

EDUCATION

The University of Minnesota, Twin Cities

Ph.D. program, Electrical and Computer Engineering

Sept. 2024—Now

Supervised by Prof. Mingyi Hong

University of Science and Technology of China

Graduate student in Electronic Engineering, School of Information Science and Technology

Sept. 2021—Jun. 2024

Supervised by Prof. Jian Li

B.Eng. in Artificial Intelligence and Electronic Engineering, School of the Gifted Young

Sept. 2017—Jun. 2021

Enrolled in AI Talent Class

RESEARCH INTEREST

My research focuses on Reinforcement Learning and its application to post-training and test-time scaling. My previous work includes designing efficient alignment algorithms (SPR) and aligning frozen LLM at inference time to achieve successive performance improvement.

PUBLICATIONS

- **Xinnan Zhang**, Chenliang Li, Siliang Zeng, Jiaxiang Li, Zhongruo Wang, Songtao Lu, Kaixiang Lin, Alfredo Garcia, Mingyi Hong. [Aligning Frozen LLMs by Reinforcement Learning: An Iterative Reweight-then-Optimize Approach](#) **In submission**
- **Xinnan Zhang**, Chenliang Li, Siliang Zeng, Jiaxiang Li, Zhongruo Wang, Songtao Lu, Kaixiang Lin, Alfredo Garcia, Mingyi Hong. [Reinforcement Learning in Inference Time: A Perspective from Successive Policy Iterations](#) **ICLR 2025 Reasoning and Planning for LLMs workshop**
- **Xinnan Zhang***, Siliang Zeng*, Jiaxiang Li*, Kaixiang Lin, Mingyi Hong. [LLM Alignment Through Successive Policy Re-weighting \(SPR\)](#), **Neurips 2024 FITML workshop**
- **Xinnan Zhang**, Jialin Wu, Junyi Xie, Tianlong Chen, Kaixiong Zhou. [Benchmark on Drug Target Interaction Modeling from a Structure Perspective](#) **In submission**
- Kun Wang*, Guibin Zhang*, **Xinnan Zhang**, Junfeng Fang, Xun Wu, Guohao Li, Shirui Pan, Wei Huang, Yuxuan Liang. [The heterophilic snowflake hypothesis: Training and empowering GNNs for heterophilic graphs](#), **KDD, 2024**
- **Xinnan Zhang**, Yuanbo Cheng, Xiaolei Shang, Jun Liu. CRB Analysis for Mixed-ADC Based DOA Estimation, **IEEE Transactions on Signal Processing, 2024**
- **Xinnan Zhang**, Yuanbo Cheng, Hing Cheung So, Jian Li. A novel mixed-ADC architecture for DOA Estimation, **Signal Processing Letter, 2024**
- **Xinnan Zhang**, Yuanbo Cheng, Xiaolei Shang, Jun Liu. [Optimal Mixed-ADC Arrangement for DOA Estimation via CRB using ULA](#), **ICASSP, 2023, poster**
- **Xinnan Zhang**, Yuanbo Cheng, Xiaolei Shang, Jun Liu. Mixed-precision Arrangement for DOA Estimation via CRB using SLA, **EUSIPCO, 2023, oral**
- Junyi Xie, Jiefeng Ma, **Xinnan Zhang**, Jianshu Zhang, Jun Du. Enhancing Math Word Problem Solving through Salient Clue Prioritization: A Joint Token-Phrase-Level Feature Integration Approach, **IALP, 2023, oral**

RESEARCH EXPERIENCE

The University of Minnesota, Twin Cities

Jun. 2024—now

Research Assistant

Advisor: Prof. Mingyi Hong

- Designed a compute- and memory-efficient alignment algorithm, Successive Policy Re-weighting (SPR), that loads only one model at a time; from a Llama-3 SFT baseline, SPR achieved iterative gains and reached 70.10% on the Open LLM Leaderboard after 5 iterations, outperforming Llama-3 DPO (69.88%) and rejection sampling (70.01%).
- Proposed Iterative Reweight-then-Optimize (IRO), a compute-efficient test-time alignment method that steers a frozen base model (no weight updates) via a three-step loop: (1) sample candidates, (2) resample with current value functions, (3) train a lightweight value model to guide the next decoding pass. Achieved SOTA on summarization and instruction following—on AlpacaEval 2.0, improved Llama-3-8B-Instruct from 30.71% → 43.80% and Meta-Llama-3-70B-Instruct from 43.11% → 49.77%, outperforming Best-of-N and weak-to-strong search.

