Course Project - Machine Learning

Elizabeth M. Trujillo

2022-11-30

BACKGROUND Using devices such as Jawbone Up, Nike FuelBand, and Fitbit it is now possible to collect a large amount of data about personal activity relatively inexpensively. These type of devices are part of the quantified self movement – a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. One thing that people regularly do is quantify how much of a particular activity they do, but they rarely quantify how well they do it. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbell 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).

knitr::opts_chunk$set(echo = TRUE)

#Check and change directory

#Load necessary packages

install.packages("writexl")

#Import training data

#Import testing data library(readxl)

#View the excel sheets View(pml_testing) dim(pml_testing) #There are 20 observations and 159 variables in the Training dataset

View(pml_training) dim(pml_training) #There are 16383 observations and 159 variables in the Testing dataset

Data Reference: "Ugulino, W.; Cardador, D.; Vega, K.; Velloso, E.; Milidiu, R.; Fuks, H. Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements"

Create a report describing how you built your model, how you used cross validation, what you think the expected out of sample error is, #and why you made the choices you did. You will also use your #prediction model to predict 20 different test cases.

#Load necessary packages install.packages("caret")

#Clean Training and Testing datasets by removing missing values

```
testingData <- pml_testing[, colSums(is.na(pml_testing)) == 0]
dim(testingData)
```

```
## [1]  20 126
```

#There are 20 observation and 126 variables

View(testingData)

```
trainingData<- pml_training[, colSums(is.na(pml_training)) == 0]
dim(trainingData)
```

```
## [1] 16383   125
```

#There are 16383 obervations and 125 variables

View(trainingData)

#Setting up the Training dataset for prediction. I split the training dataset into #65% as the train data and 35% as the test data. I can then calculate the sample error #from this split dataset.

```
set.seed(1500)
library(caret)
inTrain <- createDataPartition(trainingData$classe, p = 0.65, list = FALSE)
trainData <- trainingData[inTrain, ]
testData <- trainingData[-inTrain, ]
dim(trainData)
```

```
## [1] 10652   125
```

#There are 10652 observations and 125 variables in the training data

```
dim(testData)
```

```
## [1] 5731  125
```

#There at 5731 observations and 125 variables in the testing data

#After removing missing values above, I will remove any values that are close to zero

```
nearZero <- nearZeroVar(trainData)
trainData <- trainData[, -nearZero]
testData <- testData[, -nearZero]
dim(trainData)
```

```
## [1] 10652    57
```

#There are 10652 observations and 57 variables in the training data

```
dim(testData)
```

```
## [1] 5731   57
```

#There are 5731 observations and 57 variables in the testing data

View(trainData) View(testData)

#Create model based predictions by using Baye's Theorem to identify optimal classifiers. Use #linear disciminant analysis (lda) and Naive Bayes (nb)

