Human Activity Recognition with Smartphones

Stephen McKechnie

Motivation

As of 2015, nearly two-thirds (64%) of U.S. adults own a smartphone, up from 35% in 2011. *Source: Pew Research Center*

- Smartphones gives us access to almost everything
- Apps can manage various aspects of our lives
- Smartphones are embedded with advanced technology

Data

A study of 30 participants each performing 6 activities of daily living while wearing a smartphone on their waist.

Activities

- Walking, Walking Upstairs, Walking Downstairs
- Sitting, Standing, Laying

561 variables relating to various accelerometers and gyroscopes measurements as the participant performed each activity.

The objective is to classify the measurements into one of the six activities performed.

Methods

Classification Trees

Statistical Model:

$$\hat{\boldsymbol{y}}(\boldsymbol{x_i}) = \sum_{t=1}^{T} \left(\arg\max_{k} \boldsymbol{\pi_k} \right) \mathtt{I}(\boldsymbol{x_i} \in \boldsymbol{\mathcal{R}_t}),$$

where $\pi_k =$ Proportion of $\{y(x_i) : x_i \in \mathcal{R}_t\}$ of class k.

Growing a Tree: Find regions that minimize:

$$Error(y, \hat{y}) + \lambda T$$

$$Error(y, \hat{y}) = \text{Gini Index} = \sum_{k=1}^{K} \hat{\pi}_{tk}(1 - \hat{\pi}_{tk})$$

where λ penalizes the size of the tree, T, and $\hat{\pi}_{tk} =$ Proportion of obs in Region t of class k.

Methods

Classification Trees

Gini Index takes small values if $\hat{\pi}_{tk}$ is close to 0 or 1. Meaning a node contains predominantly observations from a single class.

Recursive Partitioning

1. Find a predictor X_p and a cut point s such that splitting space into $\{X|X_p < s\}$ and $\{X|X_p \geq s\}$ leads to greatest reduction in error
2. Repeat step 1 repeatedly but only split previously defined regions

Methods

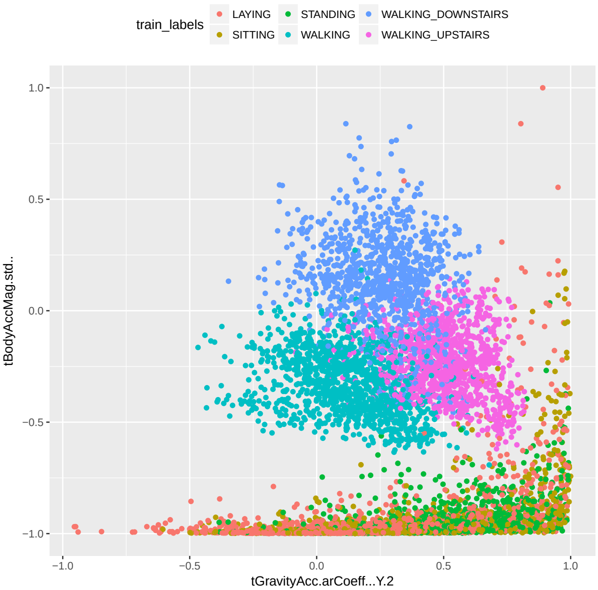
Random Forests

Similar process as Classification Trees, however selecting only $m < P$ predictors to build a tree.

Algorithm

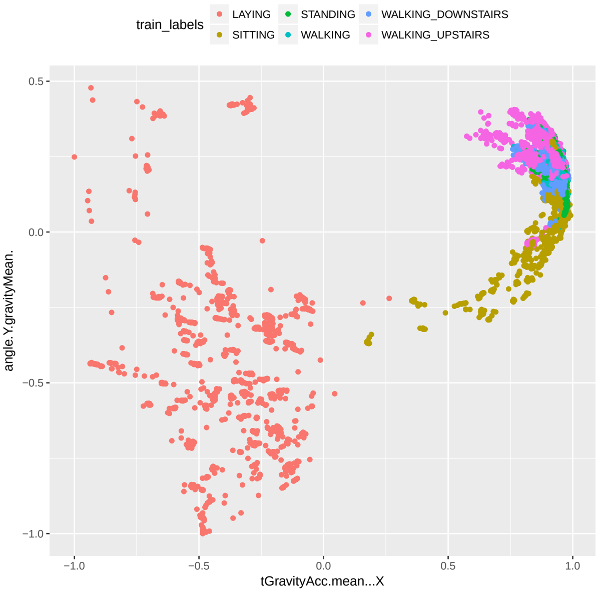
1. For $b = 1, \ldots, B$:
   1.1 Take a bootstrapped sample of size n of data
   1.2 At each node, randomly consider $m < P$ variables
   1.3 Build a tree $\mathcal{T}_b$
2. To predict a class for data point x_0
   2.1 For $b = 1, \ldots, B$: pass x_0 down to the b^{th} tree to get the bootstrapped prediction $\hat{y}^b(x_0)$
   2.2 Average predictions to get $\boldsymbol{\hat{y}(x_0)} = \text{mode}(\boldsymbol{\hat{y}^1(x_0), \ldots, \hat{y}^B(x_0)})$

Exploratory Analysis
Motion



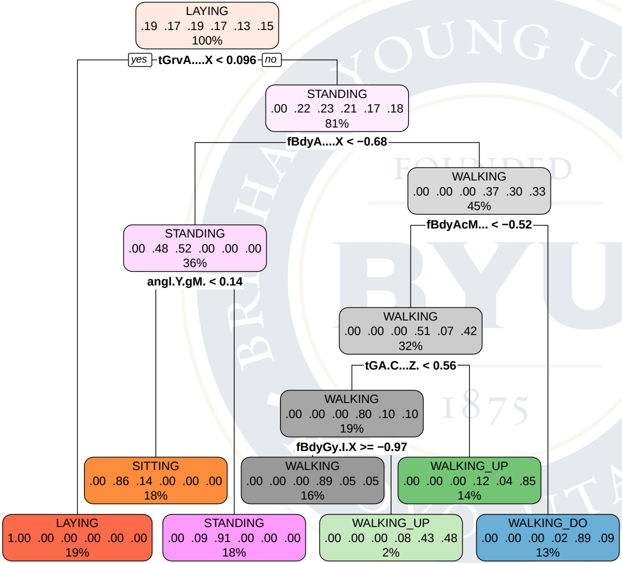
Exploratory Analysis

Stationary



Tree

Classification Trees



Predictions - Test Data

Classification Trees

Predictive Rate of **84.0%**

	True Classification					
Prediction	Laying	Sitting	Standing	Walking	Walking Downstairs	Walking Upstairs
Laying	537	0	0	0	0	0
Sitting	0	400	107	0	0	0
Standing	0	91	425	0	0	0
Walking	0	0	0	430	40	61
Walking-D	0	0	0	11	282	8
Walking-U	0	0	0	55	98	402

Predictions - Test Data

Random Forests

Predictive Rate of **92.9%**

	True Classification					
Prediction	Laying	Sitting	Standing	Walking	Walking Downstairs	Walking Upstairs
Laying	537	0	0	0	0	0
Sitting	0	437	43	0	0	0
Standing	0	54	489	0	0	0
Walking	0	0	0	481	19	29
Walking-D	0	0	0	10	358	7
Walking-U	0	0	0	5	43	435

Conclusions

- Classification Trees do well in predicting the correct activity and visualizing the decision process
- However, Random Forests performed better in predicting the correct activity

Future Research

Explore other machine learning algorithms

- Support Vector Machines and Neural Networks