd Qu.:  71.53   3rd Qu.:  3.465   3rd Qu.:  25.50   3rd Qu.:33.25
##  Max.   : 152.00   Max.   : 55.000   Max.   : 178.00   Max.   :44.00
##  var_accel_arm  avg_roll_arm   stddev_roll_arm var_roll_arm
##  Mode:logical   Mode:logical   Mode:logical    Mode:logical
##  NA's:20        NA's:20        NA's:20         NA's:20
##
##
##
##
##  avg_pitch_arm  stddev_pitch_arm var_pitch_arm  avg_yaw_arm
##  Mode:logical   Mode:logical     Mode:logical   Mode:logical
##  NA's:20        NA's:20          NA's:20        NA's:20
##
##
##
##
##  stddev_yaw_arm var_yaw_arm     gyros_arm_x      gyros_arm_y
##  Mode:logical   Mode:logical   Min.   :-3.710   Min.   :-2.0900
##  NA's:20        NA's:20        1st Qu.:-0.645   1st Qu.:-0.6350
##                                Median : 0.020   Median :-0.0400
##                                Mean   : 0.077   Mean   :-0.1595
##                                3rd Qu.: 1.248   3rd Qu.: 0.2175
##                                Max.   : 3.660   Max.   : 1.8500
##   gyros_arm_z       accel_arm_x      accel_arm_y      accel_arm_z
##  Min.   :-0.6900   Min.   :-341.0   Min.   :-65.00   Min.   :-404.00
##  1st Qu.:-0.1800   1st Qu.:-277.0   1st Qu.: 52.25   1st Qu.:-128.50
##  Median :-0.0250   Median :-194.5   Median :112.00   Median : -83.50
##  Mean   : 0.1205   Mean   :-134.6   Mean   :103.10   Mean   : -87.85
##  3rd Qu.: 0.5650   3rd Qu.:   5.5   3rd Qu.:168.25   3rd Qu.: -27.25
##  Max.   : 1.1300   Max.   : 106.0   Max.   :245.00   Max.   :  93.00
##   magnet_arm_x      magnet_arm_y     magnet_arm_z    kurtosis_roll_arm
##  Min.   :-428.00   Min.   :-307.0   Min.   :-499.0   Mode:logical
##  1st Qu.:-373.75   1st Qu.: 205.2   1st Qu.: 403.0   NA's:20
##  Median :-265.00   Median : 291.0   Median : 476.5
##  Mean   : -38.95   Mean   : 239.4   Mean   : 369.8
##  3rd Qu.: 250.50   3rd Qu.: 358.8   3rd Qu.: 517.0
##  Max.   : 750.00   Max.   : 474.0   Max.   : 633.0
##  kurtosis_picth_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm
##  Mode:logical       Mode:logical     Mode:logical      Mode:logical
##  NA's:20            NA's:20          NA's:20           NA's:20
##
##
##
##
##  skewness_yaw_arm max_roll_arm   max_picth_arm  max_yaw_arm
##  Mode:logical     Mode:logical   Mode:logical   Mode:logical
##  NA's:20          NA's:20        NA's:20        NA's:20
```

