Xiaowei **Yin**

(+86) 13717654270 | 2021201831@ruc.edu.cn | yinvvvvv.github.io

Education

Renmin University of China *Beijing, China*

B.S. in Statistics *Sep. 2021 - Jun. 2025 (expected)*

- GPA: 3.77/4.0 (89.1), Major GPA: 3.86 (90.2)
- Relevant coursework: Mathematical Analysis | Higher Algebra | Probability Theory | Real Analysis | Mathematical Statistics | Functional Analysis | Optimization | Statistical Computing | Stochastic Processes | Nonparametric Statistics | Regression Analysis | Multivariate Statistical Analysis | Experimental Design | Time Series Analysis | Machine Learning | C Programming
- 2022,2023 - Academic Excellence Award. Renmin University of China

Research Experience

RLHF with unobserved confounder *May. 2024 - Present*

Supervised by Prof. Zhuoran Yang at Yale and Prof. Zhengling Qi at GWU

- Developed a preference model using a specialized regressor approach to address identification challenges arising from binary outcomes and discrete endogenous regressor, reflecting the assumption that human preferences are binary, allowing only two options.
- Applied advanced nonparametric instrumental variable (NPIV) methods, specifically AdversarialGMM, to estimate the causal relationship between actions and preference scores.
- Currently working on generating pessimistic NPIV estimates and theoretically establishing bounds on the suboptimality of the policies derived from the model.

Integrating Subgroup Identification with Adaptive Experimentation *Mar. 2024 - Sep. 2024*

Supervised by Prof. Waverly Wei at USC and Prof. Jingshen Wang at UC Berkeley

- Developed a pipeline for identifying subgroups in clinical trial data, applying multi-stage experiments to improve subgroup detection with stronger treatment effects.
- Explored subgroup identification methods (e.g., SIDES, GUIDE, causal tree) and selected an approach (TSMCP) for consistently identifying heterogeneous treatment effects (HTE) using continuous biomarkers.
- Implemented an R algorithm to identify HTE subgroups initially, optimizing treatment allocation in later stages to confirm the most effective subgroup.
- Conducted extensive simulations to evaluate accuracy of identifying subgroups with stronger treatment effects.

Enhancing the Performance of Multi-Label Classification Models for Knowledge Tagging *Dec. 2023 - Feb. 2024*

Supervised by Prof. Xing Wang at Renmin University

- Leveraged BERT and GAT in a multi-label classification model to effectively tag knowledge points in real-world question data.
- Conducted exploratory data analysis on training data and model outputs, developing error matrices to pinpoint performance issues and utilizing Jaccard similarity as a refined evaluation metric.
- Identified label annotation errors and applied unsupervised and semi-supervised learning techniques to resolve this data quality issues, enhancing model performance.

Predicting the Secondary Market through Graph Neural Networks *May. 2023 - Dec. 2023*

Supervised by Prof. Wenbing Huang at Renmin University

- Developed a GNN-based model to predict implied volatility surfaces in option order books using bid-ask prices, contract details, and minute-level time series data, capturing complex interactions and market dynamics.
- Designed the GATDropModel by integrating APPNP (Approximate Personalized Propagation of Neural Predictions) with GAT to enhance node relationship capture and classification accuracy.
- Achieved high-precision implied volatility predictions, demonstrating model effectiveness.

Service

Statistical Investigation Association of Renmin University *Sep. 2022-Jun. 2023*

- Taught R programming to beginners within the university and promoted activities for the association.

Skills & Hobbies

Programming	R, Python, C/C++
Documentation	LaTeX, R markdown
Languages	Chinese(native), English(TOEFL 104)
Hobbies	Hiking, Running, Guitar, Watercolor Painting