# HUMAN ACTIVITY RECOGNITION FROM SMARTPHONES DEVICES, USING MACHINE LEARNING MULTICLASS CLASSIFIERS

Daniel Guimaraes, Sebastian Valencia, Carlos Hernandez

MUM – Maharishi University of Management

Fairfield, Iowa, United States, November 2018

## Abstract

Human Activity Recognition use data captured from several technological devices you can use today to help you/assist you in daily activities: fitness, heart and pulse rates, virtual reality, augmented reality and so on. Is so how you can find in the market: smart watches, smart bands, glasses and in general wearable devices. The main inconvenient about this is that this technologies in 2018 is not available to everyone because the costs.

Starting from a paper of 2012 in which author’s captures and process simple smartphones data to classify the human activity using Machine Learning algorithms, but whose major concern was about reducing the computational need of processing and classifying the information because the expensive floating-point calculations and the limited resources from mobile devices, in capacity and power, in that time; we will demonstrate in this paper that is possible to go further in two directions: 1. Applying feature selection – that will reduce in more than 70% the needed inputs for the classifier and 2. Using this reduced set of features still is possible to identify the same 6 human activities plus a new activity: Falling. With this final activity we will show the huge possibilities of real applications that can be developed to help people who suffers accidents or health crises being alone; with the help of common and no expensive mobile devices everybody has today.

**Keywords:** Human Activity Recognition, Smartphones, Feature Reduction, Lying Detection

## Introduction

Considering the high penetration of smartphones today, near to 40% worldwide and more than 65% within north-America[[1]](#footnote-1) we could say that is a device widely used and almost available for most of the people.

By the other hand, researches about human mortality reveals that falling is the second leading cause of accidental death worldwide and is a major cause of injury for people, especially for older adults. About 226 million cases of fall reported in 2.015[[2]](#footnote-2) resulted in 527.000 deaths. Each one of us has somebody to think about this: what about if he/she being alone has an accident? How will I know? Who can help him/her? Would him/her be alone in that moment? What about if happens to me?

In today’s market there are some devices that can help us but, as we mentioned in the abstract: is not available for everybody, because the introduction cost (new technological devices, recently introduced to the market, usually are expensive).

This is where both things can be together to help us: with the common devices we have today -smartphones- that includes inertial sensors like: gyroscope and accelerometer is possible to develop applications that takes that data in real time, to process it, and using Machine Learning algorithms be able to learn about the human behavior and to determine the activity is being performed, we mean: is possible to know if the person is: walking, lying, going upstairs, downstairs, just stand up, sited, or our major interest in this paper: if he/she is falling and could need some help.

In this paper we deal with two topics: 1. Feature reduction: The data captured from those internal sensors is huge (more than 500 measures and with a high frequency of generation), so this produces a high computational cost, and 2. How to be able to identify new Human Activities, starting from this data and the huge capability of exploit this to develop human-assist applications with the use of Machine Learning.

## Related Work

In 2012 Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz, wrote a paper called “Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine” in which the authors major concern was about how to reduce the high computational cost involved in processing the sensors data from smartphones (gyroscope and accelerometer too) in order to reduce the power consumption and processing time of those devices, with the capabilities existing in that year. The authors focused its work in using SVM Machine Learning algorithm but adapting it to do operations with integer values, avoiding so the expensive floating-point operations to be done in those limited capabilities devices. This work finished with the conclusion that is possible to do it, and the authors named this new method with the name “Multiclass Hardware Friendly Support Vector Machine (MC-HF-SVM)”.

However, reading the paper and making some research about his experiment, we noticed they didn’t use the feature reduction that could help them to reduce significative the amount of data to be processed.

Several other experiments have been done but we don’t find a useful and available application to solve the problem exposed in the introduction: How using Machine Learning modern techniques we can help people who suffer an accident, a fall, to be attended on time, with the urgency required with the technology we have available.

