

ZHIWEI XU

Email: zhiweixu@umich.edu | Tel: (+1) 616-227-1677 | Web: zhiweixx.github.io

RESEARCH INTERESTS

My research interests focus on understanding and improving the length and compositional generalization, and reasoning abilities of Large Language Models.

Previously, I have spent time on statistical inference and deep learning theory, where I was fortunate to collaborate with many great senior mentors and talented peers.

EDUCATION

University of Michigan

Ph.D. student, *Department of Statistics*

Advisor: Prof. Wei Hu and Prof. Yixin Wang

Ann Arbor, U.S.

2022 - present

University of Science and Technology of China

Bachelor of Science, *Department of Statistics*

Hefei, China

2018 - 2022

PUBLICATIONS AND PREPRINTS

[6] Recommendations Beyond Catalogs: Diffusion Models for Personalized Generation

Gabriel Alfonso Patron Herrera, **Zhiwei Xu**, Ishan Kapnada, Felipe Maia Polo

Submitted, 2025

[5] [Benign Overfitting in Single-Head Attention](#)

Roey Magen*, Shuning Shang*, **Zhiwei Xu**, Spencer Frei, Wei Hu, Gal Vardi

Submitted, 2025

[4] [Let Me Grok for You: Accelerating Grokking via Embedding Transfer from a Weaker Model](#)

Zhiwei Xu*, Zhiyu Ni*, Yixin Wang[◇], Wei Hu[◇]

International Conference on Learning Representations (**ICLR**) 2025

[3] [Benign Overfitting and Grokking in ReLU Networks for XOR Cluster Data](#)

Zhiwei Xu, Yutong Wang, Spencer Frei, Gal Vardi, Wei Hu

International Conference on Learning Representations (**ICLR**) 2024

[2] [Inference of Dependency Knowledge Graph for Electronic Health Records](#)

Zhiwei Xu, Ziming Gan, Doudou Zhou, Shuting Shen, Junwei Lu*, Tianxi Cai*

Under revision at Journal of the Royal Statistical Society, Series B

[1] [ARCH: Large-scale Knowledge Graph via Aggregated Narrative Codified Health Records Analysis](#)

Ziming Gan*, Doudou Zhou*, et al.

Journal of Biomedical Informatics, 2025+

(Note: authors with * contributed equally to the work)

EXPERIENCES

Simons Institute, UC Berkeley

Visiting Student Researcher

Program: [Modern Paradigms in Generalization](#)

Hybrid

Aug - Nov.2024

Cubist Systematic Strategies, LLC

Quantitative Research Intern

New York

Jun - Aug.2024

- Pretrained and fine-tuned LLMs on financial event transcripts to generate tradable signals.

Harvard University

Visiting student, *Department of Biostatistics*
Advisor: Prof. Junwei Lu and Prof. Tianxi Cai

Remote
May.2021 - May.2022

SELECTED AWARDS AND HONORS

Graduate Student Research Competition Award at IMPS 2024	<i>July 2024</i>
Outstanding GSI (teaching) Award	<i>April.2024</i>
Outstanding Graduates at USTC	<i>May.2022</i>
Gold Prize for Outstanding Scholarship at USTC	<i>Oct.2019, 2020</i>

TEACHING EXPERIENCE

STATS 500: Statistical Learning I Graded homework and conducted office hours	<i>Fall 2024</i>
STATS 413: Applied Regression Analysis Taught weekly lab classes, graded homework, and conducted office hours	<i>Winter 2024</i>
STATS/DATASCI 451: Bayesian Data Analysis Graded homework and conducted office hours	<i>Fall 2023</i>
STATS 250: Introduction To Statistics and Data Analysis Taught weekly lab classes, graded homework and lab, and conducted office hours	<i>Winter 2023</i>
Real Analysis (at USTC) Graded homework and conducted office hours	<i>Spring 2021</i>

TALKS & PRESENTATIONS

NeurIPS M3L Workshop (oral presentation on publication [5])	<i>Dec.2024</i>
MSSISS (oral presentation on publication [3])	<i>Mar.2024</i>
NeurIPS M3L Workshop (oral presentation on publication [3])	<i>Dec.2023</i>
Joint Statistical Meetings (poster presentation on preprint [2])	<i>Aug.2023</i>

PROFESSIONAL EXPERIENCE

Conference Reviewer of ICML 2024, ICLR 2024, ICML 2025

SKILLS

PyTorch, JAX, Python, R, C, SQL