

Name: José Carlos Baquero Triguero
Instructor's Name: Jeff Leek
Course: Data Analysis
Date: March 10, 2013

PREDICTION OF MOBILE ACTIVITY

Introduction

A smartphone is a [mobile phone](#) built on a [mobile operating system](#), with more advanced computing capability connectivity than a [feature phone](#)[1]. A lot of smartphones have an accelerometer used, for example, to change the camera orientation when the smartphone is turn. That accelerometer actually measure how much activity is going on in, on three different axes.

This study was performed where that data was collected for a number of individuals and each individual participated in several different types of activities. During those activities, they had the smartphone in their pocket, and the accelerometer measured the types of activity that were going on.

Prediction methods:

Find the right data:

For our analysis we use the data consist of a sample of 7.352 observations from 30 different individuals. This data is available from the UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>) where there are more details of the data set.

The data were downloaded from <https://spark-public.s3.amazonaws.com/dataanalysis/samsungData.rda> on March 1, 2013

Define your error rate:

The variable that we want to predict is the type of activity the user is doing. This variable is qualitative. For this reason the error measures we going to use (and these rates) are:

- Sensitivity > 75%
- Specificity: > 90%
- % False Positive < 5%
- % False Negative < 5%

Split data:

For this analysis we split the data in:

- Training set: the data from subjects 1, 3, 5, and 6
- Test set: the data from subjects 27, 28, 29, and 30

- Validation: we don't use validation data set.

Pick features on the training set:

Our strategic to select the features on the training set is the study of each variable's distribution for each activity and pick out the variables which their distribution don't collide between the different activities

We select the next variables:

- tGravityAcc_mean___Y
- tGravityAcc_mean___X
- angle_Y_gravityMean_
- tBodyAccJerk_energy___X
- fBodyBodyAccJerkMag_mean___
- tBodyAccJerk_max___X
- tBodyAcc_energy___X
- tBodyAccJerk_std___X
- tBodyAccJerk_std___Z

Pick prediction function on the training set:

Using the Test Training data we select the prediction function:

```
randomForest(activity ~ tGravityAcc_mean___Y
               + tGravityAcc_mean___X
               + angle_Y_gravityMean_
               + tBodyAccJerk_energy___X
               + fBodyBodyAccJerkMag_mean___
               + tBodyAccJerk_max___X
               + tBodyAcc_energy___X
               + tBodyAccJerk_std___X
               + tBodyAccJerk_std___Z
               , data = trainingSet, prox = TRUE, ntree = 50)
```

Cross-validate on the training set:

First we split the data in Training Set and Test Set describe above.

After that we run the prediction function to the Test Set, with the next results:

Activity	Sensitivity	Specificity	% False Positive	% False Negative
Laying	100.00%	100.00%	0.00%	0.00%
Sitting	75.47%	97.09%	1.82%	3.96%
Standing	87.10%	95.11%	3.96%	2.13%
Walk	95.83%	92.50%	6.40%	0.60%
Walkdown	82.89%	99.28%	0.60%	2.74%
Walkup	76.60%	100.00%	0.00%	3.96%

Only the False Positive of Walkdown activity doesn't achieve the objective

Reproducibility:

All analyses performed in this manuscript are reproduced in the R markdown file SamsungFinal1.Rmd. To reproduce the exact results presented in this manuscript the cached version of the analysis must be performed, as the data available from <https://spark-public.s3.amazonaws.com/dataanalysis/samsungData.rda>

Results:

In this data Analysis we get a prediction tree, described above, that allow us classify the smartphone activity through the variables selected in the study. We ensure the results because we executed a cross-validation with different data.

Conclusions:

We can classify accurately the activity of the user of one smartphone through the different variables register by the mobile phone. We could use this result to measure the user's activity every day and to know how many calories the user burn.

References:

1. Wikipedia "Smartphone" page. URL: <http://en.wikipedia.org/wiki/Smartphone>. Accessed 03/08/2013.