

AAHLAD MANAS PULI

✉ aahlad@nyu.edu ☎ (929) 312-7360 🌐 [aahladmanas](#) 📄 [Google Scholar](#)

My research centers on **adaptation**: building AI systems that generalize across environments and learn the mechanisms that remain useful when distributions, observations, or interventions change. As a Faculty Fellow at NYU, I work across foundation models, causal inference, and representation learning, with applications to clinical and scientific problems. My work appears at NeurIPS, ICLR, ICML, and EMNLP, and I received NYU Computer Science's Janet Fabri Prize for the outstanding dissertation.

RESEARCH

Foundation models & architectures. Controllably efficient language models [P1] • Transformer retrieval and training dynamics [P5] • Flow-map learning [P10] • Multimodal representation learning [P9]

Adaptation & learned inference. OOD generalization and shortcut learning [P6, P7, P8] • Meta-learning for causal estimation [P2] • Causal inference [P11, P12] • Interpretability [P14]

Health & scientific modeling. Transportable EHR-scale clinical prediction [P3] • Genome-wide association studies [P11] • Robust scientific discovery [P15]

EXPERIENCE

Ongoing clinical research — 2026 manuscript in preparation

- Led the experimental program and transportability analysis for an EHR-scale twinned survival study of coronary heart disease, demonstrating better risk prediction than the AHA-standard risk model [P3] (co-first author).

Center for Data Science, New York University

Faculty Fellow / Assistant Professor

New York, USA

Sep 2024 – Present

- **Controllably efficient language models:** Co-supervised CAT, an architecture that exposes an adjustable throughput knob in a single model to achieve a new Pareto frontier of performance–throughput trade-offs [P1].
- **Learned causal inference:** Co-developed BBCI, which frames causal effect estimation as dataset-level prediction, learning entire inference procedures for different causal problems [P2] (co-first author).
- **Adaptation and training dynamics:** Contributed to the development of equal-learning objectives for LLM retrieval, building on my margin-control method [P6, P5], and used them to improve transportability for risk models [P3].
- **Interpretability:** Proposed STRIPE-X, a scalable detector of explanation failures, and proved the main results; detected such failures in introspection traces from an 8B Llama model [P14].
- **Generative and multimodal learning:** Studied convergence for flow-map learning [P10]; contributed theory and experimental design for Symile, a multimodal alignment method that outperforms CLIP-based objectives by 12.5% [P9].
- Delivered the [NeurIPS 2024 tutorial](#) on spurious correlations and led the CDS capstone course (>100 students).

Courant Institute, New York University

PhD Researcher — Advisor: Rajesh Ranganath

New York, USA

Sep 2018 – Sep 2024

- Developed NURD, a method for out-of-distribution generalization with optimality guarantees [P7], and semantic corruptions, a data-augmentation approach that adjusts away *unknown* nuisances [P8].
- Developed *margin control* to mitigate shortcut learning that arises from gradient-based training with cross-entropy [P6].
- Developed General Control Functions to estimate causal effects under unobserved confounding using instrumental variables and variational autoencoders [P12].
- Proposed LODE, a scalable causal estimation algorithm to handle functionally specified confounding, and applied it to genome-wide association studies [P11].
- Helped adapt NURD to enable robust new-particle detection in particle collision data (e.g., LHC) [P15].

Adobe Research, Data Science Lab

Research Intern

San Jose, USA

Summer 2019

- Proposed a new Bayesian uplift modeling method for marketing attribution problems that makes use of order information.

Biomedical Informatics, Columbia University

Software Developer

New York, USA

Jul 2017 – Jun 2018

- Contributed to Columbia DBMI's data pipeline for the NIH *All of Us* research program on Google Cloud and BigQuery, enabling secure and scalable access to EHR data for >100 researchers.

