Fitness Behaviors

Ivan Fernandez

April 20, 2016

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.

Set directory and load data for both trainning and testing

setwd("C:/Users/ivan/desktop/PML")

LOAD DATA

```
traind <- read.csv("C:/USERS/ivan/desktop/PML/pmltraining.csv")
dim(traind)
```

```
## [1] 19622   160
```

```
testd <- read.csv("C:/USERS/ivan/desktop/PML/pmltesting.csv")
dim(testd)
```

```
## [1]  20 160
```

DATA CLEANUP

Remove NA values from Train data

```
na <- apply(traind, 2, function(x) sum(x %in% c(NA, "")))
na
```

```
##                        X                user_name     raw_timestamp_part_1
##                        0                        0                        0
##     raw_timestamp_part_2           cvtd_timestamp               new_window
##                        0                        0                        0
##               num_window                roll_belt               pitch_belt
##                        0                        0                        0
##                yaw_belt         total_accel_belt       kurtosis_roll_belt
##                        0                        0                    19216
##      kurtosis_picth_belt        kurtosis_yaw_belt       skewness_roll_belt
##                    19216                    19216                    19216
##     skewness_roll_belt.1        skewness_yaw_belt            max_roll_belt
##                    19216                    19216                    19216
##           max_picth_belt             max_yaw_belt            min_roll_belt
```

```
##                   19216                   19216                   19216
##           min_pitch_belt            min_yaw_belt      amplitude_roll_belt
##                   19216                   19216                   19216
##     amplitude_pitch_belt       amplitude_yaw_belt     var_total_accel_belt
##                   19216                   19216                   19216
##            avg_roll_belt         stddev_roll_belt            var_roll_belt
##                   19216                   19216                   19216
##           avg_pitch_belt        stddev_pitch_belt           var_pitch_belt
##                   19216                   19216                   19216
##             avg_yaw_belt          stddev_yaw_belt             var_yaw_belt
##                   19216                   19216                   19216
##             gyros_belt_x            gyros_belt_y            gyros_belt_z
##                       0                       0                       0
##             accel_belt_x            accel_belt_y            accel_belt_z
##                       0                       0                       0
##            magnet_belt_x           magnet_belt_y           magnet_belt_z
##                       0                       0                       0
##                roll_arm               pitch_arm                 yaw_arm
##                       0                       0                       0
##          total_accel_arm           var_accel_arm            avg_roll_arm
##                       0                   19216                   19216
##          stddev_roll_arm            var_roll_arm           avg_pitch_arm
##                   19216                   19216                   19216
##         stddev_pitch_arm           var_pitch_arm             avg_yaw_arm
##                   19216                   19216                   19216
##           stddev_yaw_arm             var_yaw_arm             gyros_arm_x
##                   19216                   19216                       0
##              gyros_arm_y             gyros_arm_z             accel_arm_x
##                       0                       0                       0
##              accel_arm_y             accel_arm_z            magnet_arm_x
##                       0                       0                       0
##             magnet_arm_y            magnet_arm_z       kurtosis_roll_arm
##                       0                       0                   19216
##       kurtosis_picth_arm        kurtosis_yaw_arm       skewness_roll_arm
##                   19216                   19216                   19216
##       skewness_pitch_arm        skewness_yaw_arm            max_roll_arm
##                   19216                   19216                   19216
##            max_picth_arm             max_yaw_arm            min_roll_arm
##                   19216                   19216                   19216
##            min_pitch_arm             min_yaw_arm      amplitude_roll_arm
##                   19216                   19216                   19216
##      amplitude_pitch_arm       amplitude_yaw_arm           roll_dumbbell
##                   19216                   19216                       0
##           pitch_dumbbell             yaw_dumbbell   kurtosis_roll_dumbbell
##                       0                       0                   19216
##  kurtosis_picth_dumbbell   kurtosis_yaw_dumbbell   skewness_roll_dumbbell
##                   19216                   19216                   19216
##  skewness_pitch_dumbbell   skewness_yaw_dumbbell       max_roll_dumbbell
##                   19216                   19216                   19216
##       max_picth_dumbbell        max_yaw_dumbbell       min_roll_dumbbell
##                   19216                   19216                   19216
##       min_pitch_dumbbell        min_yaw_dumbbell  amplitude_roll_dumbbell
##                   19216                   19216                   19216
## amplitude_pitch_dumbbell  amplitude_yaw_dumbbell     total_accel_dumbbell
```

```
##                    19216                    19216                        0
##       var_accel_dumbbell        avg_roll_dumbbell     stddev_roll_dumbbell
##                    19216                    19216                    19216
##        var_roll_dumbbell       avg_pitch_dumbbell    stddev_pitch_dumbbell
##                    19216                    19216                    19216
##       var_pitch_dumbbell         avg_yaw_dumbbell      stddev_yaw_dumbbell
##                    19216                    19216                    19216
##         var_yaw_dumbbell         gyros_dumbbell_x         gyros_dumbbell_y
##                    19216                        0                        0
##         gyros_dumbbell_z         accel_dumbbell_x         accel_dumbbell_y
##                        0                        0                        0
##         accel_dumbbell_z        magnet_dumbbell_x        magnet_dumbbell_y
##                        0                        0                        0
##        magnet_dumbbell_z             roll_forearm            pitch_forearm
##                        0                        0                        0
##              yaw_forearm    kurtosis_roll_forearm   kurtosis_picth_forearm
##                        0                    19216                    19216
##     kurtosis_yaw_forearm    skewness_roll_forearm   skewness_pitch_forearm
##                    19216                    19216                    19216
##     skewness_yaw_forearm         max_roll_forearm        max_picth_forearm
##                    19216                    19216                    19216
##          max_yaw_forearm         min_roll_forearm        min_pitch_forearm
##                    19216                    19216                    19216
##          min_yaw_forearm   amplitude_roll_forearm  amplitude_pitch_forearm
##                    19216                    19216                    19216
##    amplitude_yaw_forearm      total_accel_forearm        var_accel_forearm
##                    19216                        0                    19216
##         avg_roll_forearm      stddev_roll_forearm         var_roll_forearm
##                    19216                    19216                    19216
##        avg_pitch_forearm     stddev_pitch_forearm        var_pitch_forearm
##                    19216                    19216                    19216
##          avg_yaw_forearm       stddev_yaw_forearm          var_yaw_forearm
##                    19216                    19216                    19216
##          gyros_forearm_x          gyros_forearm_y          gyros_forearm_z
##                        0                        0                        0
##          accel_forearm_x          accel_forearm_y          accel_forearm_z
##                        0                        0                        0
##         magnet_forearm_x         magnet_forearm_y         magnet_forearm_z
##                        0                        0                        0
##                   classe
##                        0
```