#Prediction

View(plda) View(pnd)

```
table(plda,pnd)
```

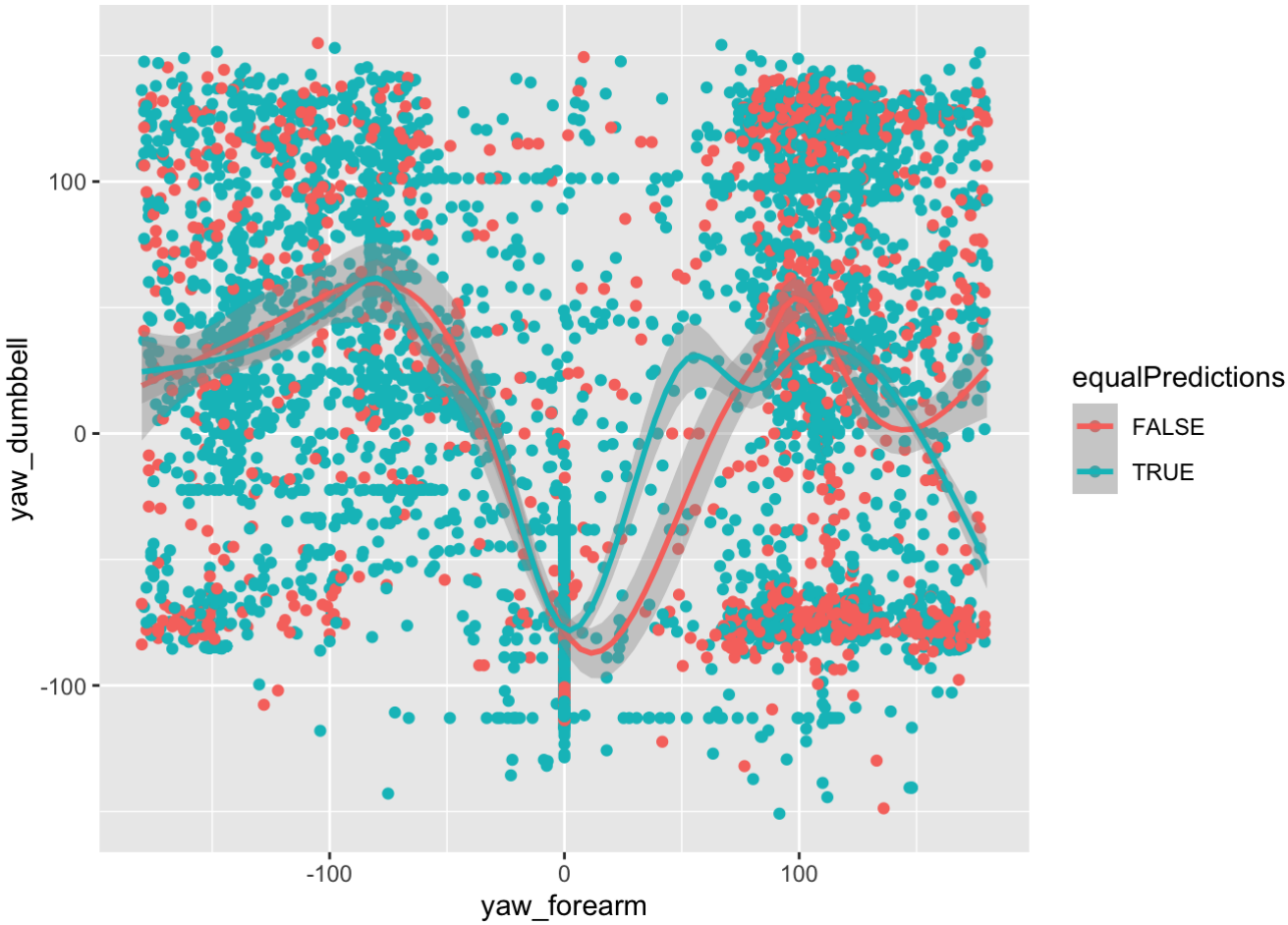
```
##     pnd
## plda    A    B    C    D    E
##    A 1770   96   78  106    0
##    B  115  827  109   75   25
##    C  375   75  827   55    3
##    D  233   14  121  740   16
##    E    0    4    0    2   65
```

#Compare the results using two exercises

```
equalPredictions = (plda == pnd)
qplot(yaw_forearm, yaw_dumbbell, geom=c("point", "smooth"), colour=equalPredictions, dat
a=testData)
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
```

```
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



#Predict with Random Forests

#Create model based predictions by using Random Forests (rf). Use data that was previously cleaned.

```
mod_rf1$finalModel
```

```
##
## Call:
##  randomForest(x = x, y = y, mtry = param$mtry)
##                Type of random forest: classification
##                      Number of trees: 500
## No. of variables tried at each split: 31
##
##         OOB estimate of  error rate: 0.1%
## Confusion matrix:
##      A    B    C    D   E  class.error
## A 3627    0    0    0   0 0.0000000000
## B    2 2466    1    0   0 0.0012150668
## C    0    1 2223    1   0 0.0008988764
## D    0    0    6 2085   0 0.0028694405
## E    0    0    0    0 240 0.0000000000
```

#Determine the accuracy by using the test data

```
library(randomForest)
```

```
## randomForest 4.7-1.1
```

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

```
##
## Attaching package: 'randomForest'
```

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

