

Xinyue Liu

Carnegie Mellon University

[Portfolio](#) [Github](#) [Google Scholar](#) [Email](#)

Education

Aug 2023	Carnegie Mellon University Master's of Computational Data Science CGPA: 4.04/4.33	Pittsburgh, USA
May 2025		
Sep 2019	Beijing University of Posts and Telecommunications B.E., Computer Science & Technology CGPA: 3.8/4.0	Beijing, China
Jun 2023		

Publications

- [C] **Customizing Large Language Model Generation Style using Parameter-Efficient Finetuning** [[🔗](#)]
Xinyue Liu, Harshita Diddee, Daphne Ippolito
17th International Natural Language Generation Conference [INLG 2024]

Research Projects

- Instruction Dataset Construction for Low-resource Text Style** Sep'24 - Present
Advisor: Daphne Ippolito
- > Developing an instruction dataset construction method that enables language models to learn specific writing styles from low-resource texts while maintaining instruction-following capabilities.
 - > Designing an instruction framework encompassing planning, writing, and editing tasks to efficiently create diverse synthetic training examples from limited source data.
- Experience-Based Coding Issue Solving** Sep'24 - Present
Advisors: Graham Neubig, Zora Zhiruo Wang
- > Developing a retrieval framework based on functional and contextual similarities to help agents learn from successfully solved issues in SWE-bench, enabling effective example reuse for new problems.
 - > Investigating embedding-based similarity approaches to capture comprehensive issue representations from patch content, related file content and location, problem descriptions, and solution trajectories.
- Parameter-Efficient Pretraining with Model Expansion** Jan'24 - Present
Advisors: Emma Strubell, Ananya Harsh Jha
- > Developed a function-preserving model expansion method that leverages trained small models to initialize larger architectures, eliminating the need for from-scratch pretraining.
 - > Implemented parameter-efficient training with LoRA and ReLoRA on expanded models, achieving comparable performance to full model pretraining while reducing computational costs.
- Customizing LLM Generation Style** Feb'24 - Jun'24
Advisor: Daphne Ippolito
- > Developed a parameter-efficient style customization method using LoRA, enabling LLMs to learn specific writing styles while mitigating content memorization through entity masking during training.
 - > Developed methods for merging style-trained and instruction-trained LoRA adapters, enabling models to generate text with specific writing styles while following user instructions.
- NL2Code Network Configuration Audit System (Undergraduate Thesis)** Feb'23 - Jun'23
Advisor: Haifeng Sun
- > Developed a natural language to code translation system for network configuration verification, enabling automated auditing through translation of user instructions into executable Batfish commands.

Teaching

11-667 | **Large Language Models: Methods and Applications** *Teaching Assistant*

Fall 2024

- > Designed and graded assignments and exam questions.
- > Held office hours and answered questions on Piazza.

Industry Experience

Jun 2024	Walmart Global Tech	Bentonville, USA
Aug 2024	<i>NLP Engineer Intern</i>	
	Designed and implemented a Product2Vec model that determines substitute or complementary product relationships for assortment planning.	
Jun 2023	Ding Intelligent Technology	Beijing, China
Aug 2023	<i>NLP Engineer Intern</i>	
	Developed an AI assistant that can summarize long meeting transcripts with multiple speakers.	

Awards and Grants

China National Scholarship granted by the Ministry of Education of P.R.C.

Outstanding Graduates granted by the Ministry of Education of Beijing

Pioneer of Erudition granted by Beijing University of Posts and Telecommunications

First-Class Scholarship granted by Beijing University of Posts and Telecommunications

Merit Student granted by Beijing University of Posts and Telecommunications