XINYU (JANET) ZHANG

✉ xinyu_zhang@ncsu.edu • ✆ (608) 504-0767• ✎ https://xinyuz1996.github.io

HIGHLIGHTS

- Ph.D. candidate in Statistics with expected graduation in May 2024
- Data Scientist Internship and Research Partnership with Meta Platforms, Inc.
- Strong publications on developing novel statistical methodology with topics on optimization, time series forecasting, reinforcement learning and machine learning
- Solid experience in statistical research and consulting
- Excellent interpersonal skills, oral presentation and written skills
- Seeking for full time opportunity starting May 2024, or internship opportunity starting Jan 2024

PROFESSIONAL EXPERIENCE

North Carolina State University, Raleigh, NC August 2019 to Present
Graduate Research Assistant

- Dissertation research on enhancing the sample efficiency of time series forecasting models using ML methods, with epidemiological application to nowcasting of Coronavirus data, reduced error by 3%~30% using 30%~60% less data
- Dissertation research on developing efficient zeroth-order algorithm for black-box global optimization
- Developed multi-objective disease control solution for COVID-19 using Reinforcement Learning with Shiny Application
- Consulted on tax analytics and automation research, specializing in missing data imputation and factor analysis

North Carolina State University & Meta Platforms, Inc., Raleigh, NC
Graduate Research Consultant September 2022 to August 2023

- Automated human-decision based reconciliation strategy with robust theoretical guarantee for short-term and long-term budget planning, reducing the time of each planning from weeks to hours for the Meta Infrastructure team
- Optimized the strategy (by Bootstrapping) and validated it with advanced time series models (ARNN, BSTS, Prophet, SARIMA, TBATS, etc.) against multiple criteria (MAE, RMSE and MAPE)
- Reduced MAE by 10% for short-term forecasting and 70% for long-term forecasting on real-world demand data

Meta Platforms, Inc., Menlo Park, CA May 2022 to August 2022
Research Data Scientist Intern

- Designed and implemented a Python API to streamline the sensitivity analysis of capacity planning, integrating time series, ML, backtesting, and SQL skills, slashing industrial-level analysis development time by up to several weeks
- Applied by Meta's infrastructure foundation team, hardware team, and strategy team

University of Wisconsin-Madison, Madison, WI August 2017 to May 2019
Graduate Researcher

- Established a clustering method for recurrent event based on change-point detection

- Provided consultancy to quantitative ecologist from National Wildlife Health Center to estimate the abundance of white-tailed deer by developing advanced mark and recapture models (Lincoln–Petersen, logistic regression, Bayesian analysis)

EDUCATION

North Carolina State University, Raleigh, NC
Ph.D. candidate, Department of Statistics, Expected May 2024

University of Wisconsin-Madison, Madison, WI
M.S. in Statistics, Department of Statistics, 2019

Nankai University, Tianjin, China
B.S. in Statistics, School of Mathematical Sciences, 2017

PUBLICATIONS & PRESENTATIONS

- Zhang, X., & Ghosh, S. (2023). ProGO: Probabilistic Global Optimizer. *Under review, arXiv:2310.04457*
- Zhang, X., & Ghosh, S. (2022). PaEBack: Pareto-Efficient Backsubsampling for Time Series Data. *Under review, arXiv:2210.15780*
- Wan, R., Zhang, X., & Song, R. (2021). Multi-objective model-based reinforcement learning for infectious disease control. In *Proceedings of the 27th ACM SIGKDD Conference* (pp. 1634-1644)
- Li, Q., Yao, K., & Zhang, X. (2020). A change-point detection and clustering method in the recurrent-event context. *Journal of Statistical Computation and Simulation*, *90*(6), 1131-1149
- Presentation on Long-Term and Short-Term Time Series Reconciliation. Meta Platforms, Inc., May 2023
- Presentation on Multi-Objective Reinforcement Learning for Infectious Disease Control with Application to COVID-19 Spread[J]. ACM Conference on Health, Inference, and Learning, Apr 2021
- Presentation on Abundance Estimation of White-tailed Deer in Shenandoah National Park. National Wildlife Health Center of the United States Geological Survey, Madison, Wisconsin, May 2019

GRANTS & AWARDS

- Meta Platforms, Inc: $60,000 Funding for Research on Time Series Forecasting Reconciliation, Sep 2022
- George Mason University: Best Narrative - Cherry Blossom Prediction Competition, May 2022
- National Training Program of Innovation and Entrepreneurship (China): First Prize for Undergraduates, May 2017
- China Undergraduate Mathematical Contest: Second National Prize, Sep 2015

TECHNICAL SKILLS

- Programming Skills: Python, R, SAS, SQL, LaTeX, JAGS
- Technical Skills: Bayesian, Causal Inference, Experimental Design, Hypothesis Tests, Machine Learning, Missing Data, Optimization, Precision Medicine, Reinforcement Learning, Statistical Research & Consulting, Time Series Forecasting