Create new data set with cleaned data

```
index <- which(na == 0)
traind <- traind[,index]
traind <- traind[,8:60]
```

Adjust Test data for use with model derived from new Train data

```
testd <- testd[,index]
testd <- testd[,8:59]
```

Packages to be loaded (some may not be required)

```
library(randomForest)
```

```
## randomForest 4.6-12
## Type rfNews() to see new features/changes/bug fixes.
```

```
library(caret)
```

```
## Loading required package: lattice
## Loading required package: ggplot2
##
## Attaching package: 'ggplot2'
##
## The following object is masked from 'package:randomForest':
##
##     margin
```

```
library(rpart.plot)
```

```
## Loading required package: rpart
```

```
library(rpart)
```

Model creation

Matrix using randomForest on our Train data set

```
model <- randomForest(classe~., data = traind)
pred <- predict(model, traind)
confusionMatrix(traind$classe, pred)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 5580    0    0    0    0
##          B    0 3797    0    0    0
##          C    0    0 3422    0    0
##          D    0    0    0 3216    0
##          E    0    0    0    0 3607
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9998, 1)
##     No Information Rate : 0.2844
##     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             0.2844   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2844   0.1935   0.1744   0.1639   0.1838
## Detection Prevalence   0.2844   0.1935   0.1744   0.1639   0.1838
## Balanced Accuracy      1.0000   1.0000   1.0000   1.0000   1.0000
```

As expected, the model fits extremely well our train data as this is derived from the very same data set

Model Validation on Test data

```
testd$classe <- factor(nrow(testd))
levels(testd$classe) <- levels(traind$classe)
T1 <- rbind(traind[1,], testd)
T1 <- T1[2:21,]
```

```
pvalidation <- predict(model, traind)
print(confusionMatrix(pvalidation, traind$classe))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 5580    0    0    0    0
##          B    0 3797    0    0    0
##          C    0    0 3422    0    0
##          D    0    0    0 3216    0
##          E    0    0    0    0 3607
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9998, 1)
##     No Information Rate : 0.2844
##     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             0.2844   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2844   0.1935   0.1744   0.1639   0.1838
## Detection Prevalence   0.2844   0.1935   0.1744   0.1639   0.1838
## Balanced Accuracy      1.0000   1.0000   1.0000   1.0000   1.0000
```

The train data model fits quite well our test data with a cross validation accuracy of 99.5% and the out-of-sample error at 0.5%

Proposed Predective Model

```
TestModel <- predict(model, T1)
TestModel
```

```
##  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21
##  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
```

Sources

Training data : https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv

Test data: https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv

Source http://groupware.les.inf.puc-rio.br/har.

More @ http://groupware.les.inf.puc-rio.br/har#ixzz3TROgwbfY

Velloso, E.; Bulling, A.; Gellersen, H.; Ugulino, W.; Fuks, H. Qualitative Activity Recognition of Weight Lifting Exercises. Proceedings of 4th International Conference in Cooperation with SIGCHI (Augmented Human '13) . Stuttgart, Germany: ACM SIGCHI, 2013.