**Background**

Using devices such as *Jawbone Up*, *Nike FuelBand*, and *Fitbit* it is now possible to collect a large amount of data about personal activity relatively inexpensively. These type of devices are part of the quantified self movement – a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. One thing that people regularly do is quantify how *much* of a particular activity they do, but they rarely quantify *how well they do it*. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbell of 6 participants. They were asked to perform barbell lifts correctly and incorrectly in 5 different ways. More information is available from the website here: <http://groupware.les.inf.puc-rio.br/har> (see the section on the Weight Lifting Exercise Dataset).

**Data**

The training data for this project are available here:

<https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv>

The test data are available here:

<https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv>

The data for this project come from this source: <http://groupware.les.inf.puc-rio.br/har>. If you use the document you create for this class for any purpose please cite them as they have been very generous in allowing their data to be used for this kind of assignment.

**What you should submit**

The goal of your project is to predict the manner in which they did the exercise. This is the "classe" variable in the training set. You may use any of the other variables to predict with. You should create a report describing how you built your model, how you used cross validation, what you think the expected out of sample error is, and why you made the choices you did. You will also use your prediction model to predict 20 different test cases.

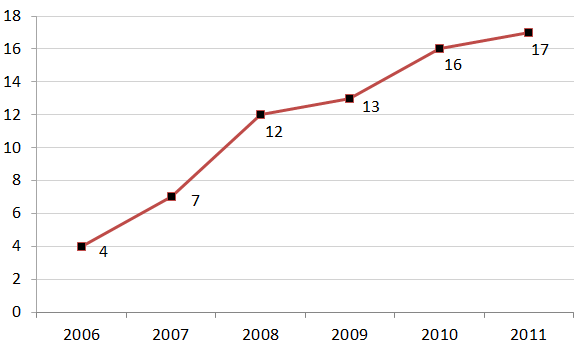
1. Your submission should consist of a link to a Github repo with your R markdown and compiled HTML file describing your analysis. Please constrain the text of the writeup to < 2000 words and the number of figures to be less than 5. It will make it easier for the graders if you submit a repo with a gh-pages branch so the HTML page can be viewed online (and you always want to make it easy on graders :-).

2. You should also apply your machine learning algorithm to the 20 test cases available in the test data above. Please submit your predictions in appropriate format to the programming assignment for automated grading. See the programming assignment for additional details.

**Reproducibility**   Due to security concerns with the exchange of R code, your code will not be run during the evaluation by your classmates. Please be sure that if they download the repo, they will be able to view the compiled HTML version of your analysis.

**Human Activity Recognition**

Human Activity Recognition - **HAR** - has emerged as a key research area in the last years and is gaining increasing attention by the pervasive computing research community (see picture below, that illustrates the increasing number of publications in HAR with wearable accelerometers), especially for the development of context-aware systems. There are many potential applications for HAR, like: elderly monitoring, life log systems for monitoring energy expenditure and for supporting weight-loss programs, and digital assistants for weight lifting exercises.



HAR: IEEE publications (2006-2011) based on wearable accelerometers' data

**Systematic-like approach for reviewing literature**

In order to enable you to replicate the literature review we made for this research, all publications assessed in this paper are [**available here, in RIS format**](http://groupware.les.inf.puc-rio.br/static/har/SystematicReview-RIS-Format.zip). In this research, the bibliographic management and publishing solution used was the EndNote X5(tm). The library in EndNote format is also [**available for download**](http://groupware.les.inf.puc-rio.br/static/har/SystematicReview-EndNote.zip).

**HAR Dataset for benchmarking**

We propose a dataset with 5 classes (sitting-down, standing-up, standing, walking, and sitting) collected on 8 hours of activities of 4 healthy subjects. We also established a baseline performance index. **You can** [**download the dataset here**](http://groupware.les.inf.puc-rio.br/static/har/dataset-har-PUC-Rio-ugulino.zip) (please, drop us a line (wugulino *at* inf *dot* puc-rio *dot* br) about your research and how we can contribute to your benchmarking).



