Lanfeng Pan

Iowa State University
Department of Statistics
Snedecor Hall
Ames, IA 50010

Phone: (515)817-3821
Email: panlanfeng@gmail.com
GitHub: github.com/panlanfeng/
Homepage: lanfeng.me

Education

Ph.D., Iowa State University, 2012 – Now.

Master, Renmin University of China, 2010 – 2012.

Bachelor, Renmin University of China, 2006 – 2010.

Research Experience

PAN, L., LI, Y., HE, K., LI, Y. and LI, Y. (2016). Latent Gaussian Mixture Models For Nationwide Kidney Transplant Center Evaluation. *The Annals of Applied Statistics* (submitted).

Research Assistant, 2014 – Now.

The project was to evaluate the performance of certain facilities on a national wide large data. The major difficulties: 1. lack of independence; 2. some facilities have too few records to be precisely evaluated; and 3. the facilities are heterogeneous. I originated a new approach which successfully improved the current solution.

Intern at Novartis Pharmaceuticals, NJ, May 2015 – August 2015.

Project 1: Building `shiny` apps to help other statisticians to visualize and analyze their data.

Project 2: Modeling and visualizing labor investment in hundreds of pharmaceutical projects, detecting potential project delays and predicting future labor investments.

First Place in the 15th Annual Data Mining Cup, May 2014.

One of the nine team members. The data was about online shopping. Task was to predict returning probability of a purchase given customer shopping records and item information. There were no explicitly useful feature available and all records were correlated. We extracted every useful information by careful data transformation and grouping. We finally ended up with the lowest prediction error rate among teams all over the world.

Agriculture Experiment Station Consulting Group, May 2014 – July 2014.

Job was answering questions from random visitors from other departments. Need to communicate with the visitors to figure out their questions, help to summarize their question in statistical language and then guide them to the solutions.

Teaching Assistant, August 2012 – May 2014.

Worked as teaching assistant for STAT 341, 342, 447, 542 and 543. Major duties were answering questions, helping with homework and grading.

Research Interests

Multiple Testing, False Discovery Rate Control

Clustering, Subgroup Analysis

High Performance Computing, Data Mining

Health Policy

Skills

Proficient with R, Julia, Python, Linux Shell and git.

Proficient with `shiny`, `ggplot2`, `knitr`, `rmarkdown` and TEX.

Contributions

Julia Package: `LatentGaussianMixtureModel.jl`. Fit a Generalized Linear Mixed Model with Gaussian mixture random effects, deciding the number of components for Gaussian mixture. And further conduct a multiple test to detect heterogeneity while control the False Discovery Rate.

Julia Package: `RFlavor.jl`. Implement a lot of useful and handy R functions in Julia. The purpose is to provide better statistical functions for Julia language as well as make it easy to translate R code into Julia.

Julia Package: `KernelEstimator.jl`. Implement kernel density estimation and kernel regression. In particular this package can deal with bounded kernel estimation using beta and gamma kernel and can choose bandwidth via cross valuation.

Julia Package: `GaussianMixtureTest.jl`. Implement the Kasahara-Shimotsu Test to decide number of components in Gaussian Mixture Model. There is very few package in this area.

Contribute to Julia Package: `Yeppp.jl`. This package ports the `Yeppp!` library into Julia, significantly speeding up several basic arithmetic operations.

Contribute to several core statistical packages in Julia community including `StatsBase.jl`, `Rmath.jl`, `DataArrays.jl` and `KernelDensity.jl`.

R package: `bignmf`. Solve the nonnegative matrix factorization problem using coordinate descent.