```
tree2 <- getTree(mod_rf1$finalModel, k=2)
tree2
```

```
##     left daughter right daughter split var   split point status prediction
## 1               2              3        49 -3.305000e+01      1          0
## 2               4              5        40  1.035000e+00      1          0
## 3               6              7        46  4.395000e+02      1          0
## 4               8              9        60  7.760000e+02      1          0
## 5               0              0         0  0.000000e+00     -1          2
## 6              10             11         8  2.415000e+02      1          0
## 7              12             13        38  5.500000e+00      1          0
## 8               0              0         0  0.000000e+00     -1          1
## 9               0              0         0  0.000000e+00     -1          2
## 10             14             15        57 -7.250000e+01      1          0
## 11             16             17        10  1.535000e+01      1          0
## 12             18             19        11 -2.940000e+00      1          0
## 13             20             21         8  8.650000e+01      1          0
## 14             22             23        10  2.575000e+01      1          0
## 15             24             25         8  1.290000e+02      1          0
## 16             26             27         8  2.600000e+02      1          0
## 17             28             29         6  1.322838e+09      1          0
## 18              0              0         0  0.000000e+00     -1          2
## 19             30             31         6  1.322838e+09      1          0
## 20              0              0         0  0.000000e+00     -1          5
## 21             32             33         6  1.322673e+09      1          0
## 22             34             35         6  1.323095e+09      1          0
## 23             36             37         8  6.850000e+01      1          0
## 24             38             39        29  2.660000e+02      1          0
## 25              0              0         0  0.000000e+00     -1          1
## 26              0              0         0  0.000000e+00     -1          2
## 27             40             41        45 -5.535000e+02      1          0
## 28              0              0         0  0.000000e+00     -1          1
## 29             42             43         6  1.323095e+09      1          0
## 30             44             45        42 -1.850000e+01      1          0
## 31              0              0         0  0.000000e+00     -1          4
## 32              0              0         0  0.000000e+00     -1          1
## 33             46             47        11  1.655000e+02      1          0
## 34             48             49         6  1.323095e+09      1          0
## 35              0              0         0  0.000000e+00     -1          4
## 36              0              0         0  0.000000e+00     -1          1
## 37             50             51         8  8.900000e+01      1          0
## 38              0              0         0  0.000000e+00     -1          5
## 39              0              0         0  0.000000e+00     -1          2
## 40             52             53        10 -4.325000e+01      1          0
## 41             54             55        47 -2.650000e+01      1          0
## 42             56             57        45  4.960000e+02      1          0
## 43             58             59         8  8.295000e+02      1          0
## 44              0              0         0  0.000000e+00     -1          3
## 45              0              0         0  0.000000e+00     -1          2
## 46             60             61         9  8.750000e-01      1          0
## 47             62             63         6  1.322833e+09      1          0
## 48             64             65         8  2.450000e+01      1          0
## 49             66             67         8  1.065000e+02      1          0
## 50              0              0         0  0.000000e+00     -1          2
## 51             68             69         6  1.323095e+09      1          0
```

```
## 52          70          71          11  1.700000e+02      1          0
## 53          72          73           8  3.645000e+02      1          0
## 54          74          75           6  1.323084e+09      1          0
## 55          76          77           8  2.785000e+02      1          0
## 56          78          79          49  4.200000e+01      1          0
## 57          80          81          54  8.650000e-01      1          0
## 58           0           0           0  0.000000e+00     -1          4
## 59           0           0           0  0.000000e+00     -1          3
## 60          82          83           8  4.535000e+02      1          0
## 61          84          85           8  2.450000e+02      1          0
## 62           0           0           0  0.000000e+00     -1          2
## 63           0           0           0  0.000000e+00     -1          3
## 64          86          87          27 -1.220000e+00      1          0
## 65          88          89           6  1.323095e+09      1          0
## 66          90          91          57 -1.060000e+02      1          0
## 67           0           0           0  0.000000e+00     -1          4
## 68           0           0           0  0.000000e+00     -1          3
## 69           0           0           0  0.000000e+00     -1          4
## 70          92          93          24 -7.765000e+01      1          0
## 71          94          95          55 -1.420000e+02      1          0
## 72          96          97          20  5.545000e+02      1          0
## 73          98          99           8  5.655000e+02      1          0
## 74         100         101           8  4.365000e+02      1          0
## 75         102         103          36  1.084798e+01      1          0
## 76           0           0           0  0.000000e+00     -1          3
## 77         104         105          35 -9.439044e+01      1          0
## 78           0           0           0  0.000000e+00     -1          2
## 79           0           0           0  0.000000e+00     -1          4
## 80         106         107          10  2.675000e+01      1          0
## 81           0           0           0  0.000000e+00     -1          2
## 82           0           0           0  0.000000e+00     -1          2
## 83         108         109          55 -8.300000e+01      1          0
## 84           0           0           0  0.000000e+00     -1          4
## 85         110         111          37 -6.750715e+01      1          0
## 86           0           0           0  0.000000e+00     -1          5
## 87           0           0           0  0.000000e+00     -1          4
## 88         112         113          18 -8.800000e+01      1          0
## 89           0           0           0  0.000000e+00     -1          2
## 90           0           0           0  0.000000e+00     -1          3
## 91           0           0           0  0.000000e+00     -1          2
## 92         114         115          34  5.940000e+02      1          0
## 93         116         117          18 -1.565000e+02      1          0
## 94           0           0           0  0.000000e+00     -1          1
## 95           0           0           0  0.000000e+00     -1          2
## 96           0           0           0  0.000000e+00     -1          4
## 97         118         119          29  1.110000e+02      1          0
## 98         120         121           6  1.323084e+09      1          0
## 99         122         123           9  1.255000e+02      1          0
## 100          0           0           0  0.000000e+00     -1          1
## 101        124         125          46  3.855000e+02      1          0
## 102        126         127           6  1.323084e+09      1          0
## 103        128         129          11 -9.300000e+01      1          0
```

