Manqing Liu

Boston, MA, 02116
manqingliu@g.harvard.edu
manqingliu.github.io

Summary

I am currently a 4th year PhD Candidate at Harvard studying causal machine learning. My research interests include using deep learning models to estimate causal effects, and how causality can be used to improve the **reasoning** capabilities and **safety** of multimodal LLMs. I am co-advised by Dr. Andrew Beam and Dr. James Robins. I am also a member of the Causal Lab.

Education

2021–Present **Ph.D. in Causal Machine Learning**, *Harvard University*

2022–Present **Secondary field in Computer Science and Engineering**, *Harvard University*

2021–Present **M.Sc in Biostatistics**, *Harvard University*

2017-2020 **Post-Baccalaureate Studies in Maths/Statistics**, *University of Pennsylvania*

2014–2016 **MHS in Epidemiology**, *Johns Hopkins University*

Relevant Coursework

Causality	Advanced Epidemiologic Methods, Models for Causal Inference
Maths and Statistics	(MIT) Matrix Methods in Data Analysis & Signal Processing, (MIT) Introduction to Functional Analysis, Probability, Statistical Inference, Advanced Regression and Statistical Learning, Bayesian Inference
Computer Science	Systems Development for Computational Science, High Performance Computing for Science and Engineering, Stochastic Methods for Data Analysis, Inference and Optimization
Machine Learning	(MIT) Machine Learning, (MIT) Quantitative Methods for NLP, Deep Learning for Biomedical Data, Geometric Methods for Machine Learning, Algorithms for Data Science

Research Experience

Feb 2025 – Present **Projected Causal Alignment: Scaling Boundless DAS for LLMs**
Optimizing Boundless DAS through Johnson-Lindenstrauss transformations, significantly reducing computational overhead while enabling cross-scale analysis of causal mechanisms in LLMs. This approach investigates the neural representations of counting and arithmetic operations, exploring why mathematical reasoning deficiencies persist even as models scale to larger parameter counts and demonstrate enhanced capabilities in other reasoning domains.

Oct 2024 – Jan 2025 **Doubly Robust MCTS for LLM reasoning**
Integrated doubly robust estimator into Monte Carlo Tree Search (MCTS), enabling large language models to perform complex, multi-step reasoning and planning with higher accuracy in real-world scenarios.

June 2023 – Dec 2024 **DAG aware Transformer**
Engineered a noval DAG-aware transformer model to precisely estimate causal effects, addressing foundational challenges in unifying causal effect estimation under various scenarios.

Publications

2025 **Doubly Robust Monte Carlo Tree Search**, *Liu M.*, Beam A. Under review at ICML. Available at: arXiv:2502.01672

2024 **DAG-Aware Transformer for Causal Effect Estimation**, *Liu M.*, Bellamy D., Robins J., Beam A. Causal Representation Learning workshop at NeurIPS 2024. Available at: arXiv:2410.10044

2022 **Development of Machine Learning Algorithms Incorporating Electronic Health Record Data, Patient-Reported Outcomes, or Both to Predict Mortality for Outpatients with Cancer**, Parikh R.B., Hasler J.S., Zhang Y., *Liu M.*, Chivers C., et al., *JCO Clinical Cancer Informatics*, 6.

2021 **Trajectories of Mortality Risk Among Patients with Cancer and Associated End-of-Life Utilization**, Parikh R.B., *Liu M.*, Li E., Li R., Chen J., *npj Digital Medicine*, 4(1):104.

2020 **Validation of a Machine Learning Algorithm to Predict 180-Day Mortality for Outpatients with Cancer**, Manz C.R., Chen J., *Liu M.*, Chivers C., Regli S.H., et al., *JAMA Oncology*, 6(11):1723-1730.

2019 **Assessment of Inpatient Time Allocation Among First-Year Internal Medicine Residents Using Time-Motion Observations**, Chaiyachati K.H., Shea J.A., Asch D.A., *Liu M.*, Bellini L.M., et al., *JAMA Internal Medicine*, 179(6):760-767.

Professional Experience

2024 summer **Technical AI safety Fellowship**, *AI safety student team*

Attended a 8-week reading group on AI safety, covering topics like neural network interpretability, learning from human feedback, goal misgeneralization in reinforcement learning agents, and eliciting latent knowledge.

Skills

Programming Languages	Python, C++, R, SAS, STATA
Libraries and Frameworks	PyTorch, Tensorflow, Pandas, NumPy
Others	Causal Inference, Machine Learning, Deep Learning