XIAONAN LUO

Phone: (+1) 312-735-1751 ⋄ Email: `xluo6@nd.edu`

Homepage: `https://xiaonan-debug.github.io/`

Google Scholar ⋄ LinkedIn

EDUCATION

University of Notre Dame *May 2029 (expected)*
Ph.D. in Computer Science and Engineering
Advisor: Xiangliang Zhang

The Hong Kong University of Science and Technology (HKUST) *July 2024*
B.S. in Data Science and Computer Science

RESEARCH INTERESTS

I am broadly interested in Generative Foundation Models, focusing on **data-centric evaluation**, **fine-tuning**, and **reinforcement learning** of large language models.

EXPERIENCE

MINE Lab Jan 2025 - present
Supervisors: Prof. Xiangliang Zhang Notre Dame

· Study generative foundation models with particular focus on reasoning, data-centric evaluation, and reinforcement learning of LLMs.
· Study Scientific AI with particular focus on AI for chemistry and biomedical science.

HKUST System Lab Jan 2023 - Sep 2023
Supervisors: Prof. Wei Wang HKUST

· Built efficient serverless machine learning inference systems.

PUBLICATIONS

[1] **Xiaonan Luo**, Y. Huang, P. He, and X. Zhang, "Better datasets start from refinelab: Automatic optimization for high-quality dataset refinement," in *The 40th Annual AAAI Conference on Artificial Intelligence (AAAI)*, 2026.

[2] Y. Huang, Z. Jiang, **Xiaonan Luo**, *et al.*, "Chemorch: Empowering LLMs with chemical intelligence via groundbreaking synthetic instructions," in *The Thirty-ninth Annual Conference on Neural Information Processing Systems (Neurips)*, 2025.

[3] X. Wang, Y. Huang, **Xiaonan Luo**, *et al.*, "Adareasoner: Adaptive reasoning enables more flexible thinking," in *The Thirty-ninth Annual Conference on Neural Information Processing Systems (Neurips)*, 2025.

[4] M. Yu, A. Wang, **Xiaonan Luo**, *et al.*, "Torpor: Gpu-enabled serverless computing for low-latency, resource-efficient inference," in *Proceedings of USENIX Annual Technical Conference (ATC)*, 2025.

SELECTED PROJECT

Better datasets start from refinelab: Automatic optimization for high-quality dataset refinement

AAAI 2026

- Proposed RefineLab, an LLM-driven framework that formulates QA dataset refinement as an integer linear program, automatically selecting coverage, difficulty, and factual-consistency edits under a token-level budget to align datasets with explicit quality targets.
- Designed and evaluated modular refinement operators, demonstrating stable LLM evaluation on benchmarks such as MMLU, GSM8K, and HumanEval.

ChemOrch: Empowering LLMs with Chemical Intelligence via Synthetic Instructions

Neurips 2025

- Developed ChemOrch, a two-stage framework that synthesizes chemically grounded instruction-response data via task-controlled instruction generation and tool-aware response construction, leveraging decomposed RDKit/PubChem sub-tools and a difficulty reward model to produce diverse, executable chemistry tasks.
- Showed that ChemOrch-generated data both evaluates and improves LLM chemical intelligence, exposing weaknesses on fine-grained tasks and significantly boosting accuracy, reasoning, and tool-use performance after fine-tuning.

Learning to Perfuse: A Physiology-Grounded Digital Twin and Reinforcement Learning Framework for Graft Preservation

Under review

- Developed a physiology-grounded digital twin for ex vivo graft perfusion that jointly models hemodynamics, gas exchange, acid–base balance, metabolism, and thermal regulation, and cast multi-input perfusion control as a deep RL problem with compositional, constraint-masked actions.
- Trained a Double-DQN perfusion controller on this simulator and showed it extends graft survival significantly.

PROFESSIONAL EXPERIENCE

Meituan   2022 Summer
Software Engineer Intern   Beijing, China

· Implemented Meituan Network Automatic Platform (MNAP) for switch operation and maintenance.

SKILLS

Programming Languages	Python, C/C++
Machine Learning Tools	Pytorch, Tensorflow, Sklearn, Pandas, Numpy