Guanghui Min

+1 (434) 409-3686 | guanghui@virginia.edu | guanghuimin.github.io

EDUCATION

University of Virginia — Charlottesville, VA
Ph.D. in Computer Science — August 2024 - Now

- Average GPA: 4.0 (out of 4.0)
- Research Interests: Graph Machine Learning, Diffusion Models.
- Advisor: Dr. Chen Chen

University of Michigan, Ann Arbor — Ann Arbor, MI
M.Sc. in Applied Statistics — September 2018 - May 2020

- Average GPA: 3.88 (out of 4.0)
- Courses: Natural Language Processing, Analysis of Time Series, Survival Time Analysis, Bayesian Modeling.

Wuhan University — Wuhan, CN
B.Sc. in Statistics — September 2014 - June 2018

- Average GPA: 3.7 (out of 4.0) | Major GPA: 3.76 (out of 4.0) (Ranking: 6/47)
- Courses: Applied Regression Analysis, Multivariate Statistical Analysis, Stochastic Process, Nonparametric Statistics, Mathematical Statistics, Optimization Theory and Methods.

PUBLICATIONS

Guanghui Min, Yinhan He, Chen Chen. Scaling Epidemic Inference on Contact Networks: Theory and Algorithms. *Proceedings of the 39th Annual Conference on Neural Information Processing Systems. (NeurIPS-25).*

Yinhan He, Wendy Zheng, **Guanghui Min**, Yunfan Wang, Patrick Soga, Zhenyu Lei, Yushun Dong, Chen Chen, Jundong Li. Towards Mechanistic Interpretability for Graph Foundation Models. *Submitted to KDD-25.*

Yinhan He, Chen Chen, Song Wang, **Guanghui Min**, Jundong Li. Demystifying Epidemic Containment in Directed Networks: Theory and Algorithms. *Proceedings of the 18th ACM International Conference on Web Search and Data Mining. (WSDM-25).*

PROJECTS

Asset Allocation using Regime-switching Hidden Markov Method — Beijing, CN
Innovation center, Yinhua Fund Management Co., Ltd — October 2020 – December 2020

- Investigated how to quantitatively characterize the assets performance rotation based on a regime-switching Hidden Markov Model other than the traditional Merrill Lynch Clock model.
- Incorporated 5 major asset classes: large-cap stocks, small-cap stocks, corporate bonds, government bonds, and gold. Conducted time-series clustering of market performance across different regimes based on the returns of these asset classes. Performed variable selection on the number of regimes using k-fold cross-validation and provided economic interpretations for the final model.
- Constructed an asset allocation portfolio based on the model combined with risk parity weight allocation and the portfolio is currently undergoing live testing and has demonstrated stable performance amidst the backdrop of significant global economic fluctuations with a 4.2% excess return in the recent year compared with the benchmark.

Coupled Mixed Model for Joint Genetic Analysis of Complex Disorders from Independently Collected Data Sets: Application to Alzheimer's Disease and Substance Use Disorder — Pittsburgh, PA
Summer Research, Carnegie Mellon University | Acknowledged by: Dr. Haohan Wang — July 2017 – August 2017

- Proposed a Bayesian maximum likelihood model on multi-correlated responses data whose design matrix has missing values and to make predictions using the same-distribution sampled data as much as possible.
- Applied ADMM algorithm for optimization of the loss function and prove the convergency by block coordinate method. The method outperformed other competing methods including the baseline linear mixed model, supported by box plots of the area under ROC curves (auROC) of identifying the SNPs that are jointly responsible for both phenotypes and for Phenotype 1.

Entity extraction on documents from the period of the ROC (1912-1949) — Wuhan, CN
Wuhan University Innovation Practice Project | Sponsored by Hanwang Technology Co., Ltd. — June 2016 – April, 2017

- Preprocessed 1774 articles from almost a century ago, developed an annotation tool and manually labeled the names of people, locations and organizations.
- Built the baseline by applying generic Conditional Random Field (CRF) algorithm to extract the three types entities from documents in Republic of China and got the accuracy, precision, recall rate and F1 score.
- Solved the optimization problem applying the bagging framework to increase the recall rate of organization by 7% and reduced the time for the model to perform entity extraction tasks on the test set by 14%.

ACADEMIC AWARDS

- **UVA Computer Science Scholarship**, 2024
- **Wuhan University Third-Class Scholarship (top 15%)**, *statistics*, 2016, 2017, 2018
- **Wuhan University Second-Class Scholarship (top 10%)**, *fundamental mathematics*, 2015
- **Wuhan University Freshman Scholarship**, *fundamental mathematics*, 2014

EXPERIENCE

Yinhua Fund Management Co., Ltd. Beijing, CN
Senior Machine Learning Engineer, Innovation Center May 2020- July 2024
Works include:

- Developed and applied machine learning-based market strategies, including unsupervised clustering of time series of market daily performance, and risk reduction and stock portfolios diversification through graph analysis.
- Applied Large Language Models (LLMs) and other generative pre-trained models to investment research and wealth management. Tried different fine-tuning methods for LLM, including Low-Rank Adaptation (LoRA) and P-tuning. Other generative pre-trained models for different modalities include Stable Diffusion(image), Whisper (speech to text), Vits (text to speech), So-Vits-Svc (singing voice conversion) and so on.
- Developed a modern active equity investment research system and related functional modules, such as simulated portfolios and periodic research reports, within the company.

Yinhua Fund Management Co., Ltd. Beijing, CN
Analyst Intern May 2019 - August 2019
Works include:

- Developed workflow for constructing single-factor effectiveness tests, factor selection, fundamental multi-factor model construction, and Brinson factor attribution.
- Preprocessed historical data for internal portfolios and performed profit and risk attribution using Barra multi-factor model (CNE5).
- Created a task-oriented dialogue chatbot using the Python-based Rasa framework, applied in the field of investment and research for question answering and communication purposes. Leveraged the pre-trained Bert-Chinese-Large model for natural language understanding (NLU) and intent classification, and utilized the Dual Intent Entity Transformer (DIET) for entity extraction.

CERTIFICATES AND HONORS

CFA (Chartered Financial Analyst) Level III Candidate

- Level II Passed Date: May 2023

Coursera Course Certificates

- Decision Making and Reinforcement Learning | Columbia University | December 2023
- Generative AI with Large Language Models | deeplearning.ai, AWS | August 2023
- The Unix Workbench | Johns Hopkins University | September 2017
- Machine Learning | Stanford University | November 2016

MCM Competition (Mathematical Contest in Modeling)

- Thesis: The effect of self-driving cars on the traffic flow in the bottleneck model.
- Honorable Mention | April 2017

Others

Fluent in: Python, R, MySQL, Oracle, Redis
Skills: Pytorch, ggplot2, MapReduce, Sklearn, Tensorflow