```
##
##
##
##
##  min_roll_arm   min_pitch_arm  min_yaw_arm    amplitude_roll_arm
##  Mode:logical   Mode:logical   Mode:logical   Mode:logical
##  NA's:20        NA's:20        NA's:20        NA's:20
##
##
##
##
##  amplitude_pitch_arm amplitude_yaw_arm roll_dumbbell      pitch_dumbbell
##  Mode:logical        Mode:logical      Min.   :-111.118   Min.   :-54.97
##  NA's:20             NA's:20           1st Qu.:   7.494   1st Qu.:-51.89
##                                        Median :  50.403   Median :-40.81
##                                        Mean   :  33.760   Mean   :-19.47
##                                        3rd Qu.:  58.129   3rd Qu.: 16.12
##                                        Max.   : 123.984   Max.   : 96.87
##   yaw_dumbbell       kurtosis_roll_dumbbell kurtosis_picth_dumbbell
##  Min.   :-103.3200   Mode:logical           Mode:logical
##  1st Qu.: -75.2809   NA's:20                NA's:20
##  Median :  -8.2863
##  Mean   :  -0.9385
##  3rd Qu.:  55.8335
##  Max.   : 132.2337
##  kurtosis_yaw_dumbbell skewness_roll_dumbbell skewness_pitch_dumbbell
##  Mode:logical          Mode:logical           Mode:logical
##  NA's:20               NA's:20                NA's:20
##
##
##
##
##  skewness_yaw_dumbbell max_roll_dumbbell max_picth_dumbbell
##  Mode:logical          Mode:logical      Mode:logical
##  NA's:20               NA's:20           NA's:20
##
##
##
##
##  max_yaw_dumbbell min_roll_dumbbell min_pitch_dumbbell min_yaw_dumbbell
##  Mode:logical     Mode:logical      Mode:logical       Mode:logical
##  NA's:20          NA's:20           NA's:20            NA's:20
##
##
##
##
##  amplitude_roll_dumbbell amplitude_pitch_dumbbell amplitude_yaw_dumbbell
##  Mode:logical            Mode:logical             Mode:logical
##  NA's:20                 NA's:20                  NA's:20
##
##
##
##
##  total_accel_dumbbell var_accel_dumbbell avg_roll_dumbbell
```

```
##  Min.   : 1.0         Mode:logical       Mode:logical
##  1st Qu.: 7.0         NA's:20            NA's:20
##  Median :15.5
##  Mean   :17.2
##  3rd Qu.:29.0
##  Max.   :31.0
##  stddev_roll_dumbbell var_roll_dumbbell avg_pitch_dumbbell
##  Mode:logical         Mode:logical      Mode:logical
##  NA's:20              NA's:20           NA's:20
##
##
##
##
##  stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
##  Mode:logical          Mode:logical       Mode:logical
##  NA's:20               NA's:20            NA's:20
##
##
##
##
##  stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x  gyros_dumbbell_y
##  Mode:logical        Mode:logical     Min.   :-1.0300   Min.   :-1.1100
##  NA's:20             NA's:20          1st Qu.: 0.1600   1st Qu.:-0.2100
##                                       Median : 0.3600   Median : 0.0150
##                                       Mean   : 0.2690   Mean   : 0.0605
##                                       3rd Qu.: 0.4625   3rd Qu.: 0.1450
##                                       Max.   : 1.0600   Max.   : 1.9100
##  gyros_dumbbell_z accel_dumbbell_x  accel_dumbbell_y accel_dumbbell_z
##  Min.   :-1.180   Min.   :-159.00   Min.   :-30.00   Min.   :-221.0
##  1st Qu.:-0.485   1st Qu.:-140.25   1st Qu.:  5.75   1st Qu.:-192.2
##  Median :-0.280   Median : -19.00   Median : 71.50   Median :  -3.0
##  Mean   :-0.266   Mean   : -47.60   Mean   : 70.55   Mean   : -60.0
##  3rd Qu.:-0.165   3rd Qu.:  15.75   3rd Qu.:151.25   3rd Qu.:  76.5
##  Max.   : 1.100   Max.   : 185.00   Max.   :166.00   Max.   : 100.0
##  magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z  roll_forearm
##  Min.   :-576.0    Min.   :-558.0    Min.   :-164.00   Min.   :-176.00
##  1st Qu.:-528.0    1st Qu.: 259.5    1st Qu.: -33.00   1st Qu.: -40.25
##  Median :-508.5    Median : 316.0    Median :  49.50   Median :  94.20
##  Mean   :-304.2    Mean   : 189.3    Mean   :  71.40   Mean   :  38.66
##  3rd Qu.:-317.0    3rd Qu.: 348.2    3rd Qu.:  96.25   3rd Qu.: 143.25
##  Max.   : 523.0    Max.   : 403.0    Max.   : 368.00   Max.   : 176.00
##  pitch_forearm      yaw_forearm       kurtosis_roll_forearm
##  Min.   :-63.500   Min.   :-168.000   Mode:logical
##  1st Qu.:-11.457   1st Qu.: -93.375   NA's:20
##  Median :  8.830   Median : -19.250
##  Mean   :  7.099   Mean   :   2.195
##  3rd Qu.: 28.500   3rd Qu.: 104.500
##  Max.   : 59.300   Max.   : 159.000
##  kurtosis_picth_forearm kurtosis_yaw_forearm skewness_roll_forearm
##  Mode:logical           Mode:logical         Mode:logical
##  NA's:20                NA's:20              NA's:20
##
##
##
```

