Ollie Liu

https://ollieliu.com/

Contact Information

Salvatori Computer Science Center
Los Angeles, CA, 90089

olliezliu
me@ollieliu.com

Research Interests

I am interested in multimodal foundation models (FM). My current research foci involve exploring and understanding their potential in facilitating scientific discovery. I am particularly excited about:

- Designing new architectures and algorithms that are broadly applicable to scientific modalities, such as (meta)genomics, protein, multiphysics, chemistry, and material sciences.
- Expanding methods in LLM post-training – such as preference learning and representation learning – to elicit controllable and interpretable behavior from multimodal scientific FMs.

Education

University of Southern California 08/2022 - *Present*
Ph.D. in Computer Science
Advisor(s): Dani Yogatama, Willie Neiswanger

Northwestern University 09/2019 - 06/2021
Master of Science in Industrial Engineering and Management Sciences
Advisor(s): Jorge Nocedal

Carnegie Mellon University 08/2013 - 08/2019
Master of Science in Machine Learning
Bachelor of Science in Statistics and *Mathematics* (with Honors)

Publications

$\star$: equal contribution; α-β: equal contribution, alphabetical order

IsoBench: Benchmarking Multimodal Foundation Models on Isomorphic Representations
Deqing Fu$^{\alpha\text{-}\beta}$, Ruohao Guo$^{\alpha\text{-}\beta}$, Ghazal Khalighinejad$^{\alpha\text{-}\beta}$, **Ollie Liu**$^{\alpha\text{-}\beta}$, Bhuwan Dhingra, Dani Yogatama, Robin Jia, Willie Neiswanger
In: *Proceedings of COLM 2024*. [pdf][website]

DeLLMa: A Framework for Decision Making Under Uncertainty with Large Language Models
Ollie Liu*, Deqing Fu*, Dani Yogatama, Willie Neiswanger
In: *arXiv Preprint*. [pdf][website]

On Retrieval Augmentation and the Limitations of Language Model Training
Ting-Rui Chiang, Xinyan Velocity Yu, Joshua Robinson, **Ollie Liu**, Isabelle Lee, Dani Yogatama
In: *Proceedings of NAACL 2024 (short)*. [pdf]

Interpretable Diffusion via Information Decomposition
Xianghao Kong*, **Ollie Liu***, Han Li, Dani Yogatama, Greg Ver Steeg
In: *Proceedings of ICLR 2024*. [pdf][code]

Approximating CKY with Transformers
Ghazal Khalighinejad , **Ollie Liu**, Sam Wiseman
In: *Proceedings of Findings of EMNLP 2023*. [pdf][code]

How Does GPT-2 Compute Greater-Than?: Interpreting Mathematical Abilities in a Pre-Trained Language Model
Michael Hanna, **Ollie Liu**, Alexandre Variengien

In: *Proceedings of NeurIPS 2023*. [pdf]

Positions

New York University & Polymathic AI, Research Scientist — Fall 2024
Microsoft Research, AI Frontiers, Research intern — Summer 2024
Reka AI, Research intern — Summer 2023
Redwood Research, Research Resident — Winter 2023
Meta Reality Labs, Assistant Technologies, Applied Research Intern — Summer 2022
You.com, Software Engineer Intern — Spring 2022
Amazon Web Services, GenAI Innovation Center, Applied Scientist Intern — 2021

Honors and Awards

University Organizer Fellowship, Open Philanthropy, 2024
Provost's Fellowship, University of Southern California, 2022
Data Science Fellowship, Northwestern University, 2019
Royal E. Cabell Fellowship, Northwestern University, 2019
Senior Leadership Recognition, Carnegie Mellon University, 2018

Invited Talks

USC Information Science Institute NLG Seminar, DeLLMa: A Framework for Decision Making Under Uncertainty with Large Language Models, [video]

Teaching Experiences

TA, Machine Learning (Master), University of Southern California, Fall 2023
Co-Instructor, Machine Learning (Ph.D Elective), Northwestern University, Fall 2021
Co-Instructor, Mathematical Statistics (Ph.D Core), Northwestern University, Fall 2020
TA, Introduction to Machine Learning (Master), Carnegie Mellon University, 2 Semesters.
TA, Principles of Computing (Undergraduate), Carnegie Mellon University, 4 Semesters.

Doctoral Courseworks

Machine Learning, Natural Language Processing, Computer Vision, Learning Theory, Scalable Learning Systems, Theoretical Optimization

Services and Activities

Reviewer, ICML (2023, 2024); NeurIPS (2024); ACL Rolling Review (2024)
President, AI Safety Group, University of Southern California
Student organizer, Center for Optimization and Statistical Learning, Northwestern University

Skills

Software: Python (JAX, PyTorch, Hugging Face), Linux, R, LaTeX
Language: Chinese (*native*), English (*proficient*, GRE V169+Q168)