Chunyi Zhao

Phone: +1 (207) 720 0755
Email: jessica.cy.zhao@gmail.com;
Homepage: https://jesscyzhao.github.io/
Github: jesscyzhao

Education

2017- **Ph.D. Statistical Science,**
University of California, Santa Cruz, *Santa Cruz, CA*
Research field: *Bayesian Nonparametrics, Mixture model, Point Processes*

2011-2015 **B.A. Mathematics,**
Bowdoin College, *Brunswick, ME*

Skills

Programming	Python, Julia, SQL
Statistical Packages	R, Stan, Turing.jl
Markup / Visualization	Tableau, LaTeX, R Markdown
Languages	English, Chinese

Employment

Jun 2021 - Sep 2021 **Data Scientist Intern (Causal Machine Learning)**, Netflix
Designed a causal segment discovery workflow for the Netflix experimentation platform that utilizes state-of-the-art machine learning algorithms to provide inference on segment specification and subgroup treatment effect. The main goal was to help A/B test users to identify cohorts of streaming accounts or devices that may experience large treatment effect to guard against potential engineering regression. Developed a prototype for an end-to-end analysis tool that automatically runs the analysis and presents a heterogeneous treatment effect (HTE) by feature visualization and a decision-rule-based HTE report at the end of each A/B test.

Jun 2020 - Sep 2020 **Data Scientist Intern (Sequential AB Testing)**, Netflix
Developed a novel sequential A/B testing method applied to real-time data in a sequential update setting where decisions are updated iteratively as data streams in. Implemented the new model in Python and benchmarked it against the existing model in production. The comparison showed significant speed (in terms of observations used) and accuracy (in terms of power) gain in the new model.

2018-Present **Research Assistant**, University of California, Santa Cruz
Spatial Point process modeling with Bayesian nonparametric approaches, supervised by Professor Athanasios Kottas .

Fall 2020-Fall 2021 **Graduate Student Instructor**, University of California, Santa Cruz
Instructor for STAT 266A: Data Visualization and Statistical Programming in R. Created labs and assignments using Google Colab. Supervised students on course projects.

2019-2021 **Graduate Student Representative**, University of California, Santa Cruz
Organized and coordinated multiple department-wide events that include faculty hiring, admission, and faculty-student socials. Point of contact for graduate students in department-related affairs.

Fall 2017- **Teaching Assistant**, University of California, Santa Cruz
Created and led discussion sections, graded exams, and presented lectures in a series of undergraduate and graduate statistics classes.

2015 - 2017 **Data Scientist**, Two Six Capital

Statistical modeling and programming for private equity consulting.
• Built the company's cloud-based and distributed data science platform capable of scaling and parallelizing with multiple computational back-ends in Python, Cython and Spark in a small team. Shortened the analysis delivery cycle from 2 weeks to 4 days.

• Implemented and tuned statistical models for forecasting customer acquisition, retention and purchasing behaviors in SMB services, subscription and retail businesses. Produced monthly and quarterly top-line revenue projection by aggregating model forecasts.

• Implemented a Bayesian hierarchical model for a more robust forecasting of the retail customers' repeat purchase behavior. Such model improved prediction accuracy by more than 200% compared to existing models in situations where the customers are churned off.

• Created data processing pipelines to automate the due diligence analysis on 3+ TBs of transaction-level data.

Research

Research Interests

Bayesian Statistics • Spatial Statistics • Mixture Modeling • Point Processes • Experimentation & Causal Inference

Work in Progress

2. Zhao, C. and Kottas, A. *Bayesian nonparametric modeling for spatial Hawkes process*

1. Zhao, C. and Kottas, A. *Modelling for Poisson process intensities over irregular spatial domains.* [arXiv]

Teaching

F20, S20, F21	**STAT 266A (graduate level): Data Visualization and Statistical Programming in R**
W19, W20	**STAT 132: Classical and Bayesian Inference**, TA for multiple instructors
S19	**STAT 204 (graduate level): Introduction to Statistical Data Analysis**, TA for Prof. R. Prado
Multiple Quarters	**STAT 5 (Introductory Statistics)**, TA for multiple instructors
Multiple Quarters	**STAT 7 (Introductory Biostatistics)**, TA for multiple instructors

References

Athanasios Kottas Professor, Department of Statistics UCSC	Juhee Lee Associate Professor, Graduate Director, Department of Statistics UCSC
Julie (Novak) Beckley Manager, Experimentation Etsy	Andy Rhines Senior Streaming Experimentation Data Scientist Netflix
Simon Ejdemyr Senior Data Scientist Netflix	Martin Tingley Manager, Experimentation Platform Netflix

Last updated: September 30, 2021
https://jesscyzhao.github.io/