```
##
##  skewness_pitch_forearm skewness_yaw_forearm max_roll_forearm
##  Mode:logical           Mode:logical         Mode:logical
##  NA's:20                NA:s:20              NA's:20
##
##
##
##
##  max_picth_forearm max_yaw_forearm min_roll_forearm min_pitch_forearm
##  Mode:logical      Mode:logical    Mode:logical     Mode:logical
##  NA's:20           NA's:20         NA's:20          NA's:20
##
##
##
##
##  min_yaw_forearm amplitude_roll_forearm amplitude_pitch_forearm
##  Mode:logical    Mode:logical           Mode:logical
##  NA's:20         NA's:20                NA's:20
##
##
##
##
##  amplitude_yaw_forearm total_accel_forearm var_accel_forearm
##  Mode:logical          Min.   :21.00       Mode:logical
##  NA's:20               1st Qu.:24.00       NA's:20
##                        Median :32.50
##                        Mean   :32.05
##                        3rd Qu.:36.75
##                        Max.   :47.00
##  avg_roll_forearm stddev_roll_forearm var_roll_forearm avg_pitch_forearm
##  Mode:logical     Mode:logical        Mode:logical     Mode:logical
##  NA's:20          NA's:20             NA's:20          NA's:20
##
##
##
##
##  stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm stddev_yaw_forearm
##  Mode:logical         Mode:logical      Mode:logical    Mode:logical
##  NA's:20              NA's:20           NA's:20         NA's:20
##
##
##
##
##  var_yaw_forearm gyros_forearm_x   gyros_forearm_y   gyros_forearm_z
##  Mode:logical    Min.   :-1.0600   Min.   :-5.9700   Min.   :-1.2600
##  NA's:20         1st Qu.:-0.5850   1st Qu.:-1.2875   1st Qu.:-0.0975
##                  Median : 0.0200   Median : 0.0350   Median : 0.2300
##                  Mean   :-0.0200   Mean   :-0.0415   Mean   : 0.2610
##                  3rd Qu.: 0.2925   3rd Qu.: 2.0475   3rd Qu.: 0.7625
##                  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
```

```
# Remove NAs
naTrain <- apply(training,2,function(x){sum(is.na(x))})
trainClean <- training[,which(naTrain==0)]

naTest <- apply(testing,2,function(x){sum(is.na(x))})
testClean <- testing[,which(naTest==0)]

# Remove Near Zero Variables
nzvTrain <- nearZeroVar(trainClean,saveMetrics=TRUE)
trainClean <- trainClean[,which(nzvTrain$nzv==FALSE)]
dim(trainClean)
```

```
## [1] 19622    59
```

```
nzvTest<- nearZeroVar(testClean,saveMetrics=TRUE)
testClean <- testClean[,which(nzvTest$nzv==FALSE)]
dim(testClean)
```

```
## [1] 20 59
```

```
# Remove first six columns - we want to include the data from accelerometers on the belt, forearm, arm,
trainClean <- trainClean[7:NCOL(trainClean)]
testClean <- testClean[7:NCOL(testClean)]

