

# ZHIWEI XU

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

## RESEARCH INTERESTS

---

My research interests focus on designing data-efficient Transformer architectures and improving the length generalization 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

---

[7] Terminal-Bench: A Benchmark for Agents on Hard, Realistic Tasks in Command Line Interfaces

Mike Merrill, et al.

Submitted, 2025

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

Gabriel Patron, **Zhiwei Xu**, Ishan Kapnadak, Felipe Maia Polo

Submitted, 2025

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

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

Neural Information Processing Systems (**NeurIPS**), 2025

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

**Zhiwei Xu**\*, Zhiyu Ni\*, Yixin Wang<sup>◊</sup>, Wei Hu<sup>◊</sup>

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\*

Journal of the Royal Statistical Society, Series B, 2025+ (To Appear)

[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

---

### Google Research

Student Researcher

Host: Hrayr Harutyunyan

*New York*

*May.2025 - Now*

**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*

**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  
Outstanding GSI (teaching) Award  
Outstanding Graduates at USTC  
Gold Prize for Outstanding Scholarship at USTC

*July 2024*  
*April.2024*  
*May.2022*  
*Oct.2019, 2020*

## TEACHING EXPERIENCE

---

**STATS 500:** Statistical Learning I  
Graded homework and conducted office hours

*Fall 2024*

**STATS 413:** Applied Regression Analysis  
Taught weekly lab classes, graded homework, and conducted office hours

*Winter 2024*

**STATS/DATASCI 451:** Bayesian Data Analysis  
Graded homework and conducted office hours

*Fall 2023*

**STATS 250:** Introduction To Statistics and Data Analysis  
Taught weekly lab classes, graded homework and lab, and conducted office hours

*Winter 2023*

**Real Analysis** (at USTC)  
Graded homework and conducted office hours

*Spring 2021*

## TALKS & PRESENTATIONS

---

NeurIPS **M3L** Workshop (oral presentation on publication [5])

*Dec.2024*

**MSSISS** (oral presentation on publication [3])

*Mar.2024*

NeurIPS **M3L** Workshop (oral presentation on publication [3])

*Dec.2023*

Joint Statistical Meetings (poster presentation on preprint [2])

*Aug.2023*

## PROFESSIONAL EXPERIENCE

---

Conference Reviewer of ICML 2024, ICLR 2025

## SKILLS

---

PyTorch, JAX, Python, R, C, SQL