Benchmark on Drug Target Interaction Modeling from a Structure Perspective

Sept. 2023—May. 2024

Remote Research Assistant

Advisor: Prof. Tianlong Chen, Prof. Kaixiong Zhou

- Integrate different methods into one framework to compare them fairly; Propose our method by utilizing both Transformer and GNN, which achieves SOTA on six datasets

The Heterophily Snowflake Hypothesis: Training and Powering GNN for Heterophilic Graphs

Sept. 2023— May. 2024

Remote Research Assistant, NUS

Advisor: Dr. Kun Wang

- Apply Snowflake Hypothesis on different GNN, such as MGNN, H2GCN, to the heterophily dataset; Apply Heterophily Snowflake Hypothesis on different GNN, such as MGNN, H2GCN, to the heterophily dataset.

University of Science and Technology of China

Jun. 2021— Nov. 2023

Research Assistant, USTC

Hefei, China

Advisor: Prof. Jian Li

- Analyze the placement of one-bit and high-precision ADCs via Cramér-Rao bound using a uniform linear array, and demonstrate the optimal mixed-ADC arrangement theoretically.
- Explore the placement problem to the sparse array with any geometries and convert the problem to a 0-1 combinatorial optimization. The problem is efficiently solved by using the ADMM algorithm.
- Study the placement of mixed-ADC problem systematically. Use a maximum likelihood-based method to estimate the parameter by modifying the Sparse Parameters via the Iterative Minimizing algorithm to mixed-ADC.
- Using FFT and advanced mathematical tools (GS factorization) to accelerate the IAA algorithm, including multiplication and inversion in the case of linear array and sparse array.

TAT-QA: A Question Answering Benchmark on a Hybrid of Tabular and Textual Content in Finance

Jun. 2020—Nov.2020

Research Intern

Advisor: Dr. Fengbin Zhu, Prof. Wengqiang Lei (Next++, National University of Singapore)

- Using the VUE program language to develop a data annotation system for a financial question answering dataset and maintaining the process; The data annotation platform further leads to ACL 2021 work.

Verb Duration Classification:Corpus Construction, Model Selection, and Empirical Investigation

Feb.2020—Jul.2020

Research Intern

Advisor: Yuanxin Xiang, Prof. Wengqiang Lei, Prof.Min-Yen Kan (Next++, National University of Singapore)

- Use Stanford CoreNlp to extract information from wiki data and Gigaword according to some specific syntactic roles and construct a verb duration classification corpus; Leverage the corpus to demonstrate that neural modeling is more effective than traditional feature-based classifiers for duration classification.

SKILL

Programming

Python, Matlab, C/C++, Linux

Tools

Verl, TRL, Tinker-code, Git, Pytorch, Latex.

SELECTED AWARDS AND FELLOWSHIPS

- Amazon Machine Learning Systems Fellowship 2024
- First-class Master's Scholarship of USTC 2021- 2023
- 2020 Microsoft Student Club Practice Space Outstanding Winner 2020
- Outstanding Student Scholarship in USTC 2019
- Endeavour Scholarships and Fellowships 2019
- Outstanding Student Officer 2018
- National Maths Competition,third class prize(Anhui) 2018

ACTIVITIES

- **Teaching assistant**, University of science and technology of China

Worked on undergraduate courses: *Signal and System*,*Mathematical equation B*,*Numerical analysis B*