Mark Goldstein

goldstein@nyu.edu
marikgoldstein.github.io

Education

NYU, Courant Institute, PhD Candidate (w/ Rajesh Ranganath and Thomas Wies) New York 2018 -

Harvard University, Special Student (mix of undergrad and PhD coursework) Cambridge 2016 - 2018

New England Conservatory of Music, Bachelor of Music Boston 2011 - 2015

Experience

Google DeepMind, Student Researcher (w/ Will Grathwohl, Arnaud Doucet) New York Summer 2024

NYU Langone Health, Non-traditional Volunteer (Radiology and Population Health) New York 2020 -

Apple, machine learning intern in health AI New York 2021-2022

RIKEN Center for Advanced Intelligence Project, Research Intern (w/ Emtiyaz Khan) Tokyo, Japan Summer 2019

MIT Brain and Cognitive Sciences, Research Assistant (w/ Pedro Tsividis, Josh Tenenbaum) Cambridge 2018

Teaching Assistant / Teaching Fellow

NYU, Machine Learning. Rajesh Ranganath. Spring 2022.
NYU, Machine Learning. Rajesh Ranganath. Spring 2021.
NYU, Deep Learning. Yann LeCun. Spring 2020.
NYU, Machine Learning. Rajesh Ranganath. Fall 2019.
Harvard, CS 181: Machine Learning. Finale Doshi-Velez and David Parkes. Spring 2021.$^{\star +}$
Harvard, CS 252: Programming Languages and Artificial Intelligence. Nada Amin. Fall 2020.$^{\dagger +}$
Harvard, CS 181: Machine Learning. Finale Doshi-Velez. Spring 2018.$^{\star +}$
Harvard, CS 281: Advanced Machine Learning. Sasha Rush. Fall 2017.$^{\star \dagger +}$
Harvard, CS 121: Intro to Theoretical CS. Boaz Barak and Salil Vadhan. Fall 2017.$^{+}$
Harvard, CS 181: Machine Learning. David Parkes and Sasha Rush. Spring 2017.$^{+}$
Harvard, CS 61: Systems Programming and Machine Organization. Margo Seltzer and Eddie Kohler. Fall 2016.$^{+}$

*Head Teaching Fellow, †Graduate Level, $^{+}$Harvard Distinction in Teaching Award

Additional TA'ing for Continuing Education classes

Harvard Extension School. COMPSCI 109A: Data Science 1: Introduction to Data Science 2020.
Harvard Extension School. CSCI E-83 Fundamentals of Data Science. 2017
Harvard Extension School. CSCI E-7 Introduction to Programming with Python. 2017
Harvard Extension School. CSCI E-63c. Elements of Data Science and Statistical Learning with R. 2016.
Harvard Extension School. COMPSCI 1: Great Ideas in Computer Science. 2016.

Invited Talks and Guest Lectures

October 2024. Pieter Abbeel's group (Berkeley Robot Learning Lab) at UC Berkeley. Continuous and Discrete Diffusion.

October 2024. Chi Jin's group at Princeton. Continuous and Discrete Diffusion.

August 2024. Google DeepMind, Info theory group. Topic: Stochastic Interpolants and Mutual Info estimation.

July 2024. Google DeepMind, Kevin Murphy's group. Topic: Stochastic Interpolants.

Nov and Dec 2023. Columbia Decisions, Risk and Operations group. Topic: diffusions and stochastic interpolants.

Nov 2023. NYU Inference and Representations course, guest lecture. Topic: diffusions and stochastic interpolants.

Fall 2022. NYU, Yann LeCun's Deep Learning course, guest lecture. Fall 2022. Topic: diffusion models.

Fall 2022. Flatiron Institute's workshop on Sampling, Transport, and Diffusions. Topic: Auxiliary Variable Diffusion Models.

Workshop Organization

Co-organized the second iteration of Workshop on Spurious Correlations, Invariance, and Stability @ ICML 2023.

Co-organized Workshop on Spurious Correlations, Invariance, and Stability @ ICML 2022.

Invited Workshop Participation

Participated in the 2nd Flatiron workshop on Measure Transport, Sampling, and Diffusions in Dec, 2023.

Participated in the 1st Flatiron workshop on Measure Transport, Sampling, and Diffusions in Dec, 2022.

