Activity tracking

Rona1d

august 6, 2016

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.2.5
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.2.5
```

Introduction

*This report was written in fulfillment of the assignment of week 4 of the 'Machine Learning' course within the Coursera specialization track 'Data Science'. Goal of this assignment is to investigate activity {lorem ipsum}.

The main questions we want to answer is: "Can we predict which activity is performed based on accellorometer data?"

Executive summary

Downloading and exploring data

For this research we will be using data on exercices provided through the Coursera course website. The data is devided in a training and test set.

```
# First, make sure a folder named 'data' exists

if(!file.exists("data")) {
        dir.create("data")
}
setwd("./data")

# Downloading file...

fileURL1 <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"
fileURL2 <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"

# You need this piece of code for knitting Rmd doc: setInternet2(use = TRUE)

if(!file.exists("HAT_traindata.csv")) {
        download.file(fileURL1, "./HAT_traindata.csv")
}
if(!file.exists("HAT_testdata.csv")) {
        download.file(fileURL2, "./HAT_testdata.csv")
```

```
}

# A quick preliminary scan of the data reveiled variables with a lot of blanks.
# It is therefore wise to name the blanks as 'NA' in the read.csv step

trainset <- read.csv("HAT_traindata.csv", na.strings=c("","  ","NA"))
valset <- read.csv("HAT_testdata.csv")
```