ZHENG GAO +1 734 846 5221 | gaozheng92@gmail.com WEBPAGE: http://pill-gz.github.io/

Skills

Computer skills:	Python, R, SQL , Metaflow, git, Docker.
Machine learning:	scikit-learn, xgboost, PyTorch.
Statistical knowledge:	experimentation and causal inference, high-dimensional statistics.

WORK EXPERIENCE

Apr. 2022 -	Senior Research Scientist at UPSTART. Machine Learning, Credit Risk
	• Led end-to-end development and deployment of machine learning models for real time credit risk assessment and pricing, overseeing algorithm design (boosted trees, ensembles), deployment, backtesting, and performance monitoring. Improved predictive accuracy by a cumulative 3.46% through four major model updates, driving 20 times deal volume increase over two years.
	• Led research on high-dimensional credit data features, incorporating both customer-provided and third- party data, to optimize model architectures, improve predictive accuracy, robustness, and enhance decision- making efficiency.
	• Led research to quantify uncertainty in model performance metrics, built and deployed user-friendly web app [link] enabling stakeholders to uncover business opportunities within operational constraints, bringing down excess loss by 50%.
	• Conducted ancillary research in causal inference methods in underwriting, led research workshops for 7 senior research scientists.
Jul. 2020 -	Quantitative Researcher at DRW, Chicago, IL.
Apr. 2022	JAX Delta One Research
	• Conducted research and predictive analysis on market microstructure, Treasury futures, and Treasury cash pricing. Developed statistical models to build pricing models for mid- to long-term Treasury products.
	• Performed post-trade statistical analysis for Treasury spread trading and developed data dashboards to automate the reporting of key performance metrics and execution statistics.
Feb - Apr.	Research Intern at POINT72 Asset MANAGEMENT, New York, NY.
2020	Market Intelligence research
May - Aug. 2017	Research Intern at MERIT NETWORKS INC., Ann Arbor, MI. Real-time Network Traffic Monitoring and Anomaly Detection
2014 - 2015	Research Scientist at SAW SWEE HOCK SCHOOL OF PUBLIC HEALTH, Singapore. Infectious Diseases Modeling and Statistical Analysis
EDUCATIO	N
2015 2020	DhD Statistics University of Michigan Ann Arbor Advisor Stillion Staay

2015 - 2020	 PhD Statistics. University of Michigan, Ann Arbor. Advisor: Stilian Stoev. Best Student Paper Award at Extreme Value Analysis (EVA) 2019. 1st place in Michigan Datathon hosted by Citadel & CorrelationOne. 	
2010 - 2014	B.Eng. National University of Singapore, Singapore. Majors: Mathematics, Industrial & Systems Engineering, Economics	
2013	 Lee Ruan Yew Gold Medal (1rd in graduating class), First Class Honors with Distinction. Winner of 2014 Revolution Analytics R Programming Competition. Undergraduate Exchange University of Wisconsin, Madison, United States. 	

SOFTWARE

U-PASS:	Unified Power Analysis and Forensics for Association Studies
	• An interactive, user-friendly web application, hosted at https://pill-gz.shinyapps.io/U-PASS/; source available at https://github.com/Pill-GZ/U-PASS
CASC:	Covariate-assisted Spectral Clustering
	• An R (Rcpp) package implementing the Norbert Binkiewicz's Covariate-assisted Spectral Clustering (2016) algorithm, runs 30 times faster than the original authors' package. Source available at https://github.com/Pill-GZ/CASC

SELECTED PUBLICATIONS

- Z.Gao, S.Stoev, Concentration of Maxima and Fundamental Limits in High-Dimensional Testing and Inference SBPMS Springer Monograph (2021).
- Z.Gao, S.Stoev, Fundamental Limits of Exact Support Recovery in High Dimensions Best Student Paper at Extreme Value Analysis (EVA) 2019, Bernoulli 26.4 (2020): 2605-2638.
- **Z.Gao**, *Five Shades of Grey: Phase Transitions in High-dimensional Multiple Testing* 1st place research award at the 50-year Anniversary Symposium of UM Statistics, arXiv:1910.05701 (2019).
- **Z.Gao**, J.Terhorst, C.Van Hout, S.Stoev, *U-PASS: Unified Power Analysis and Forensics for Qualitative Traits in Genetic Association Studies*, Bioinformatics 36.3 (2020): 974-975. [software link]