# Split training data into training and cross-validation
# We need a separate cross-validation data set to see how well the model predicts
inTrain <- createDataPartition(trainClean$classe,p=0.6,list=FALSE)
trainFinal <- trainClean[inTrain,]
trainCV <- trainClean[-inTrain,]
dim(trainFinal)
```

```
## [1] 11776    53
```

```
dim(trainCV)
```

```
## [1] 7846   53
```

3. Train the model and check out-of-sample error

```r
# Train the model - commented out in the HTML for performance reasons
# Use the saved model in modFit.rda instead
# I'm using the Random Forest model as it seemed to be a strong performer in the lecture notes and in K
# The training method will be Cross Validation with 5 folds.  We want parallel processing if possible.
#modFit <- train(classe~., method='rf',data=trainFinal,trControl=trainControl(method='cv',number=5),all
#modFit

# Save the model
#save(modFit,file="modFit.rda")
load("modFit.rda")
modFit
```

```
## Random Forest
##
## 11776 samples
##    52 predictor
##     5 classes: 'A', 'B', 'C', 'D', 'E'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 10600, 10599, 10598, 10598, 10598, 10598, ...
## Resampling results across tuning parameters:
##
##   mtry  Accuracy   Kappa
##    2    0.9903202  0.9877551
##   27    0.9906598  0.9881843
##   52    0.9800452  0.9747544
##
## Accuracy was used to select the optimal model using  the largest value.
## The final value used for the model was mtry = 27.
```

```r
# Predict using the training data
trainPred <- predict(modFit,trainFinal)
```

```
## 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:dplyr':
##
##     combine

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

```r
confusionMatrix(trainPred,trainFinal$classe)
```