**Important**: you are free to use this dataset for any purpose. **This dataset is licensed under the Creative Commons license (CC BY-SA)**. The CC BY-SA license means you can remix, tweak, and build upon this work even for commercial purposes, as long as you credit the authors of the original work and you license your new creations under the identical terms we are licensing to you. This license is often compared to "copyleft" free and open source software licenses. All new works based on this dataset will carry the same license, so any derivatives will also allow commercial use.

**Detailed Accuracy**

|  |  |  |
| --- | --- | --- |
| Correctly Classified Instances | 164662 | 99.4144 % |
| Incorrectly Classified Instances | 970 | 0.5856 % |
| Root mean squared error | 0.0463 |  |
| Relative absolute error | 0.7938 % |  |
| Relative absolute error | 0.7938 % |  |

**Detailed Accuracy by Class**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TP Rate** | **FP Rate** | **Precision** | **Recall** | **F-Measure** | **ROC Area** | **Class** |
| 0.999 | 0 | 1 | 0.999 | 0.999 | 1 | Sitting |
| 0.971 | 0.002 | 0.969 | 0.971 | 0.970 | 0.999 | Sitting down |
| 0.999 | 0.001 | 0.998 | 0.999 | 0.999 | 1 | Standing |
| 0.962 | 0.003 | 0.969 | 0.962 | 0.965 | 0.999 | Standing up |
| 0.998 | 0.001 | 0.998 | 0.998 | 0.998 | 1 | Walking |
| 0.994 | 0.001 | 0.994 | 0.994 | 0.994 | 1 | **Weighted Avg.** |

**Please, cite this publication to refer this dataset and literature review**

[Ugulino, W.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=ugulino); [Cardador, D.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=debora); [Vega, K.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=katia); [Velloso, E.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=evelloso); Milidiu, R.; [Fuks, H.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=hugo) [**Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements**](http://groupware.les.inf.puc-rio.br/work.jsf?p1=10335). Proceedings of 21st Brazilian Symposium on Artificial Intelligence. Advances in Artificial Intelligence - SBIA 2012. In: Lecture Notes in Computer Science. , pp. 52-61. Curitiba, PR: Springer Berlin / Heidelberg, 2012. ISBN 978-3-642-34458-9. DOI: 10.1007/978-3-642-34459-6\_6.  Cited by 2 (Google Scholar)

[Documento](http://groupware.les.inf.puc-rio.br/public/papers/2012.Ugulino.WearableComputing.HAR.Classifier.RIBBON.pdf) [Apresentação](http://groupware.les.inf.puc-rio.br/public/2012.SBIA.Ugulino.WearableComputing-Presentation.pdf)

**Other HAR Related Publications**

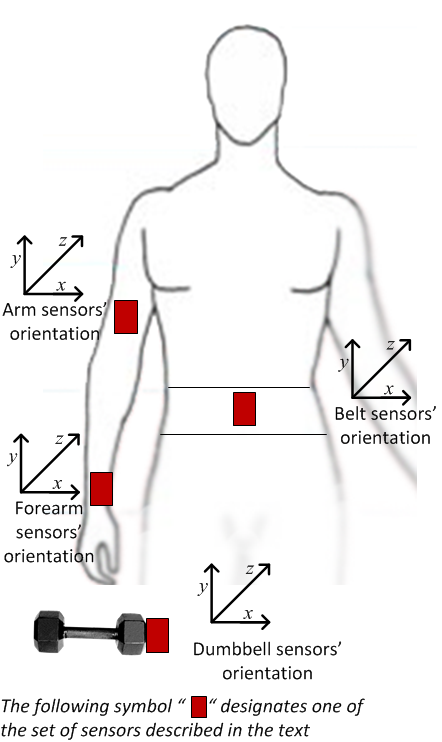
[Velloso, E.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=evelloso); Bulling, A.; Gellersen, H.; [Ugulino, W.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=ugulino); [Fuks, H.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=hugo) [**Qualitative Activity Recognition of Weight Lifting Exercises**](http://groupware.les.inf.puc-rio.br/work.jsf?p1=11201). Proceedings of 4th International Conference in Cooperation with SIGCHI (Augmented Human '13) . Stuttgart, Germany: ACM SIGCHI, 2013.

