

Yueyan Li

School of Artificial Intelligence

Beijing University of Posts and Telecommunications, P.R. China

Email: siriuslala@bupt.edu.cn | **Github:** <https://github.com/Siriuslala> | **Blog:** <https://siriuslala.github.io>

Education

Beijing University of Posts and Telecommunications

Fall 2024 - Present

Ph.D in Intelligent Science and Technology

Beijing University of Posts and Telecommunications

Fall 2020 - Fall 2024

BE in Electronic Information Engineering

- GPA: 3.7/4.0
- Beijing Outstanding (Bachelor) Graduate (top 5%)
- The 13th National College Students' Mathematics Competition (Beijing Division)/First Prize

Experience

Zhipu.AI, Inc

October 2023 – May 2024

Research Intern

Mentor: Xiao Liu

*(Phase I) Participated in the development of **GLM-4**, responsible for model fine-tuning and alignment work.*

- Trained and optimized the PRM (Process-Supervised Reward Model).
- Implemented the decoding algorithm with PRM-supervised sampling (Stepwise Generation).
- Trained and evaluated Math CritiqueLLM, a scoring model for mathematical tasks.

*(Phase II) Participated in the development of **AutoGLM**, a series of agents for autonomous control of digital devices.*

- Participated in the design of the Android Agent pipeline.
- Responsible for the maintenance of the icon library for Web/Android Agent.

Research

Beijing University of Posts and Telecommunications

Student

Supervised by Caixia Yuan and Xiaojie Wang

Mechanistic Interpretability

September 2024 - January 2025

- Purposed **circuit-tuning**, an algorithm for fine-tuning in a mechanistic way.
- Provided an analysis on training dynamics and revealed the self-organization mechanism behind learning.

Persona Dialogue

April 2023 - December 2024

- Designed a two-stage pipeline for persona dialogue based on BART and BERT.

Publications

1. **Yueyan Li**, Caixia Yuan, Xiaojie Wang. *Circuit-tuning: A Mechanistic Approach for Identifying Parameter Redundancy and Fine-tuning Neural Networks*. (Preprint)
2. Yifan Xu, Xiao Liu, Xinghan Liu, Zhenyu Hou, **Yueyan Li**, Xiaohan Zhang, Zihan Wang, Aohan Zeng, Zhengxiao Du, Zhao Wenyi, Jie Tang, Yuxiao Dong. *ChatGLM-Math: Improving Math Problem-Solving in Large Language Models with a Self-Critique Pipeline*. In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024.
3. Team GLM: Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Diego Rojas, Guanyu Feng, Hanlin Zhao, Hanyu Lai, Hao Yu, Hongning Wang, Jiadai Sun, Jiajie Zhang, Jiale Cheng, Jiayi Gui, Jie Tang, Jing Zhang, Juanzi Li, Lei Zhao, Lindong Wu, Lucen Zhong, Ming-yue Liu, Minlie Huang, Peng Zhang, Qinkai Zheng, Rui Lu, Shuaiqi Duan, Shudan Zhang, Shulin Cao, Shuxun Yang, Weng Lam Tam, Wenyi Zhao, Xiao Liu, Xiaoyu Xia, Xiaohan Zhang, Xiaotao Gu, Xin Lv, Xinghan Liu, Xinyi Liu, Xinyue Yang, Xixuan Song, Xunkai Zhang, Yi An, Yifan Xu, Yilin Niu, Yuantao Yang, **Yueyan Li**, Yushi Bai, Yuxiao Dong, Zehan Qi, Zhaoyu Wang, Zhenyi Yang, Zhengxiao Du, Zhen-Ping Hou and Zihan Wang. *ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools*. (Preprint)

Technical Skills

Languages: IELTS: 7.0

Frameworks: Pytorch

Developer Tools: Python, Matlab, C/C++

Personal

- I was a member in the art troupe of Beijing University of Posts and Telecommunications, specializing in **street dance (breaking)**. I won the second prize in the Beijing College Students Fitness Dance Competition (winter 2022).