

Statistics 101C Discussion Week 5

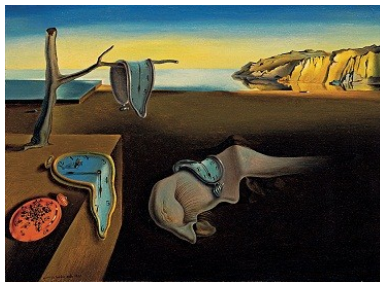
Brice Randolph

University of California, Los Angeles

github.com/briceran

April 26, 2016

Agenda



What's the plan?

- Building an R package
- Cross Validation with caret package
- Talk about article
- Spotipy API
- Robotics Club

The Dude Recommends



- <https://www.coursera.org/learn/practical-machine-learning/lecture/HZKcr/cross-validation>

Unifying framework for Machine Learning

- Preprocessing (`preProcess`)
- Data Splitting (`createDataPartition` `createResample`)
- Training/testing functions (`train` `predict`)
- Model Comparisons (`confusionMatrix`)
- ML algorithms (LDA, SVM, Regression Trees, Naive Bayes, Random Forests, Boosting Regression)

Article - Statistical Modeling: The Two Cultures

What do you think?

Generate a user token online

Join

Exercise in R

See my github for wk5 R code: github.com/briceran

Practice Job interview Question

Use R or Python:

First think of helper functions that you could use.