```
## 104          130          131          42 -1.700000e+01      1          0
## 105          132          133          10 -4.255000e+01      1          0
## 106            0            0           0  0.000000e+00     -1          1
## 107            0            0           0  0.000000e+00     -1          2
## 108            0            0           0  0.000000e+00     -1          4
## 109          134          135           8  6.320000e+02      1          0
## 110          136          137          55 -1.070000e+02      1          0
## 111            0            0           0  0.000000e+00     -1          2
## 112            0            0           0  0.000000e+00     -1          1
## 113          138          139           6  1.323084e+09      1          0
## 114            0            0           0  0.000000e+00     -1          3
## 115            0            0           0  0.000000e+00     -1          2
## 116            0            0           0  0.000000e+00     -1          2
## 117            0            0           0  0.000000e+00     -1          3
## 118            0            0           0  0.000000e+00     -1          2
## 119            0            0           0  0.000000e+00     -1          1
## 120          140          141           6  1.322833e+09      1          0
## 121            0            0           0  0.000000e+00     -1          4
## 122          142          143           8  6.125000e+02      1          0
## 123          144          145           8  6.125000e+02      1          0
## 124          146          147           9  1.230000e+02      1          0
## 125          148          149           6  1.322838e+09      1          0
## 126            0            0           0  0.000000e+00     -1          2
## 127            0            0           0  0.000000e+00     -1          4
## 128            0            0           0  0.000000e+00     -1          4
## 129            0            0           0  0.000000e+00     -1          2
## 130          150          151          10  1.495000e+01      1          0
## 131          152          153          15 -9.000000e-02      1          0
## 132          154          155           9  1.225000e+02      1          0
## 133          156          157          55 -9.950000e+01      1          0
## 134            0            0           0  0.000000e+00     -1          3
## 135            0            0           0  0.000000e+00     -1          2
## 136            0            0           0  0.000000e+00     -1          4
## 137            0            0           0  0.000000e+00     -1          2
## 138            0            0           0  0.000000e+00     -1          2
## 139            0            0           0  0.000000e+00     -1          5
## 140          158          159          57 -1.230000e+02      1          0
## 141          160          161          11 -9.355000e+01      1          0
## 142            0            0           0  0.000000e+00     -1          4
## 143          162          163          34  5.925000e+02      1          0
## 144            0            0           0  0.000000e+00     -1          4
## 145            0            0           0  0.000000e+00     -1          3
## 146          164          165          36  6.018112e+01      1          0
## 147            0            0           0  0.000000e+00     -1          2
## 148            0            0           0  0.000000e+00     -1          2
## 149            0            0           0  0.000000e+00     -1          3
## 150          166          167          15 -2.200000e-01      1          0
## 151            0            0           0  0.000000e+00     -1          4
## 152          168          169          31 -1.440000e+02      1          0
## 153          170          171          11 -8.825000e+01      1          0
## 154          172          173           8  6.405000e+02      1          0
## 155          174          175           6  1.322833e+09      1          0
```

