Xingjian Zhang

jimmyzxj@umich.edu || Personal Website || LinkedIn || 734-604-4231

Education

University of Michigan, School of Information, *Advised by Prof. Qiaozhu Mei* — Sep 2022 – Present
Ph.D. in Information Science **GPA**: 3.89/4.00 — Ann Arbor, MI

University of Michigan, EECS Department — Sep 2020 – May 2022
B.S. in Computer Science (Minor in Mathematics) **GPA**: 3.90/4.00 — Ann Arbor, MI

Work Experience

Graduate Research Assistant - University of Michigan, Ann Arbor, MI — Sep 2022 – Present

- **Foundation Model for Scientific Innovation** [6, 9]
  - Curated a large-scale dataset of structured scientific workflow summaries using LLMs, enabling AI-driven analysis and benchmarking, such as new idea generation. [6]
  - Proposed a novel task that transforms low-dim visualizations into relevant textual content, enabling intuitive exploration and new content generation from complex textual datasets. [9]
- **Graph Machine Learning** [3–5, 7]
  - Conducted an in-depth analysis of how LLMs leverage graph structural information in prompts, revealing that performance gains stem from contextual cues related to node labels rather than the graph structure. [3]
  - Developed a contributor-friendly platform that curates metadata-rich graph learning benchmarks to enhance dataset usability and incentivize contributions. [5]
  - Proposed a scalable method for approximating MNL likelihood from partial rankings, enabling efficient learning of choice-based network formation models without requiring complete temporal edge data. [4]

Research Intern - Google DeepMind, Mountain View, CA — May 2023 – Aug 2023

- **Efficient Transformer Fine-tuning and Serving** [8]
  - Proposed a GPU/TPU-friendly adaptive layer-skipping algorithm based on input sequence complexity.
  - Leveraged the correlation between loss at an early stage and the effective predictive depth of an input sequence.
  - Achieved the same performance with only 60% computation cost on MovieLens25M and ListOps.
  - Delivered the proposed algorithm in TensorFlow Recommenders (`TFRS`).

Deep Learning Engineer Intern - Intel, Shanghai, China — May 2021 - Aug 2021

- **CPU-Based Computer Vision Benchmark**
  - Participated in developing Analytics Zoo, an open-source big data AI platform.
  - Implemented and benchmarked AI models for semantic segmentation and object detection with PyTorch.

Selected Awards

Rackham Doctoral Fellowship Sep 2022
Tang Junyuan Scholarship Oct 2021
James B. Angell Scholar Mar 2022
J. Wu & J. Sun Sunshine Scholarship Oct 2020

Technical Skills

- **Programming Language**: Python (Professional); C, C++, R, Matlab, Mathematica.
- **Framework and Tools**: TensorFlow, PyTorch, Lightning, Deep Graph Library (DGL), TensorFlow Recommenders.

Additional Experience

- **Workshop Organizer**: Graph Learning Benchmark@KDD'23, LLMs for Individuals, Groups, and Society@WSDM'24.
- **Teaching Assistant**: Mathematical Analysis (Honor level), Data Mining, Data-Driven Web Application Development.
- **Competition**: Foxconn AI Big Data Competition (Gold Medalist), The University Physics Competition (Silver Medalist).

Publications/Preprints

[1] Junwei Deng, Ting-Wei Li, Shiyuan Zhang, Shixuan Liu, Yijun Pan, Hao Huang, Xinhe Wang, Pingbang Hu, Xingjian Zhang, and Jiaqi W Ma. 2024. dattri: A Library for Efficient Data Attribution. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2024.

[2] Benhao Huang, Yingzhuo Yu, Jin Huang, Xingjian Zhang, and Jiaqi Ma. 2024. DCA-Bench: A Benchmark for Dataset Curation Agents. *arXiv preprint arXiv:2406.07275* (2024).

[3] Jin Huang, Xingjian Zhang, Qiaozhu Mei, and Jiaqi Ma. 2024. Can LLMs effectively leverage graph structural information: when and why. *Transactions on Machine Learning Research (TMLR)* (2024).

[4] Jiaqi Ma*, Xingjian Zhang*, and Qiaozhu Mei. 2022. Fast learning of MNL model from general partial rankings with application to network formation modeling. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (WSDM)*, 2022. 715–725.

[5] Jiaqi Ma*, Xingjian Zhang*, Hezheng Fan, Jin Huang, Tianyue Li, Ting Wei Li, Yiwen Tu, Chenshu Zhu, and Qiaozhu Mei. 2022. Graph learning indexer: A contributor-friendly and metadata-rich platform for graph learning benchmarks. In *Proceedings of the Learning on Graphs Conference (LoG)*, 2022. 7–1.

[6] Xingjian Zhang*, Yutong Xie*, Jin Huang, Jinge Ma, Zhaoying Pan, Qijia Liu, Ziyang Xiong, Tolga Ergen, Dongsub Shim, Honglak Lee, and others. 2024. MASSW: A New Dataset and Benchmark Tasks for AI-Assisted Scientific Workflows. *arXiv preprint arXiv:2406.06357* (2024).

[7] Xingjian Zhang, Shixuan Liu, Yixin Wang, and Qiaozhu Mei. 2024. Leveraging LLM-Generated Structural Prior for Causal Inference with Concurrent Causes. In *Proceedings of the Causality and Large Models Workshop at NeurIPS 2024*, 2024.

[8] Xingjian Zhang, Jiaxi Tang, Yang Liu, Xinyang Yi, Li Wei, Lichan Hong, Qiaozhu Mei, and Ed H Chi. 2024. Conditional Transformer Fine-Tuning by Adaptive Layer Skipping. In *Proceedings of the 5th Workshop on Practical ML for Limited/Low Resource Settings (ML4LowResource)*, 2024.

[9] Xingjian Zhang, Ziyang Xiong, Shixuan Liu, Yutong Xie, Tolga Ergen, Dongsub Shim, Hua Xu, Honglak Lee, and Qiaozhu Mei. 2024. Map2Text: New Content Generation from Low-Dimensional Visualizations. *arXiv preprint arXiv:2412.18673* (2024).