

EDUCATION

- **University of Wisconsin - Madison** Sept. 2021 - Expected 2025
Doctor of Philosophy in Statistics, Minor in Computer Science CGPA: 4.0/4.0
- **University of Wisconsin - Madison** Sept. 2020 - May 2021
Master of Science in Statistics, Speciality in Data Science CGPA: 4.0/4.0
- **University of Wisconsin - Madison** Sept. 2019 - May 2020
Visiting International Student CGPA: 4.0/4.0
- **Zhejiang University** Sept. 2016 - Jun. 2020
Bachelor of Natural Sciences in Statistics CGPA: 3.8/4.0

RESEARCH EXPERIENCE

- **Pass-to-rank(PTR) Robust Spectral Clustering** - In manuscript Sept. 2022 - Present
Novel robust spectral clustering method utilizing rank statistics, eigendecomposition, and k-means UW - Madison
 - Established tail bounds for the spectral norm of random matrices of normalized rank statistics by using random matrix theory.
 - Proved the strong and weak consistency of the Pass-to-rank (PTR) spectral clustering method, providing optimal theoretical foundations.
 - Evaluated the performance of the PTR spectral clustering by applying it to simulated gene co-expression data.
- **Censored AcF Model for Financial Market Risk Studies** May 2020 - May 2021
Innovative stochastic time series model based on Extreme Value Theory UW - Madison
 - Developed the Censored Autoregressive conditional Fréchet (AcF) model, an extension of the classic AcF, for modeling block maxima of negative log returns in stock markets with daily price limits.
 - Captured the dynamics of systemic risk by recovering time-varying GARCH-like tail index series using the CAcF model.
 - Applied the CAcF model to diverse stock markets including CSI300, S&P500, DJI30, SET100, revealing interpretable market relationships.
- **Sports Game Result Prediction using Latent Factor Model(LFM)** Jan. 2020 - May 2020
Application of LFM with Logit function and penalized loss function Zhejiang University
 - Utilized classical LFM on home team - away team interaction matrices, achieving average 59.7% prediction accuracy for NBA game results.
 - Extended the approach to three-dimensional tensor decomposition and introduced temporal factors for improved prediction accuracy.

SELECTED PROJECTS

- **Yelp Review Data Analysis** Nov. 2020 - Dec. 2020
Languages & Tools: Python, R, Git, Github, Pandas
 - Employed Natural Language Processing techniques (LDA, LSA, NMF) to extract review topics from extensive Yelp review data.
 - Quantified the correlation between ratings and reviews, constructing a multivariate regression model for rating prediction using Step-wise selection and cross-validation.
 - Generated data-driven business advice for related industries.
- **Spectrum Similarity Analysis for cB58 Galaxy** Oct. 2020
Languages & Tools: R, High Throughput Computing, Distributed Computing
 - Devised an efficient algorithm using statistical characteristics to identify spectra closely resembling the cB58 galaxy from a 25GB spectrum dataset containing 2.5 million spectra.
 - Utilized exponential smoothing for noise reduction and introduced Temporal Correlation Coefficient to evaluate spectral trend similarity.
- **Server Storage Prediction** Sept. 2020
Languages & Tools: R
 - Designed a robust Bayesian linear regression algorithm in R for predicting precise server storage capacity depletion times, outperforming ARIMA algorithm in speed.

AWARDS AND HONORS

- Aug. 2020** Academic Excellence Award, UW-Madison, Department of Statistics
- Oct. 2019** Hailiang First Class Scholarship (Ranked 2nd), Zhejiang University
- Oct. 2019** Academic First Class Scholarship, Zhejiang University
- Sept. 2018** Third Prize, China Undergraduate Mathematical Modeling Contest
- Sept. 2018** Outstanding Student Leaders, Zhejiang University