Xiaoyue Li, Ph.D. Candidate

CONTACT INFORMATION
+1-650-772-2359
xyrli@ucdavis.edu

SUMMARY

Ph.D. candidate in Statistics with strong technical skills and 5+ year experience in research, consulting and data analysis. Extensive knowledge of statistical modeling (generalized linear models, change-point detection, covariance estimation, computational statistics), machine learning (graph theory, matrix completion, clustering, tree-based learning), optimization (SGD, duality, ADMM, parallel optimization), and effectively communicating results to audience from various backgrounds.

PROGRAMMING SKILLS

Proficient in *Python, R, Julia* and *Matlab*
Experienced in *SQL, C++, Java, SAS, LaTeX and Linux*

EDUCATION

University of California - Davis, CA, United States

Ph.D. Statistics, *Expected:* June 2020 (GPA 3.964/4.0)

- Advisor: Prof. James Sharpnack

Hong Kong University of Science and Technology, Hong Kong

BSc. Risk Management and Business Intelligence, 2015 (GPA: 3.752/4.3) (A: 4.0, A+: 4.3)

- Minor in Information Technology
- Minor in Mathematics

RESEARCH EXPERIENCE

Compression of Spatio-Temporal Networks via Point-to-Point Process Models Jun 2016 - Present
Python, Julia, SQL

- Work published in Proceedings of International Workshop on Mining and Learning with Graphs, 2017
- Developed a novel framework to model the stochastic process of spatio-temporal networks
- Applied the framework to NYC taxi dataset to compress the trip demand spatially and temporally:
  - Cleaned, transformed and analyzed large scale datasets consisting information of more than 100 million taxi trips
  - Estimated a spatially smoothed community structure and localized temporal change-points for the network
  - Derived and implemented an ADMM optimizer to solve a group-fused LASSO program
  - Visualized spatial clustering of 8000 grid areas with GIS shape file in Python with interactive display of cluster ID with data cursor

Taxi Pickups Near Subway Stations: A GLM Approach Jan 2017 - Mar 2017
Python, R, SQL

- Modeled how taxi pickups near subway entrances during rush hours covary with characteristics of the station, trip date, and subway arrival information
- Fitted a generalized additive partial linear model with negative binomial family after diagnostics and model selection
- Processed complex dataset integrated from 3 data sources involving gtfs-realtime, SQL and web scrapping

Copula Based Modelling in Geostatistics Sep 2013 - Aug 2015
Extreme value theory, Copula, Variogram, R

- Initiated a copula based method to model dependence structure in spatial data with the presence of extreme values
- Incorporated a skewed version of t copula to address asymmetric tail dependence
- Studied various advanced topics including Bayesian estimation, geostatistics, multivariate copula, extreme value theory

Consulting Experience

Research Assistant of Prof. SO, Mike Ka Pui, HKUST Oct 2014 - Mar 2015
Dept of Information Systems, Business Statistics and Operations Management

- Conducted exploratory data analysis and correlation analysis for contract list data and call list data
- Summarized results from over 50 papers to carry out literature review regarding employees' unethical behaviors in information systems

Teaching Experience

Teaching Assistant, UC Davis Sep 2015 - Present

- Responsibilities: preparing materials, leading discussions, holding office hours and managing teamwork.
- Courses taught: Data and Web Technologies for Data Analysis (Python for Statistics), ANOVA, Elementary Statistics, Applied Statistics for Biological Sciences etc.
- Vast majority gave 'excellent' or 'very good' ratings.
- Positive student feedback on effectiveness in communication, helpfulness in office hours and enthusiasm in teaching.