```
## 156          176          177           6  1.322673e+09      1          0
## 157          178          179           8  3.135000e+02      1          0
## 158          180          181          36 -5.045274e+01      1          0
## 159            0            0           0  0.000000e+00     -1          2
## 160            0            0           0  0.000000e+00     -1          4
## 161          182          183          16 -2.350000e+01      1          0
## 162          184          185          30  1.115000e+02      1          0
## 163          186          187           9  6.810000e+01      1          0
## 164            0            0           0  0.000000e+00     -1          1
## 165            0            0           0  0.000000e+00     -1          4
## 166          188          189          31 -3.450000e+01      1          0
## 167          190          191          11  4.390000e+00      1          0
## 168            0            0           0  0.000000e+00     -1          3
## 169            0            0           0  0.000000e+00     -1          4
## 170          192          193          36 -1.906337e+00      1          0
## 171            0            0           0  0.000000e+00     -1          1
## 172          194          195          23  3.010000e+01      1          0
## 173            0            0           0  0.000000e+00     -1          1
## 174          196          197           6  1.322833e+09      1          0
## 175          198          199          40 -2.400000e-01      1          0
## 176          200          201           6  1.322673e+09      1          0
## 177          202          203           8  3.760000e+02      1          0
## 178            0            0           0  0.000000e+00     -1          4
## 179          204          205           6  1.322673e+09      1          0
## 180            0            0           0  0.000000e+00     -1          3
## 181            0            0           0  0.000000e+00     -1          1
## 182            0            0           0  0.000000e+00     -1          4
## 183            0            0           0  0.000000e+00     -1          3
## 184          206          207          39 -1.950000e-01      1          0
## 185          208          209          51  2.450000e+01      1          0
## 186            0            0           0  0.000000e+00     -1          4
## 187            0            0           0  0.000000e+00     -1          1
## 188            0            0           0  0.000000e+00     -1          4
## 189            0            0           0  0.000000e+00     -1          3
## 190          210          211          15  4.000000e-02      1          0
## 191            0            0           0  0.000000e+00     -1          4
## 192            0            0           0  0.000000e+00     -1          2
## 193            0            0           0  0.000000e+00     -1          3
## 194            0            0           0  0.000000e+00     -1          4
## 195            0            0           0  0.000000e+00     -1          2
## 196            0            0           0  0.000000e+00     -1          1
## 197            0            0           0  0.000000e+00     -1          2
## 198            0            0           0  0.000000e+00     -1          3
## 199            0            0           0  0.000000e+00     -1          4
## 200          212          213          59  6.585000e+02      1          0
## 201            0            0           0  0.000000e+00     -1          3
## 202          214          215           6  1.323084e+09      1          0
## 203          216          217           9  1.255000e+02      1          0
## 204          218          219          10  7.000000e-02      1          0
## 205          220          221          42 -4.550000e+01      1          0
## 206            0            0           0  0.000000e+00     -1          2
## 207          222          223           7  8.883515e+05      1          0
```

