

Desafio Opcional: Introdução ao projeto alternativo

Imagine uma pulseira capaz de transmitir dados em forma de coordenadas (X,Y, Z) gerados por um acelerômetro acoplado. Haverá uma central capaz de interpretar estes dados como certos movimentos (imagine o quanto esta solução será útil para pessoas idosas morando sozinhas). Quando houver algum movimento brusco (uma queda), um alarme seria disparado.

Neste projeto você montará um modelo preditivo associado à descoberta de um tipo de movimento a partir de dados gerados por um acelerômetro. Veja os tipos possíveis:

- Escovar os dentes
- Subir escadas
- Descer escadas
- Pentear cabelo
- Beber um copo d'água
- Comer uma refeição
- Tomar sopa
- Deitar-se na cama
- Levantar-se da cama
- Derramar água
- Sentar-se em uma cadeira
- Levantar-se de uma cadeira
- Utilizar o telefone
- Caminhar

Conheça um pouco mais os experimentos inspiradores deste projeto:

<https://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer>
(<https://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer>).

E não deixe de conferir estes!

<https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer>
(<https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer>).
<http://mldata.org/repository/data/viewslug/wearable-accelerometers-activity/>.
(<http://mldata.org/repository/data/viewslug/wearable-accelerometers-activity/>).

Coleta

Os dados serão disponibilizados em um arquivo compactado, **HMP_Dataset.zip**, que você pode baixar [aqui](https://s3.amazonaws.com/caelum-online-public/731-pipeline-big-data/01/arquivos/HMP_Dataset.zip)
(https://s3.amazonaws.com/caelum-online-public/731-pipeline-big-data/01/arquivos/HMP_Dataset.zip). Uma vez destrinchado, uma série de diretórios será criada:

Este Computador > LENOVO (D:) > HMP_Dataset

Nome	Data de modificação	Tipo	Tam
Brush_teeth	18/09/2017 14:44	Pasta de arquivos	
Climb_stairs	18/09/2017 13:16	Pasta de arquivos	
Comb_hair	18/09/2017 13:16	Pasta de arquivos	
Descend_stairs	18/09/2017 13:16	Pasta de arquivos	
Drink_glass	18/09/2017 13:17	Pasta de arquivos	
Eat_meat	18/09/2017 13:17	Pasta de arquivos	
Eat_soup	18/09/2017 13:17	Pasta de arquivos	
Getup_bed	18/09/2017 13:17	Pasta de arquivos	
Liedown_bed	18/09/2017 13:17	Pasta de arquivos	
MODELS	18/09/2017 13:17	Pasta de arquivos	
Pour_water	18/09/2017 13:17	Pasta de arquivos	
Sitdown_chair	18/09/2017 13:17	Pasta de arquivos	
Standup_chair	18/09/2017 13:17	Pasta de arquivos	
Use_telephone	18/09/2017 13:17	Pasta de arquivos	
Walk	18/09/2017 13:17	Pasta de arquivos	

Cada diretório armazena diversos arquivos coletados a partir de um determinado movimento. Por exemplo, vejamos a relação disponível para escovar os dentes:

Este Computador > LENOVO (D:) > HMP_Dataset > Brush_teeth

Nome	Data de modificação	Tipo	Tamanho
Accelerometer-2011-04-11-13-28-18-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	22 KB
Accelerometer-2011-04-11-13-29-54-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	9 KB
Accelerometer-2011-05-30-08-35-11-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	22 KB
Accelerometer-2011-05-30-09-36-50-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	22 KB
Accelerometer-2011-05-30-10-34-16-brush_teeth-m1.txt	15/09/2017 09:38	Documento de Texto	32 KB
Accelerometer-2011-05-30-21-10-57-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	28 KB
Accelerometer-2011-05-30-21-55-04-brush_teeth-m2.txt	15/09/2017 09:38	Documento de Texto	30 KB
Accelerometer-2011-05-31-15-16-47-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	25 KB
Accelerometer-2011-06-02-10-42-22-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	25 KB
Accelerometer-2011-06-02-10-45-50-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	29 KB
Accelerometer-2011-06-06-10-45-27-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	31 KB
Accelerometer-2011-06-06-10-48-05-brush_teeth-f1.txt	15/09/2017 09:38	Documento de Texto	23 KB

Perceba a convenção utilizada para denominar cada arquivo:

- "Accelerometer-"
- Data e hora na qual aconteceu a coleta (AAA-MM-DD-HH-MI-SS)
- Denominação do movimento (exemplo: **brush_teeth**)
- Gênero (**m** ou **f**)
- Número do voluntário

Depois de extrair os arquivos, a coleta propriamente dita acontecerá ao executar o seguinte código em R:

```
## Diretório de trabalho
setwd("D:\\HMP_Dataset")

##### 1) Carrega dados
df_medidas <- data.frame()
arqs = list.files(pattern = ".txt$", recursive = TRUE)
for (a in 1:length(arqs))
{
  arquivo = arqs[a]
  # buscamos arquivos em subdiretórios
  if (!(length(i <- grep('/', arquivo))) == 0)
  {
    movimento = unlist(strsplit(arquivo, "/")[1][1])[1]
    # remove subdiretório de arquivo
    arquivo = unlist(strsplit(arquivo, "/"))[2][1]
  }
}
```

```
genero = substr(unlist(strsplit(arquivo, " "))[9], 1, 1)
data <- read.table(arquivo)
data$arquivo <- arquivo
data$movimento <- movimento
data$genero <- genero
df_medidas <- rbind(df_medidas, data)
}
}
colnames(df_medidas) <- c("X", "Y", "Z", "Arquivo", "Movimento", "Genero")
```