## Methodology

**Feature reduction**

In our experiment with feature selection we started with the data available in UCI, with information from 6 subjects (6 different persons) who tied to their waist a common smartphone and recorded the gyroscope and accelerometer measures, doing the 6 different activities: walking, siting, laying, upstairs, downstairs and standing. This information already was classified into that 6 different activities and presented in two different files: a train and a test file.

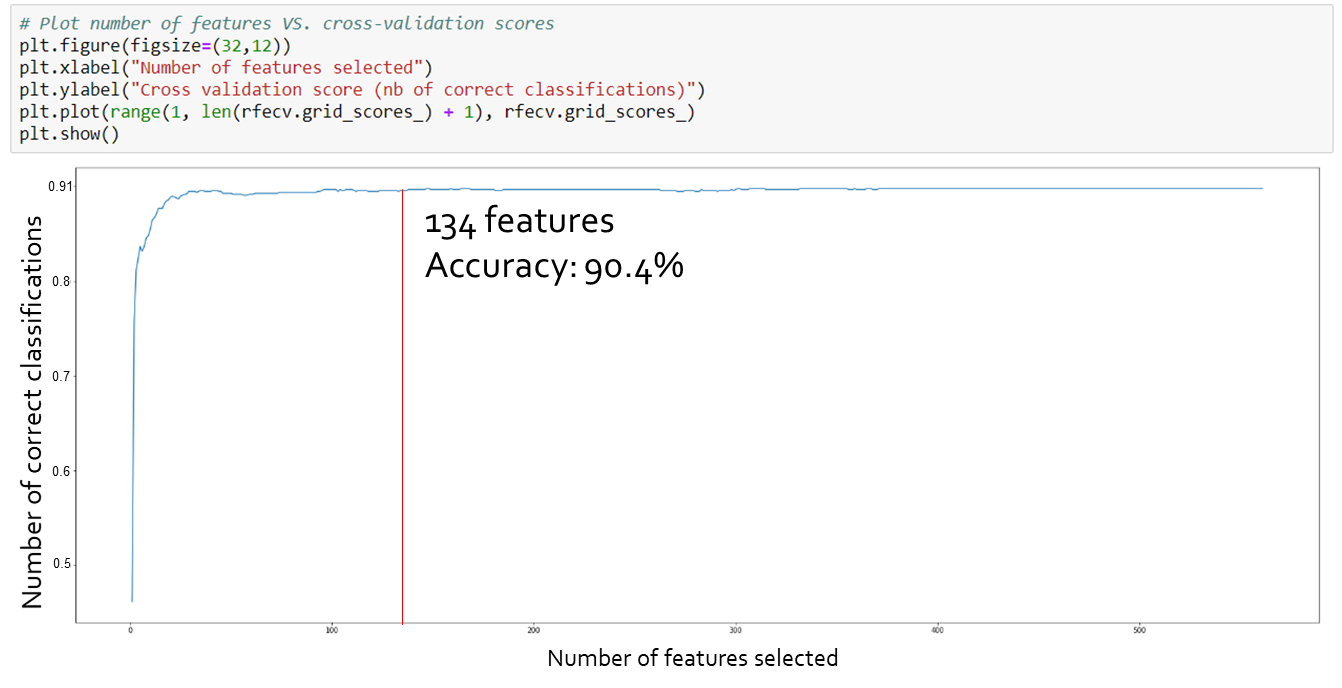
We explored several Machine Learning algorithms including: K-nearest neighbor, Logistic Regression, SGD, SVM, Keras Neural Network and Random Forrest; and reproduced the classification with a good accuracy in a common laptop but, in difference with the paper author’s experiment our best algorithm was Random Forrest with an acceptable accuracy of 91%.

**Figure 01: Machine Learning Algorithms Accuracy Comparison**

Noticing the considerable amount of time taking to process a small size file, and the huge features these sensors (gyroscope and accelerometer together) produce: 561 columns, we decided to apply some feature selection technique and check if accuracy would be affected.

So, we decided to apply PCA because of exploring the data, plot some columns and notice that some columns were relevant to try to differentiate the activity being performed but some others no, pointing us that these columns could be correlated with each other.