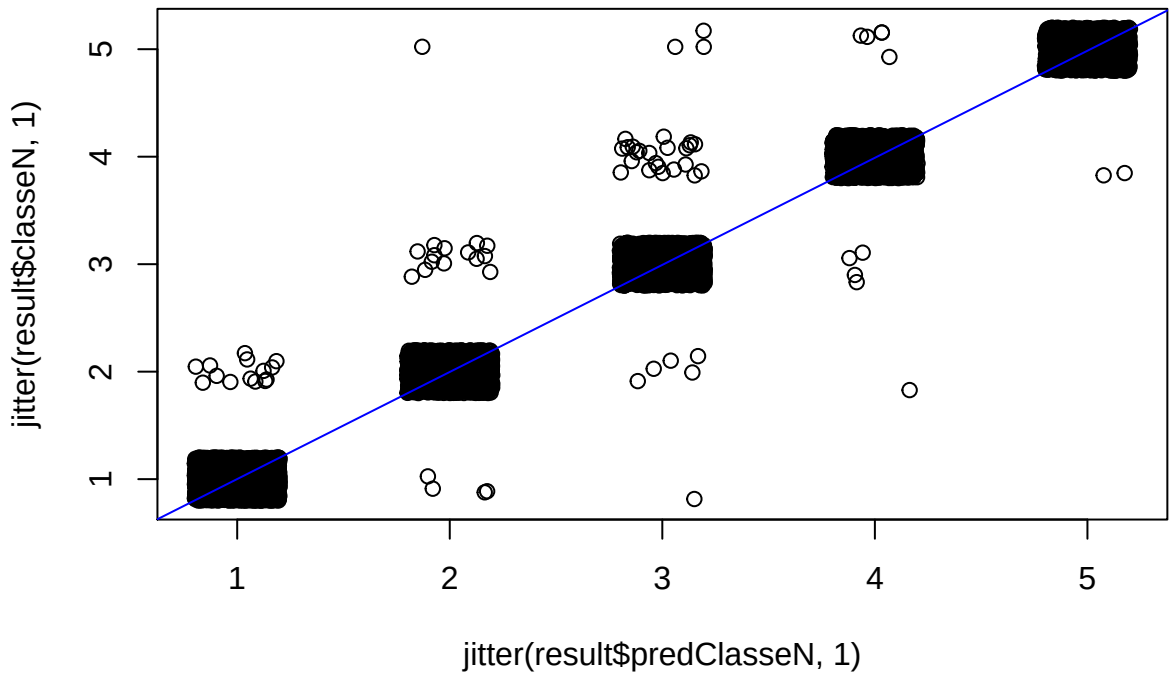
```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 3348    0    0    0    0
##          B    0 2279    0    0    0
##          C    0    0 2054    0    0
##          D    0    0    0 1930    0
##          E    0    0    0    0 2165
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9997, 1)
##     No Information Rate : 0.2843
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            1.0000   1.0000   1.0000   1.0000   1.0000
## Specificity            1.0000   1.0000   1.0000   1.0000   1.0000
## Pos Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Neg Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Prevalence             0.2843   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2843   0.1935   0.1744   0.1639   0.1838
## Detection Prevalence   0.2843   0.1935   0.1744   0.1639   0.1838
## Balanced Accuracy      1.0000   1.0000   1.0000   1.0000   1.0000
```

```r
# Cross-Validation using the model
crossValidPred <- predict(modFit,trainCV)
confusionMatrix(crossValidPred,trainCV$classe)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 2227   14    0    0    0
##          B    4 1498   14    0    1
##          C    1    5 1350   23    3
##          D    0    1    4 1261    5
##          E    0    0    0    2 1433
##
## Overall Statistics
##
##                Accuracy : 0.9902
##                  95% CI : (0.9877, 0.9922)
##     No Information Rate : 0.2845
```

```
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9876
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.9978   0.9868   0.9868   0.9806   0.9938
## Specificity            0.9975   0.9970   0.9951   0.9985   0.9997
## Pos Pred Value         0.9938   0.9875   0.9768   0.9921   0.9986
## Neg Pred Value         0.9991   0.9968   0.9972   0.9962   0.9986
## Prevalence             0.2845   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2838   0.1909   0.1721   0.1607   0.1826
## Detection Prevalence   0.2856   0.1933   0.1761   0.1620   0.1829
## Balanced Accuracy      0.9976   0.9919   0.9910   0.9895   0.9967
```

```
# Accuracy is Accuracy : 0.9902
# Let's plot the predicted classe against the actual classe to compare
source("toNum.R")
result <- data.frame(predClasse=as.character(crossValidPred),classe=as.character(trainCV$classe))
result <- result %>% rowwise() %>% mutate(predClasseN=toNum(predClasse))
result <- result %>% rowwise() %>% mutate(classeN=toNum(classe))
plot(jitter(result$predClasseN,1),jitter(result$classeN,1),pch=1)
abline(lm(result$predClasseN~result$classeN),col='blue')
```




```
# Compute correlation between predicted and actual classe values
cor(result$predClasseN,result$classeN)
```

```
## [1] 0.9970786
```

```
# The result is 0.9970786 which shows that the model is doing a good job
# of predicting.  It's good enough to move to the next phase in the project
# and predict using the testing data.
```

4. Predict using the testing data

```
# Now predict the values using testing data set
testPred <- predict(modFit,testClean)
testPred
```

```
##  [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
```

```
# Recombine with original test data set
testFinal <- select(testing,c(user_name,problem_id))
testFinal$classe <- testPred
testFinal
```

```
##    user_name problem_id classe
## 1      pedro          1      B
## 2     jeremy          2      A
## 3     jeremy          3      B
## 4     adelmo          4      A
## 5     eurico          5      A
## 6     jeremy          6      E
## 7     jeremy          7      D
## 8     jeremy          8      B
## 9   carlitos          9      A
## 10   charles         10      A
## 11  carlitos         11      B
## 12    jeremy         12      C
## 13    eurico         13      B
## 14    jeremy         14      A
## 15    jeremy         15      E
## 16    eurico         16      E
## 17     pedro         17      A
## 18  carlitos         18      B
## 19     pedro         19      B
## 20    eurico         20      B
```