```
## 208   0   0   0  0.000000e+00   -1  3
## 209   0   0   0  0.000000e+00   -1  2
## 210   0   0   0  0.000000e+00   -1  3
## 211   0   0   0  0.000000e+00   -1  2
## 212 224 225  56 -1.545000e+02    1  0
## 213 226 227  11 -8.900000e+01    1  0
## 214   0   0   0  0.000000e+00   -1  1
## 215   0   0   0  0.000000e+00   -1  2
## 216 228 229  34 -3.715000e+02    1  0
## 217 230 231  33  2.095000e+02    1  0
## 218   0   0   0  0.000000e+00   -1  4
## 219 232 233   6  1.322673e+09    1  0
## 220 234 235   8  5.650000e+02    1  0
## 221 236 237  42  3.500000e+00    1  0
## 222   0   0   0  0.000000e+00   -1  3
## 223 238 239  51  3.600000e+01    1  0
## 224 240 241  37 -7.877347e+01    1  0
## 225   0   0   0  0.000000e+00   -1  1
## 226   0   0   0  0.000000e+00   -1  2
## 227   0   0   0  0.000000e+00   -1  4
## 228 242 243  47  3.550000e+01    1  0
## 229 244 245  32 -3.520000e+02    1  0
## 230 246 247  10 -4.180000e+01    1  0
## 231   0   0   0  0.000000e+00   -1  3
## 232 248 249  47  3.375000e+02    1  0
## 233   0   0   0  0.000000e+00   -1  2
## 234 250 251  46  3.785000e+02    1  0
## 235 252 253  33  3.790000e+02    1  0
## 236 254 255   8  3.775000e+02    1  0
## 237 256 257  11 -9.325000e+01    1  0
## 238   0   0   0  0.000000e+00   -1  2
## 239   0   0   0  0.000000e+00   -1  1
## 240   0   0   0  0.000000e+00   -1  1
## 241   0   0   0  0.000000e+00   -1  2
## 242   0   0   0  0.000000e+00   -1  3
## 243 258 259   1  5.000000e-01    1  0
## 244   0   0   0  0.000000e+00   -1  3
## 245 260 261  40 -8.050000e-01    1  0
## 246   0   0   0  0.000000e+00   -1  2
## 247   0   0   0  0.000000e+00   -1  4
## 248 262 263  46  2.505000e+02    1  0
## 249   0   0   0  0.000000e+00   -1  2
## 250 264 265  49  2.135000e+01    1  0
## 251 266 267  49  1.221000e+01    1  0
## 252 268 269   9  7.000000e-01    1  0
## 253   0   0   0  0.000000e+00   -1  3
## 254 270 271   8  3.540000e+02    1  0
## 255 272 273   6  1.322833e+09    1  0
## 256   0   0   0  0.000000e+00   -1  4
## 257   0   0   0  0.000000e+00   -1  2
## 258   0   0   0  0.000000e+00   -1  4
## 259   0   0   0  0.000000e+00   -1  2
```