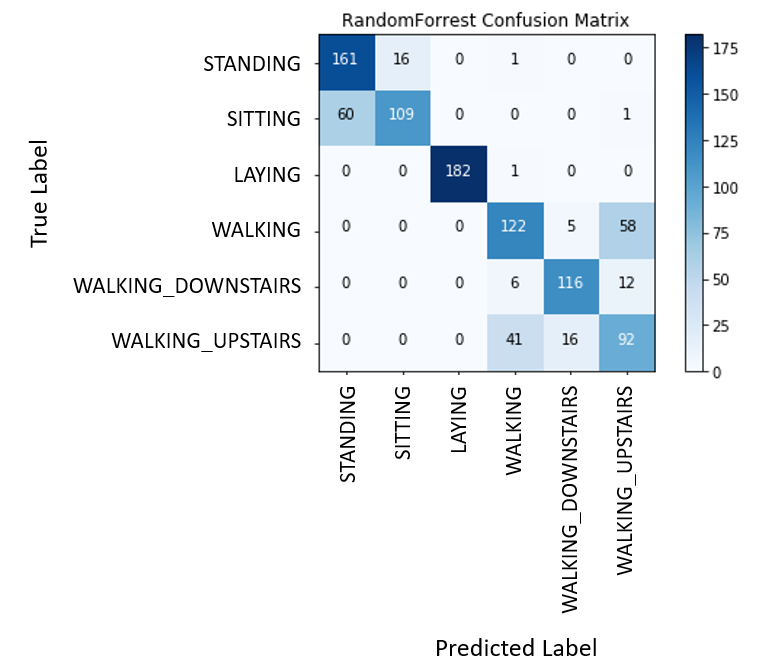
The experiment was successful: we reduced the initial number of features from 561 to 134 in the original UCI files, ran again the algorithms selected and no relevant accuracy problems arise, in fact in the Random Forrest we have an accuracy of 90.4%.

**Figure 02: Feature selection graph**

**New activity recognition: Falling**

We conducted a real experiment with our smartphones: tied the smartphone to the waist, started to record the activity being performed, with a help of an available application for android devices that records the sensors measures as a csv file, and record manually the time in order to have similar amount of records, in order to have the data balanced for the training file. Between those activities we added a new one: simulating the person was falling because a faint, recorded and added to the train file.

Later we extracted the information from smartphones, processed and split it into the train and the test files, ran the algorithms again with this own information and obtained similar results with good accuracy too using Random Forrest, and to conclude: We were able to classify correctly the new desired activity at processing the test file: Falling.



**Figure 03: Confusion Matrix from experiment**

## Conclusion

The next step here could be to develop a mobile application exploiting this Activity Recognition: Falling and to integrate functionalities to assist the person like: call automatically an emergency contact preconfigured, send automatically a message to a contact, call an emergency health or recue service, after detect the person doesn’t move after the fall and have no response from him/she after some seconds.

Further steps could conduce to mix smartphone data with smartwatches or smart bands connected to detect and send vital signs to provide more information, when those devices become more popular and accessible.

## Technologies Used

Programming Language: Python.

Machine: HP Laptop. Model: Envy. Processor: Intel Core i5-8250U CPU @ 1.60GHz. Memmory: 8.00 GB RAM. System type: 64-bit OS.

Software: Physics Toolbox Sensor Suite. Anaconda 3. Jupiter Notebook.

Smartphone: Sony XA 1.

## References

Statista – the Statistics Portal, <https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/>, December 2017.

GBD 2015 Disease and Injury Incidence and Prevalence, Collaborators. (8 October 2016). "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015"

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

Ruder, Sebastian. “An overview of gradient descent optimization algorithms.” CoRR abs/1609.04747 (2016).

Ho, Tin Kam (1995). Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282. Archived from the original (PDF) on 17 April 2016. Retrieved 5 June 2016

1. Statista – the Statistics Portal, <https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/>, December 2017. [↑](#footnote-ref-1)
2. GBD 2015 Disease and Injury Incidence and Prevalence, Collaborators. (8 October 2016). "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015" [↑](#footnote-ref-2)