

TAO JIN

Email: taoj@outlook.com Links: Homepage: tao-j.me, LinkedIn, Google Scholar, GitHub: tao-j

EDUCATION

University of Virginia
Ph.D. in Computer Science

May 2025

Zhejiang University
B.S. in Computer Science

June 2017

EXPERIENCE

Applied Scientist *Amazon*

July 2024 - Now

- Designed and implemented an agentic workflow to automate dispute processing, including contact handling, document retrieval and creation, case assembly, and risk evaluation.
- Research and apply efficient training and data selection methods for fraud detection models.
- Integrate few-shot learning techniques with large language models to enhance fraud detection performance.

Applied Scientist Intern *Trinity AI*

Jan. 2024 - May 2024

- Improved recall for sensitive-information protection by 30% using open large language models (LLMs), prompt engineering, and efficient fine-tuning on a small dataset.

Applied Scientist Intern *Amazon*

June 2023 - Aug. 2023

- Trained and deployed sequential recommendation transformer models that improved recall by 8% in offline benchmarks and increased click-through rate by 5% in A/B tests.
- Analyzed model performance across sizes (10M–100M parameters) and training data scales (5M–200M samples), and designed deployment strategies under compute and latency constraints.
- Profiled and optimized the training data pipeline, reducing data preparation time by 90%.

Graduate Research Assistant *University of Virginia*

Feb. 2018 - July 2024

- Focused on efficient estimation of rankings and utilities from noisy pairwise preference data collected from multiple information sources.
- Developed adaptive algorithms for active ranking under these preference models.
- Designed adaptive learning strategies for noisy ranking in multi-armed bandit and dueling bandit settings, achieving up to 40% performance gains over baseline methods in benchmarks.

Computer Vision Scientist *Momenta Technology Ltd.*

Feb. 2017 - Feb. 2018

- Led a team of 4 engineers and 8 data operations interns.
- Developed traffic light and traffic sign detection and tracking algorithms for autonomous driving using an iterative data collection and model training cycle.

Coordinated with the data operations team, defined labeling guidelines, and optimized collection and management processes for million-scale datasets; delivered models trained on this data after quality control and cleaning.

Deployed traffic light and traffic sign detection and tracking to road-test vehicles, enabling the decision module to achieve near 100% correctness in closed-road test environments.

PUBLICATIONS (* INDICATES EQUAL CONTRIBUTION)

- Tao Jin***, Yue Wu*, Quanquan Gu, Farzad Farnoud: *Ranking with Multiple Oracles: From Weak to Strong Stochastic Transitivity* [ICML'25]
- Yue Wu, **Tao Jin***, Qiwei Di*, Farzad Farnoud, Quanquan Gu: *Borda regret minimization for generalized linear dueling bandits* [ICML'24]
- Qiwei Di*, **Tao Jin***, Yue Wu, Farzad Farnoud, Quanquan Gu: *Variance-Aware Regret Bounds for Stochastic Contextual Dueling Bandits* [ICLR'24]
- Hao Lou, **Tao Jin**, Yue Wu, Pan Xu, Farzad Farnoud, Quanquan Gu: *Active Ranking without Strong Stochastic Transitivity* [NeurIPS'22]
- Yue Wu*, **Tao Jin***, Hao Lou, Pan Xu, Farzad Farnoud, Quanquan Gu: *Adaptive Sampling for Heterogeneous Rank Aggregation from Noisy Pairwise Comparisons* [AISTATS'22]
- Tao Jin***, Pan Xu*, Quanquan Gu, Farzad Farnoud: *Rank Aggregation via Heterogeneous Thurstone Preference Models* [AAAI'20, Oral]
- Chenghao Liu, **Tao Jin**, Steven Hoi, Jianling Sun, Peilin Zhao: *Collaborative Topic Regression for Online Recommender Systems: An Online and Bayesian Approach* [ACML'16, Machine Learning Journal]