

Michael Li

Homepage: ml5885.github.io
Email: ml6@andrew.cmu.edu

EDUCATION	Northeastern University <i>Ph.D., Computer Science, incoming</i> Advisors: David Bau & Byron Wallace.	2026–Present
	Carnegie Mellon University <i>B.S., Statistics & Machine Learning</i> Dean’s List with High Honors in all semesters.	2022–2026
RESEARCH	Carnegie Mellon University <i>Language Technologies Institute</i> Mentor: Nishant Subramani Research on interpreting language models, with a focus on understanding how their internal representations relate to model behavior.	2025–Present
	Carnegie Mellon University <i>Institute for Strategy & Technology</i> Advisor: Hannah Bailey Developed an embedding-based method for measuring agenda-setting power in U.S. state legislation and political social media.	2025–2026
	University of Victoria <i>Department of Mathematics & Statistics</i> Advisor: Xuekui Zhang Developed TCN, CNN, and LSTM forecasting models for U.S. COVID-19 trends using demographic and mobility data.	2022–2023
AWARDS	Best Paper, LTI Student Research Symposium, Preliminary Track	2026
	1 st Place, Alameda County Science Fair	2021, 2022
	Presenter, California Science Fair	2021, 2022
	U.S. Open Music Competition Gold Medal, Showcase Piano Senior Division	2021, 2022
INTERNSHIPS	Google: <i>Software Engineering Intern</i> Developed data pipelines and trained a GNN autoencoder for anomaly detection on Google’s B4 network.	Summer 2025
	Epirus: <i>Machine Learning Intern</i> Developed reinforcement-learning policies for heterogeneous drone swarms. Built a vectorized environment and training pipeline that reduced training time from days to minutes.	Summer 2024
	Beaver Health: <i>Software Engineering Intern</i> Led development of generative-AI dialogue systems for evidence-based dementia interventions and implemented multimodal data and conversation infrastructure.	Summer 2023
PUBLICATIONS	[Link] Michael Li , Nishant Subramani. “How Much Do Circuits Tell Us? Measuring the Consistency and Specificity of Language Model Circuits,” Oral presentation at the SciFM Workshop at COLM, 2026.	
	[Link] Michael Li , Nishant Subramani. “Model Internal Sleuthing: Finding Lexical Identity and Inflectional Features in Modern Language Models,” ACL, 2026.	
	[Link] Michael Li , Nishant Subramani. “BERTology in the Modern World,” INTERPLAY Workshop at COLM, 2025.	
	[Link] Michael Li , Fatemeh Esfahani, Li Xing, Xuekui Zhang. “Predicting the Daily Counts of COVID-19 Infection Using Temporal Convolutional Networks,” Journal of Global Health, 2023.	

LEADERSHIP	Carnegie AI Safety Initiative 2024–Present <i>Technical Reading Group Co-Lead</i> Co-lead weekly discussions on RLHF, scalable oversight, mechanistic interpretability, robustness, technical governance, and AI control.
PROJECTS	CLaiM: <i>Cal Hacks 11.0</i> Hyperbolic Prize (\$1,000) Automated natural-disaster insurance-claim processing using SAM 2 and YOLOv8. Shipworthy: <i>HackItShipIt</i> 1st Place Used OpenCV to turn a physical steering wheel into a real-time controller for a Unity sailing game. Stance: <i>Data Day Grind</i> Best Visualization Built and interpreted toxic-comment classifiers using scikit-learn and LIME. Ad Lunam: <i>To the Moon and Hack</i> 3rd Place Developed a virtual-reality physics simulation with procedurally generated planets in Unity.