

Stats and ML in Julia: Status Update

March 17, 2015

Julia's got a good foundation for stats and ML

But it needs a lot of work on packages (esp. data infrastructure)

Doing things with floating-point arrays is pretty smooth

Everything else is pretty rough

Basic stats packages

- ▶ StatsBase
- ▶ Distributions
- ▶ Distances

- ▶ HypothesisTests
- ▶ PowerAnalysis
- ▶ StreamStats
- ▶ MultivariateStats
- ▶ Bootstrap

- ▶ DataArrays
- ▶ DataFrames

- ▶ GLM
- ▶ MixedModels

Generic model fitting via optimization

- ▶ Calculus
- ▶ Optim
- ▶ NLOpt
- ▶ JuMP
- ▶ Convex

Basic ML tools

► Clustering

- ▶ FLANN
- ▶ KDTrees
- ▶ NearestNeighbors
- ▶ kNN

- ▶ GLMNet
- ▶ Lasso
- ▶ LARS
- ▶ LIBSVM
- ▶ SVM

Graphics

- ▶ Gadfly
- ▶ Interact
- ▶ Loess

- ▶ TopicModels
- ▶ TextAnalysis