[Documento](http://groupware.les.inf.puc-rio.br/public/papers/2013.Velloso.QAR-WLE.pdf)

[Ugulino, W.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=ugulino); Ferreira, M.; [Velloso, E.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=evelloso); [Fuks, H.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=hugo) [**Virtual Caregiver: Colaboração de Parentes no Acompanhamento de Idosos**](http://groupware.les.inf.puc-rio.br/work.jsf?p1=10657). Anais do SBSC 2012, IX Simpósio Brasileiro de Sistemas Colaborativos , pp. 43-48. São Paulo, SP: IEEE, 2012. ISBN 978-0-7695-4890-6.

[Documento](http://groupware.les.inf.puc-rio.br/public/papers/2012.SBSC.Ugulino.VirtualCaregiver.pdf)

**Weight Lifting Exercises Dataset**



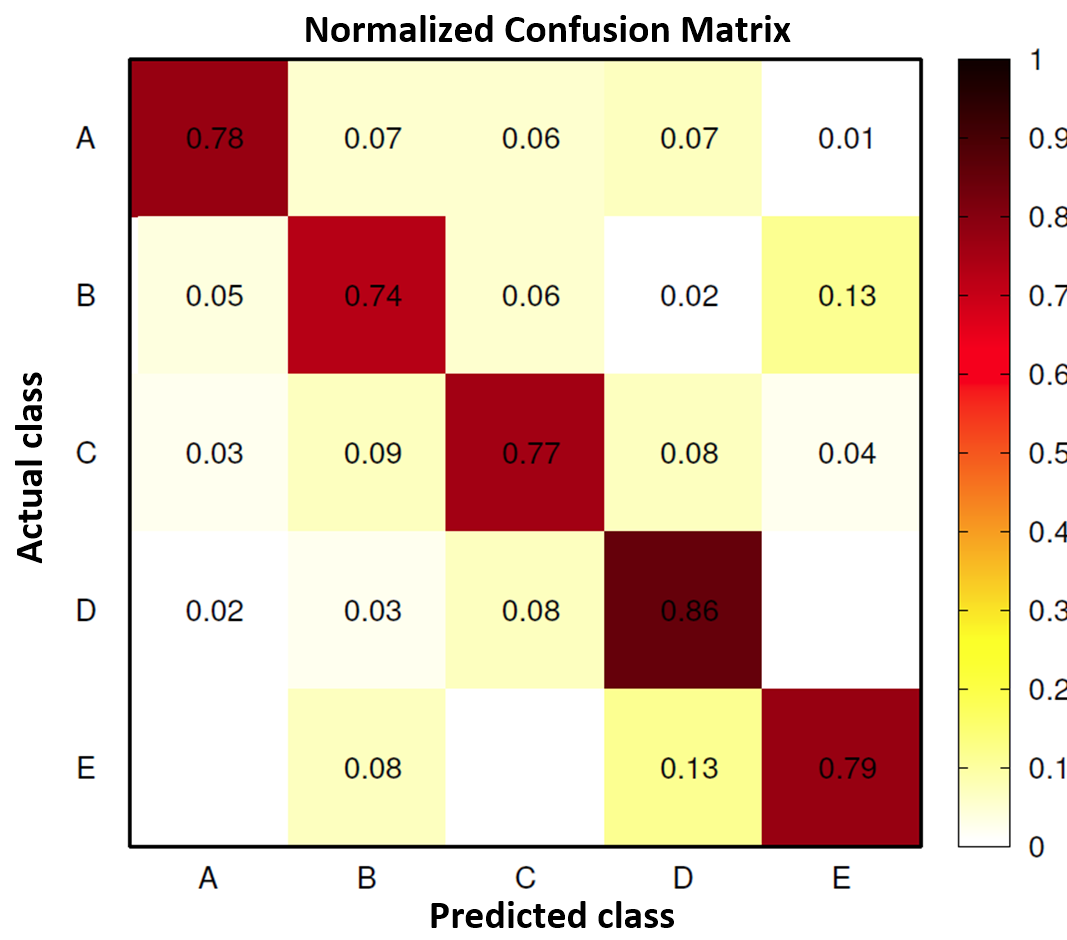
This human activity recognition research has traditionally focused on discriminating between different activities, i.e. to predict "which" activity was performed at a specific point in time (like with the Daily Living Activities dataset above). The approach we propose for the Weight Lifting Exercises dataset is to investigate "how (well)" an activity was performed by the wearer. The "how (well)" investigation has only received little attention so far, even though it potentially provides useful information for a large variety of applications,such as sports training.

