

Machine Learning Course Project

Coursera Student A

15 February 2015

Summary

Using devices such as Jawbone Up, Nike FuelBand, and Fitbit it is now possible to collect a large amount of data about personal activity relatively inexpensively. These type of devices are part of the quantified self movement – a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. One thing that people regularly do is quantify how much of a particular activity they do, but they rarely quantify how well they do it. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbbell of 6 participants. They were asked to perform barbell lifts correctly and incorrectly in 5 different ways. More information is available from the website here: <http://groupware.les.inf.puc-rio.br/har> (see the section on the Weight Lifting Exercise Dataset).

Data Analysis

```
require(caret)
```

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

```
require(ggplot2)
require(randomforest)
```

```
## Loading required package: randomforest
```

```
## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,
## logical.return = TRUE, : there is no package called 'randomforest'
```

```
require(stargazer)
```

```
## Loading required package: stargazer
##
```

```
## Please cite as:
##
```

```
## Hlavac, Marek (2014). stargazer: LaTeX code and ASCII text for well-formatted regression and summary
## R package version 5.1. http://CRAN.R-project.org/package=stargazer
```

```
#install.packages("stargazer")
#library(stargazer)
```

#The training data for this project are available here:

```
trainingURL<-"http://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"
testURL<-"http://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"
training<-read.csv(trainingURL,na.strings=c("NA",""))
test<-read.csv(testURL,na.strings=c("NA",""))

?read.csv
#Clean data
clean_train<-apply(!is.na(training),2,sum)==19622
training2<-training[,clean_train]
testing2<-test[,clean_train]

test<-test[,clean_train]
stargazer(test[1:4],summary=FALSE)
```

```
##
## % Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
## % Date and time: Tue, Feb 17, 2015 - 09:39:19
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}} ccccc}
## \hline
## \hline \hline
##   & X & user\_name & raw\_timestamp\_part\_1 & raw\_timestamp\_part\_2 & \\
## \hline \hline
## 1 & $1$ & pedro & $1,323,095,002$ & $868,349$ & \\
## 2 & $2$ & jeremy & $1,322,673,067$ & $778,725$ & \\
## 3 & $3$ & jeremy & $1,322,673,075$ & $342,967$ & \\
## 4 & $4$ & adelmo & $1,322,832,789$ & $560,311$ & \\
## 5 & $5$ & eurico & $1,322,489,635$ & $814,776$ & \\
## 6 & $6$ & jeremy & $1,322,673,149$ & $510,661$ & \\
## 7 & $7$ & jeremy & $1,322,673,128$ & $766,645$ & \\
## 8 & $8$ & jeremy & $1,322,673,076$ & $54,671$ & \\
## 9 & $9$ & carlitos & $1,323,084,240$ & $916,313$ & \\
## 10 & $10$ & charles & $1,322,837,822$ & $384,285$ & \\
## 11 & $11$ & carlitos & $1,323,084,277$ & $36,553$ & \\
## 12 & $12$ & jeremy & $1,322,673,101$ & $442,731$ & \\
## 13 & $13$ & eurico & $1,322,489,661$ & $298,656$ & \\
## 14 & $14$ & jeremy & $1,322,673,043$ & $178,652$ & \\
## 15 & $15$ & jeremy & $1,322,673,156$ & $550,750$ & \\
## 16 & $16$ & eurico & $1,322,489,713$ & $706,637$ & \\
## 17 & $17$ & pedro & $1,323,094,971$ & $920,315$ & \\
## 18 & $18$ & carlitos & $1,323,084,285$ & $176,314$ & \\
## 19 & $19$ & pedro & $1,323,094,999$ & $828,379$ & \\
## 20 & $20$ & eurico & $1,322,489,658$ & $106,658$ & \\
## \hline \hline
## \end{tabular}
## \end{table}
```

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.