EDUCATION

Ph.D. in Computer Science , New York University	2018–2024
M.S. in Computer Science , New York University	2015–2017
B.Tech. + M.Tech. in Electrical Engineering , Indian Institute of Technology Madras	2010–2015

SELECTED WORK

-
- [P1] **Controllably Efficient Language Models** — *In submission, 2026*
Jatin Prakash, **Aahlad Puli**, Rajesh Ranganath.
- [P2] **Black Box Causal Inference: Effect Estimation via Meta Prediction** — *arXiv:2503.05985, 2025*
Lucius E.J. Bynum*, **Aahlad Puli***, Diego Herrero-Quevedo, Nhi Nguyen, Carlos Fernandez-Granda, Kyunghyun Cho, Rajesh Ranganath. (* equal contribution)
- [P3] **Performant and Transportable Risk Models for Coronary Heart Disease from EHR Data** — *In preparation, 2026*
Aahlad Puli*, Shreyas Bhawe*, Mark Goldstein*, Mert Ketenci*, Pierre Elias, Sumit Chopra, Glenn Fishman, Stuart Katz, John Dodson, Yindalon Aphinyanaphongs, Shalmali Joshi, Noémie Elhadad, Adler Perotte, Rajesh Ranganath.
- [P4] **Let the Experts Speak: Improving Survival Prediction & Calibration via Mixture-of-Experts Heads** — *ML4H 2025*
Todd Morrill, **Aahlad Puli**, Murad Meghani, Soojin Park, Richard Zemel.
- [P5] **Learning Is Not A Race: Improving Retrieval in Language Models via Equal Learning** — *EMNLP 2025*
Wanqian Yang, **Aahlad Puli**, Rajesh Ranganath.
- [P6] **Don't Blame Dataset Shift! Shortcut Learning due to Gradients and Cross Entropy** — *NeurIPS 2023*
Aahlad Puli, Lily Zhang, Yoav Wald, Rajesh Ranganath.
- [P7] **OOD Generalization in the Presence of Nuisance-Induced Spurious Correlations** — *ICLR 2022*
Aahlad Puli, Lily Zhang, Eric Oermann, Rajesh Ranganath.
- [P8] **Nuisances via Negativa: Adjusting for Spurious Correlations via Data Augmentation** — *TMLR 2024*
Aahlad Puli, Nitish Joshi, Yoav Wald, He He, Rajesh Ranganath.
- [P9] **Contrasting with Symile: Simple Model-Agnostic Representation Learning for Unlimited Modalities** — *NeurIPS 2024*
Adriel Saporta, **Aahlad Puli**, Mark Goldstein, Rajesh Ranganath.
- [P10] **Flow Map Learning via Non-Gradient Vector Flow** — *ICLR 2026*
Mark Goldstein, Anshuk Uppal, Raghav Singhal, **Aahlad Puli**, Rajesh Ranganath.
- [P11] **Causal Estimation with Functional Confounders** — *NeurIPS 2020*
Aahlad Puli, Adler J. Perotte, Rajesh Ranganath.
- [P12] **General Control Functions for Causal Effect Estimation from IVs** — *NeurIPS 2020*
Aahlad Puli, Rajesh Ranganath.
- [P13] **X-CAL: Explicit Calibration for Survival Analysis** — *NeurIPS 2020*
Mark Goldstein*, Xintian Han*, **Aahlad Puli***, Adler J. Perotte, Rajesh Ranganath. (* equal contribution)
- [P14] **Explanations that Reveal All through the Definition of Encoding** — *NeurIPS 2024*
Aahlad Puli*, Nhi Nguyen*, Rajesh Ranganath. (* equal contribution)
- [P15] **Robust Anomaly Detection for Particle Physics Using Multi-Background Representation Learning** — *Machine Learning: Science and Technology, 2024*
Abhijith Gandrakota, Lily Zhang, **Aahlad Puli**, Kyle Cranmer, Jennifer Ngadiuba, Rajesh Ranganath, Nhan Tran.

AWARDS

Janet Fabri Prize — Outstanding PhD Dissertation, NYU Computer Science (2025)
Apple Scholars in AI/ML Fellowship (2022)
Trustworthy ML Rising Star (2021)
MacCracken Doctoral Fellowship (2018)

SELECTED INVITED TALKS

Out-of-Distribution Generalization: Shortcuts, Spuriousness, and Stability — Tutorial, NeurIPS 2024.
The Spectre of Spurious Correlations in OOD Generalization and Interpretability — Fermilab, 2026.
Making the Most of Your Data for Reliable AI — Columbia DBMI, UCLA, UC Irvine, UC Riverside, 2026.

ACADEMIC SERVICE

Organizer: Spurious Correlations, Invariance and Stability (SCIS) Workshop, ICML 2022–2023.
Reviewer: NeurIPS, ICLR, ICML, AISTATS, JMLR, TMLR, MLHC, UAI (Top Reviewer, UAI 2022).