```
## 260              0              0          0  0.000000e+00     -1           3
## 261            274            275         26 -4.200000e+00      1           0
## 262            276            277         45 -4.485000e+02      1           0
## 263            278            279         53  4.135000e+00      1           0
## 264              0              0          0  0.000000e+00     -1           3
## 265            280            281          8  4.765000e+02      1           0
## 266              0              0          0  0.000000e+00     -1           3
## 267              0              0          0  0.000000e+00     -1           4
## 268              0              0          0  0.000000e+00     -1           3
## 269            282            283         19  1.655000e+02      1           0
## 270              0              0          0  0.000000e+00     -1           1
## 271            284            285          7  6.803120e+05      1           0
## 272            286            287         46  3.360000e+02      1           0
## 273            288            289         27  9.800000e-01      1           0
## 274              0              0          0  0.000000e+00     -1           3
## 275            290            291         20  5.945000e+02      1           0
## 276            292            293         39  3.600000e-01      1           0
## 277              0              0          0  0.000000e+00     -1           1
## 278              0              0          0  0.000000e+00     -1           1
## 279              0              0          0  0.000000e+00     -1           2
## 280              0              0          0  0.000000e+00     -1           3
## 281              0              0          0  0.000000e+00     -1           4
## 282              0              0          0  0.000000e+00     -1           4
## 283              0              0          0  0.000000e+00     -1           3
## 284              0              0          0  0.000000e+00     -1           3
## 285              0              0          0  0.000000e+00     -1           2
## 286              0              0          0  0.000000e+00     -1           3
## 287            294            295         33  1.500000e+02      1           0
## 288            296            297         43 -3.650000e+01      1           0
## 289              0              0          0  0.000000e+00     -1           2
## 290            298            299         29 -1.905000e+02      1           0
## 291            300            301         29 -2.965000e+02      1           0
## 292              0              0          0  0.000000e+00     -1           3
## 293            302            303         47  2.480000e+02      1           0
## 294              0              0          0  0.000000e+00     -1           3
## 295              0              0          0  0.000000e+00     -1           1
## 296              0              0          0  0.000000e+00     -1           3
## 297              0              0          0  0.000000e+00     -1           4
## 298              0              0          0  0.000000e+00     -1           4
## 299              0              0          0  0.000000e+00     -1           3
## 300              0              0          0  0.000000e+00     -1           3
## 301              0              0          0  0.000000e+00     -1           4
## 302              0              0          0  0.000000e+00     -1           4
## 303              0              0          0  0.000000e+00     -1           2
```

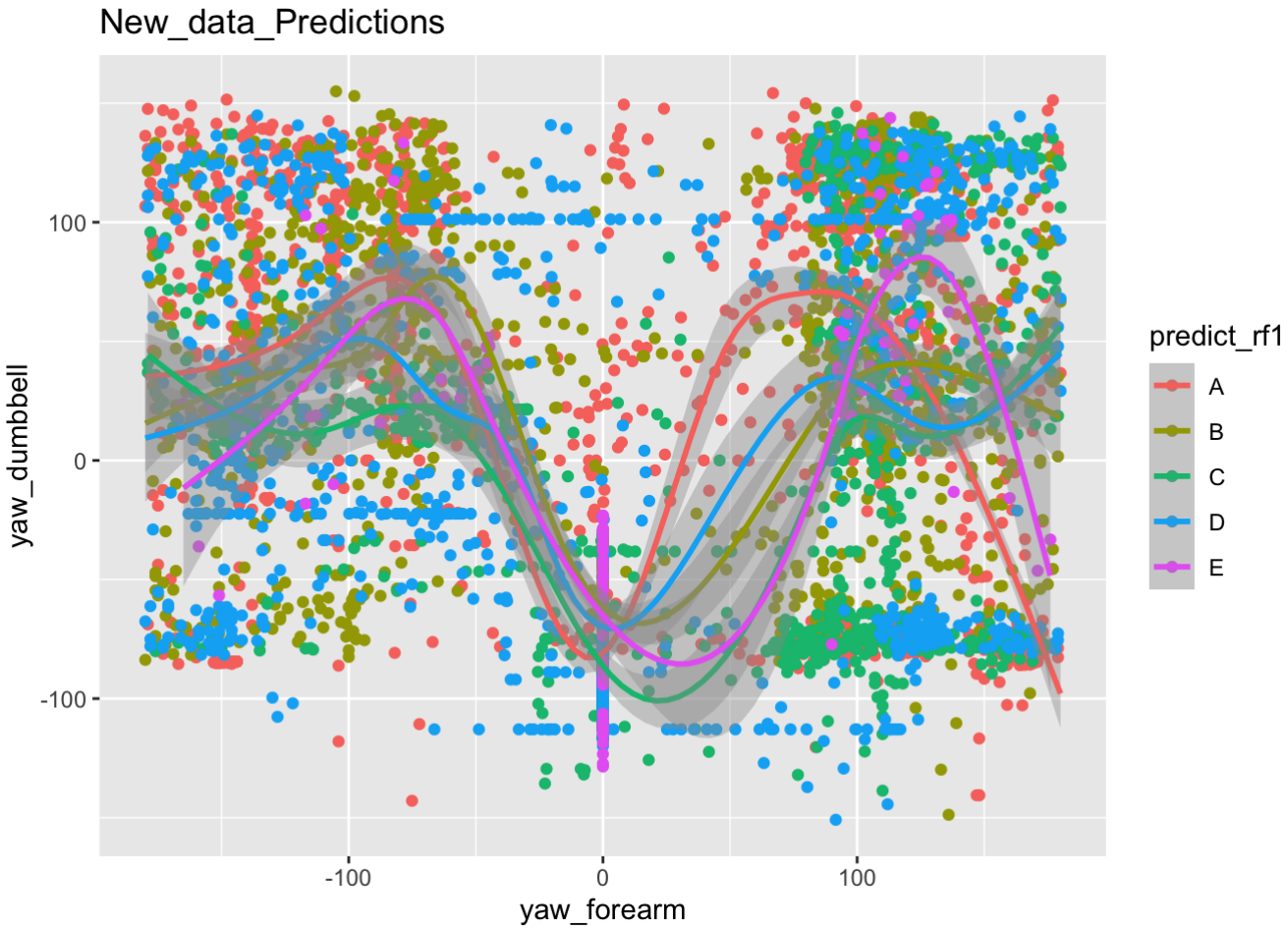
#Predict new values

```
table(predict_rf1, testData$classe)
```

```
##
## predict_rf1    A    B    C    D    E
##           A 1953    0    0    0    0
##           B    0 1327    4    0    0
##           C    0    1 1193    1    0
##           D    0    0    0 1124    0
##           E    0    0    0    0  128
```

```
qplot(yaw_forearm, yaw_dumbbell, geom=c("point", "smooth"), colour=predict_rf1, data=tes
tData, main= "New_data_Predictions")
```

```
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



#Validate the data from random forest and use prediction model for the Final quiz Results <- predict(mod_rf1, newdata=testData) Results