

Andrew Cron

3217 Appling Way
Durham, NC 27705

Email: andrew.j.cron@gmail.com
Phone: (919) 627-8564

Education

DUKE UNIVERSITY
Ph.D. Statistics, 2012

UNIVERSITY OF SOUTH CAROLINA
B.S. Statistics-Mathematics, 2009

Experience & Employment

Data Scientist, *February 2017 to Current*
CITADEL, LLC

Senior Research Scientist, *August 2015 to February 2017*
85.51°

Lead the development of a predictive modeling framework for massive retail forecasting problems. This included developing new modeling approaches, advising multiple research projects, supervising projects within this initiative, and managing collaborations with academia.

CTO, *March 2014 to August 2015*
WEINRAUB ANALYTICS

Managed implementation and research for quantitative financial modeling of US equities. This included maintaining backend data systems and redundancy, building machine learning and time series models for asset management, and developing robust software solutions for business execution.

Adjunct Faculty Member, *January 2015 to Present*
DEPARTMENT OF STATISTICAL SCIENCE, DUKE UNIVERSITY

Data Scientist, *December 2012 - March 2014*
MAXPOINT

Developed and implemented statistical and machine learning methods on massive datasets to deliver cutting-edge targeted advertising solutions using hadoop and high performance computing frameworks.

Research Intern, *Summer 2012*
IBM WATSON RESEARCH CENTER

Developed methodology for online fitting of dynamic clustering point processes applied to crime data. Created a Bayesian approach for dynamic coregionalization models for climate sensor analysis with missing data.

Research Intern, *Summer 2011*

YAHOO! LABS

Developed generalized latent variable regression models for recommendation systems using high throughput, sequential learning algorithms resulting in significantly improved recommendation performance.

Research Assistant, *Fall 2009 - Fall 2012*

DUKE UNIVERSITY DEPARTMENT OF STATISTICS

Used GPU technology to address mixture model fitting issues on massive datasets in flow cytometry. Efficiently addressed label switching issues in MCMC algorithms. Developed an informative Bayesian approach to metabolite detection in mass spectrometry applications for metabolomics. Created novel methods for flexible, sparse covariance modeling on large datasets.

Skills

High dimensional predictive modeling, Bayesian methods, semi-parametric analysis, statistical and machine learning algorithms, online learning, high throughput computing, distributed and parallel computing, GPU computing, data visualization, web development, software development, project management

Computing

Python, C/C++, R, Julia, Matlab, Hadoop, Javascript, HTML/CSS, SQL, Bash/Unix Utilities.

Consulting Experience

ANGULAR, *Spring 2011 - March 2014*

Implemented a Latent Dirichlet Allocation tool in a parallel computing framework on the Amazon EC2 Cloud for assessing document relevance to improve web search rank performance.

TRIALRATE, *Spring 2010 - Fall 2012*

Developed the methodology and software to analyze multi-scale survey data using multivariate probit models with standard latent normal models and latent factor models.

Professional Service

Section Chair, Junior ISBA, 2012-2014

Guest Editor, XRDS Crossroads: ACM Magazine for Students, 2012

Papers

M. Suchard, Q. Wang, C. Chan, J. Frelinger, A. Cron and M. West. "Understanding GPU programming for statistical computation: Studies in massively parallel massive

mixtures.” *Journal of Computational and Graphical Statistics* 19 (2010): 419-438

A. Cron and M. West. “Efficient Classification-Based Relabeling in Mixture Models.” *The American Statistician* 65 (2011): 16-20.

W. McKinney and A. Cron. “gpustats: GPU Library for Statistical Computing” *SciPy Conference*, Austin, TX, July 2011. The proceedings of the 10th Python in Science Conference.

A. Cron, L. Zhang, and D. Agarwal. “Collaborative filtering for massive multinomial data.” *Journal of Applied Statistics* (2013): 1-15.

A. Cron, C. Gouttefangeas, J. Frelinger, L. Lin, S. K. Singh, C. M. Britten, M. J. P. Welters, S. H. van der Burg, M. West, C. Chan. “Hierarchical Modeling for Rare Event Detection and Cell Subset Alignment across Flow Cytometry Samples.” *PLOS Computational Biology* 9 (2013).

A. Cron and M. West. “Modeling Sparse Orthogonal Matrices using Givens Rotations.” *Statistical Analysis for High Dimensional Data* (2016).

Presentations

“GPU Computing in Statistics.” *Interface*, Seattle, WA, June 2010.

“Bayesian Statistics in Metabolomics.” *Bayesian Biostatistics Conference*, Houston, TX, January 2011.

“Modeling Sparse Orthogonal Matrices.” *ISBA World Meeting*, Kyoto, Japan, June 2012.

Software

DPMIX: Python library for Bayesian inference for hierarchical and standard Dirichlet process mixtures of normals using GPUs.

Available at: <http://github.com/andrewcron/dpmix/>.

GPUSTATS: Python library for GPU implementations of statistical functions in python.

Available at: <https://github.com/dukestats/gpustats/>.

GPUMIX: C++ application for fitting Dirichlet process mixture models using GPUs.

Available at: <http://www.stat.duke.edu/gpustatsci/>.

CY-ARMADILLO: C++ to Python port of the Armadillo linear algebra library using Cython.

Available at: https://github.com/andrewcron/cy_armadillo/.

References

Available upon request.