

Caso 3

Luis Cabrera

14 de junio de 2016

Caso: Reconocimiento de Actividad Humana con Celulares

Introducción

Compañías como Fitbit, Nike y Jawbone Up compiten para desarrollar los algoritmos más avanzados para atraer a nuevos usuarios. Los datos de este caso provienen de los acelerómetros ubicados en celulares Samsung Galaxy S.

Una mayor descripción de la base de datos y su recopilación puede ser hallada en la siguiente liga:

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

El propósito de este proyecto es demostrar mi habilidad para recolectar, trabajar y limpiar base de datos. El objetivo es preparar un conjunto ordenado de información que pueda ser trabajado en análisis posteriores.

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Archivos que debe contener

```
"README.md"
"CodeBook.md"
"correr_analisis.R"
```

Función

```
"correr_analisis.R"
```

1. Une los datos de test con los de training, para crear un solo conjunto de datos. .

```
xt1 <- read.table("./Dataset/train/X_train.txt")
yt1 <- read.table("./Dataset/train/y_train.txt")
st1 <- read.table("./Dataset/train/subject_train.txt")

xt2 <- read.table("./Dataset/test/X_test.txt")
yt2 <- read.table("./Dataset/test/y_test.txt")
st2 <- read.table("./Dataset/test/subject_test.txt")

xt3 <- rbind(xt1, xt2)
yt3 <- rbind(yt1, yt2)
st3 <- rbind(st1, st2)
```

- Borra las variables que ya no son requeridas.

```
rm(xt1,yt1,st1)
rm(xt2,yt2,st2)
```

2. Extrae únicamente las medidas de media y desviación estándar de cada medición.

```
caracteristicas <- read.table("./Dataset/features.txt")
mds <- grep("mean\\(\\)|std\\(\\)", caracteristicas[, 2])
xt3 <- xt3[, mds]
```

4. Coloca etiquetas apropiadas en la base de datos con nombres de variables que las describan.(Antes del paso 3 realiamos el 4.)

```
names(xt3) <- gsub("\\(\\)", "", caracteristicas[mds, 2])
names(xt3) <- gsub("mean", "Mean", names(xt3))
names(xt3) <- gsub("std", "Std", names(xt3))
names(xt3) <- gsub("t", "Time", names(xt3))
names(xt3) <- gsub("f", "Frequency", names(xt3))
names(xt3) <- gsub("\\(\\)", "", names(xt3))
names(xt3) <- gsub("-", "", names(xt3))
names(xt3) <- gsub("BodyBody", "Body", names(xt3))
names(xt3) <- gsub("^", "MeanOf", names(xt3))
```

3. Usa nombres de actividad para describir los nombres de las actividades en la base de datos.

```
actividades_train <- read.table("./Dataset/train/y_train.txt")
actividades_test <- read.table("./Dataset/test/y_test.txt")
actividades <- rbind(actividades_train,actividades_test)[,1]
labels <- c("WALKING", "WALKING_UPSTAIRS", "WALKING_DOWNSTAIRS", "SITTING", "STANDING", "LAYING")
actividades <- labels[actividades]
xt3 <- cbind(Actividad = actividades,xt3)
```

5. Con los datos del paso 4, crea una segunda base de datos independiente con el promedio de cada variable para cada actividad y cada sujeto.

```
voluntarios_1 <- read.table("./Dataset/train/subject_train.txt")
voluntarios_2 <- read.table("./Dataset/test/subject_test.txt")
voluntarios <- rbind(voluntarios_1,voluntarios_2)[,1]
xt3 <- cbind(Voluntario = voluntarios,xt3)
```

- Se requiere del paquete dplyr ya instalado.

```
library(dplyr)
```

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

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##   filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
prom_xt3 <- xt3 %>% group_by(Voluntario,Actividad) %>% summarise_each(funs(mean))
```

- Y creamos un archivo llamado “tidy_data_set.txt” donde aparecerá la base de datos ordenada.

```
write.table(prom_xt3,row.name = FALSE,file = "tidy_data_set.txt")
```