In this work ([see the paper](http://groupware.les.inf.puc-rio.br/work.jsf?p1=11201)) we first define quality of execution and investigate three aspects that pertain to qualitative activity recognition: the problem of specifying correct execution, the automatic and robust detection of execution mistakes, and how to provide feedback on the quality of execution to the user. We tried out an on-body sensing approach ([dataset here](http://groupware.les.inf.puc-rio.br/static/WLE/WearableComputing_weight_lifting_exercises_biceps_curl_variations.csv)), but also an "ambient sensing approach" (by using Microsoft Kinect - dataset still unavailable)

Six young health participants were asked to perform one set of 10 repetitions of the Unilateral Dumbbell Biceps Curl in five different fashions: exactly according to the specification (Class A), throwing the elbows to the front (Class B), lifting the dumbbell only halfway (Class C), lowering the dumbbell only halfway (Class D) and throwing the hips to the front (Class E).

Class A corresponds to the specified execution of the exercise, while the other 4 classes correspond to common mistakes. Participants were supervised by an experienced weight lifter to make sure the execution complied to the manner they were supposed to simulate. The exercises were performed by six male participants aged between 20-28 years, with little weight lifting experience. We made sure that all participants could easily simulate the mistakes in a safe and controlled manner by using a relatively light dumbbell (1.25kg).

[**Download the WLE dataset here**](http://groupware.les.inf.puc-rio.br/static/WLE/WearableComputing_weight_lifting_exercises_biceps_curl_variations.csv)





**Important**: you are free to use this dataset for any purpose. **This dataset is licensed under the Creative Commons license (CC BY-SA)**. The CC BY-SA license means you can remix, tweak, and build upon this work even for commercial purposes, as long as you credit the authors of the original work and you license your new creations under the identical terms we are licensing to you. This license is often compared to "copyleft" free and open source software licenses. All new works based on this dataset will carry the same license, so any derivatives will also allow commercial use.

Read more: <http://groupware.les.inf.puc-rio.br/har#ixzz34tug16A5>

**Please, cite this paper to refer the WLE dataset**

[Velloso, E.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=evelloso); Bulling, A.; Gellersen, H.; [Ugulino, W.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=ugulino); [Fuks, H.](http://groupware.les.inf.puc-rio.br/collaborator.jsf?p1=hugo) [**Qualitative Activity Recognition of Weight Lifting Exercises**](http://groupware.les.inf.puc-rio.br/work.jsf?p1=11201). Proceedings of 4th International Conference in Cooperation with SIGCHI (Augmented Human '13) . Stuttgart, Germany: ACM SIGCHI, 2013.

Read more: <http://groupware.les.inf.puc-rio.br/har#ixzz34tuomUON>