YUNXIANG ZHANG

Phone: (+1) 734-834-5492 ⋄ Email: yunxiang@umich.edu ⋄ Homepage: yunx-z.github.io

EDUCATION

University of Michigan – Ann Arbor *09/2022 – 04/2027 (Expected)*
Ph.D. in Computer Science and Engineering
Advisor: Prof. Lu Wang

University of Michigan – Ann Arbor *09/2022 – 04/2024*
M.S. in Computer Science and Engineering
Advisor: Prof. Lu Wang

Peking University *09/2018 – 07/2022*
B.S. in Geographic Information Science
Advisor: Prof. Xiaojun Wan

PUBLICATIONS

1. MLRC-Bench: Can Language Agents Solve Machine Learning Research Challenges?
   Yunxiang Zhang, Muhammad Khalifa, Shitanshu Bhushan, Grant D Murphy, Lajanugen Logeswaran, Jaekyeom Kim, Moontae Lee, Honglak Lee, Lu Wang. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025 (NeurIPS Datasets and Benchmarks Track 2025)*.

2. Logit Arithmetic Elicits Long Reasoning Capabilities Without Training
   Yunxiang Zhang, Muhammad Khalifa, Lechen Zhang, Xin Liu, Ayoung Lee, Xinliang Frederick Zhang, Farima Fatahi Bayat, Lu Wang. In *The 1st Workshop on Test-time Scaling and Reasoning Models (SCALR @ COLM 2025)*.

3. Skill-Aware Data Selection and Fine-Tuning for Data-Efficient Reasoning Distillation
   Lechen Zhang, **Yunxiang Zhang**, Wei Hu, Lu Wang. In *The 5th Workshop on Mathematical Reasoning and AI (MATH-AI @ NeurIPS 2025)*.

4. Small Language Models Need Strong Verifiers to Self-Correct Reasoning
   Yunxiang Zhang, Muhammad Khalifa, Lajanugen Logeswaran, Jaekyeom Kim, Moontae Lee, Honglak Lee, Lu Wang. In *Findings of the Association for Computational Linguistics: ACL 2024 (Findings of ACL 2024)*.

5. Merging Generated and Retrieved Knowledge for Open-Domain QA
   Yunxiang Zhang, Muhammad Khalifa, Lajanugen Logeswaran, Moontae Lee, Honglak Lee, Lu Wang. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*.

6. SituatedGen: Incorporating Geographical and Temporal Contexts into Generative Commonsense Reasoning
   Yunxiang Zhang, Xiaojun Wan. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023 (NeurIPS Datasets and Benchmarks Track 2023)*.

7. Investigating Zero- and Few-shot Generalization in Fact Verification
   Liangming Pan*, **Yunxiang Zhang*** and Min-Yen Kan. In *The 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AACL 2023)*.

8. MOVER: Mask, Over-generate and Rank for Hyperbole Generation
   Yunxiang Zhang, Xiaojun Wan. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2022).*

9. Interpreting the Robustness of Neural NLP Models to Textual Perturbations
   Yunxiang Zhang, Liangming Pan, Samson Tan, Min-Yen Kan. In *Findings of the Association for Computational Linguistics: ACL 2022 (Findings of ACL 2022).*

10. BiRdQA: A Bilingual Dataset for Question Answering on Tricky Riddles
    Yunxiang Zhang, Xiaojun Wan. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI 2022).*

PREPRINTS & SUBMISSIONS

1. Learning to Ideate for Machine Learning Engineering Agents
   Yunxiang Zhang, Kang Zhou, Zhichao Xu, Kiran Ramnath, Yun Zhou, Sangmin Woo, Haibo Ding, Lin Lee Cheong. In submission to *ACL Rolling Review (ARR).*

2. LLMs as Rules Oracles: Exploring Real-World Multimodal Reasoning in Tabletop Strategy Game Environments
   Joseph J Peper, Sai Krishna Gandra, **Yunxiang Zhang**, Vaibhav Chennareddy, Shloki Jha, Ali Payani, Lu Wang. In submission to *The Fourteenth International Conference on Learning Representations (ICLR 2026).*

3. LiveOIBench: Can Large Language Models Outperform Human Contestants in Informatics Olympiads?
   Kaijian Zou, Aaron Xiong, **Yunxiang Zhang**, Xinliang Frederick Zhang, Yueqi Ren, Jirong Yang, Ayoung Lee, Shitanshu Bhushan, Lu Wang. In submission to *The Fourteenth International Conference on Learning Representations (ICLR 2026).*

4. CC-Riddle: A Question Answering Dataset of Chinese Character Riddles
   Fan Xu, **Yunxiang Zhang**, Xiaojun Wan. In *ArXiv abs/2206.13778 (2022).*

INTERNSHIP EXPERIENCE

Applied Scientist Intern at Amazon AWS AI, Santa Clara   *May. 2025 – Aug. 2025*
Mentor: Kang Zhou, Haibo Ding
Topics: RL for MLE Agents

Research Intern at Microsoft Research, Montreal   *May. 2024 – Aug. 2024*
Mentor: Xingdi Yuan, Alessandro Sordoni, Tong Wang
Topics: reward model for LLM reasoning

HONORS & AWARDS

Rackham Conference Travel Grant, University of Michigan	2023-2025
Rackham Professional Development Grant, University of Michigan	2023
Best Lightning Talk, 3rd Data Science Student Conference (DISC) at MSU	2023
AAAI-22 Student Scholarship	2022
May 4th Scholarship, Peking University	2020
Robin Li Scholarship, Peking University	2019

PROFESSIONAL SERVICE

Organizer, 1st Workshop on Test-Time Scaling and Reasoning Models (SCALR), COLM 2025

Student Volunteer, ACL 2024

Program Committee Member / Reviewer: ICML (2025), AISTATS (2025), ICLR (2025), NeurIPS (2024-25), ACL 2024 KnowledgeNLP Workshop, ACL Rolling Review, COLM (2024-25), EMNLP (2022-24), ACL (2023), CoNLL (2023-24), ML4Health (2023), ACM Computing Surveys

TALKS

Merging Generated and Retrieved Knowledge for Open-Domain QA. EMNLP	Dec. 2023
Merging Generated and Retrieved Knowledge for Open-Domain QA. MSU	Nov. 2023
MOVER: Mask, Over-generate and Rank for Hyperbole Generation. NAACL	Jul. 2022

MENTORING

Lechen Zhang (MS @ UMich 25' → PhD @ UIUC), Xinnuo Li (BS @ UMich 25' → PhD @ UArizona), Yufan Wu (BS @ UMich & SJTU 24' → PhD @ OSU), Fan Xu (BS @ PKU 23' → PhD @ PKU)

SKILLS

Programming: Python, C/C++, SQL
Tools & Frameworks: Pytorch, Huggingface, NLTK, LaTeX, Git, Vim, Linux