Ollie Liu

🌐 https://ollieliu.com/

Contact Information	Salvatori Computer Science Center Los Angeles, CA, 90089	🐦 olliezliu ✉ me@ollieliu.com

Research Interests

I'm passionate about designing multimodal foundation models that accelerate scientific discovery. Towards this, my research foci include foundation model pre-training, understanding and exploring their capabilities, as well as continual training to align them with human desiderata. By focusing on these areas, I aim to develop trustworthy models that not only advance artificial intelligence but also serve practical purposes in various scientific domains.

Education

University of Southern California — 08/2022 - *present*
Ph.D in Computer Science
Advisors: Dani Yogatama, Willie Neiswanger

Carnegie Mellon University — 08/2013 - 08/2019
B.S + M.S in Machine Learning
Magna Cum Laude

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: *arXiv Preprint*. [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

Microsoft Research, Research intern with Corby Rosset		Summer 2024
Reka AI, Research intern with Qi Liu		Summer 2023
Redwood Research, Research Resident with Alexandre Variengien		Winter 2023

Meta Reality Labs, Applied Research Intern with Stan Peshterliev	Summer 2022
You.com, Software Engineer Intern with Saahil Jain	Spring 2022
Amazon Web Services, Applied Scientist Intern with Mehdi Noori	2021
Northwestern University, Research Associate with Jorge Nocedal	2019 - 2021

Honors and Awards	**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, [vi deo]
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
Activities	**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)