Awards and Fellowships

Fall 2024: NYU Henning Biermann Prize. This award honors the memory of Henning Biermann, a brilliant and much-loved Ph.D. student whose dedication to teaching, mentoring, and service enriched academic and extracurricular life for everyone in the department. The award is made to a Computer Science Ph.D. student who exemplifies this spirit through outstanding contributions to education or service to the department.

Fall 2021: selected as a recipient of the NeurIPS 2021 Outstanding Reviewer Award.

Fall 2018: MacCracken Fellow, NYU Graduate School of Arts and Sciences. Five years of PhD funding.

Mentoring

Abhipsha Das (Master's Thesis at NYU, 2024, on diffusions for text)

Shraddha Jain (current Master's student, 2024, research on VAEs/Diffusion Models)

Nina Mortensen (Masters Thesis at NYU, 2024, research on VAEs, now at Fauna Robotics)

Kyle Adams (Undergradate from University of Florida, 2022, research on survival analysis, now a math phd at UF)

Mentor for Women in Data Science Datathon, Cambridge, 2021.

Tutored at a few iterations of NYU AI School, meant for first and second year undergraduates in the NYC area as a first exposure to AI/ML, prioritizing the under-represented in ML.

Publications

Adriel Saporta, Mark Goldstein, Aahlad Puli, and Rajesh Ranganath. Contrasting with Symile: Simple Model-Agnostic Representation Learning for Unlimited Modalities.
Conference paper @ NeurIPS, 2024.

Nanye Willis Ma, Mark Goldstein, Michael Albergo, Nick Boffi, Eric Vanden-Eijnden, and Saining Xie. SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers (preprint version).
Conference paper @ European Conference on Computer Vision (ECCV), 2024.

Mark Goldstein, Raghav Singhal, and Rajesh Ranganath. What's the score? Automated Denoising Score Matching for Nonlinear Diffusions.
Conference paper @ International Conference on Machine Learning (ICML), 2024.

Mark Goldstein, Yifan Chen, Mengjian Hua, Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Probabilistic Forecasting with Stochastic Interpolants and Föllmer Processes.
Conference paper @ International Conference on Machine Learning (ICML), 2024.

Mark Goldstein, Michael Albergo, Nick Boffi, Rajesh Ranganath, and Eric Vanden-Eijnden. Stochastic interpolants with data-dependent couplings. **Spotlight Paper.**
Conference paper @ International Conference on Machine Learning (ICML), 2024.

Yuxuan Hu, Mark Goldstein, Rajesh Ranganath, and others. A dynamic risk score for early prediction of cardiogenic shock using machine learning (arxiv).
European Heart Journal: Acute Cardiovascular Care. 2024.

Hao Zhang, Mark Goldstein, Rajesh Ranganath, and others. QTNet: Predicting Drug-Induced QT Prolongation with Artificial Intelligence-Enabled Electrocardiograms.
Journals of the American College of Cardiology, Clinical Electrophysiology. 2023.

Mark Goldstein, Raghav Singhal, Rajesh Ranganath. Where to Diffuse, How to Diffuse and How to get back: Learning in Multivariate Diffusions.
Conference paper @ International Conference on Learning Representations. 2023.

Xintian Han, Mark Goldstein, Rajesh Ranganath. Survival Mixture Density Networks.
Conference paper @ Machine Learning for Healthcare Conference. PMLR, 2022.

Mark Goldstein, Jörn-Henrik Jacobsen, Olina Chau, Adriel Saporta, Aahlad Puli, Rajesh Ranganath, Andrew C. Miller. Learning Invariant Representations with Missing Data (full version).
Conference paper @ CLeaR (Causal Learning and Reasoning) 2022.

Mark Goldstein, Xintian Han, Aahlad Manas Puli, Thomas Wies, Adler J. Perotte, Rajesh Ranganath. Inverse-Weighted Survival Games.
Conference paper @ NeurIPS 2021.

Lily H. Zhang, Mark Goldstein, Rajesh Ranganath. Understanding Failures in Out-of-Distribution Detection with Deep Generative Models.
Conference paper @ International Conference on Machine Learning. 2021.

Mark Goldstein, Xintian Han, Aahlad Manas Puli, Adler J. Perotte, Rajesh Ranganath. X-CAL: Explicit Calibration for Survival Analysis.
Conference paper @ NeurIPS 2020.

Thomas Pasquier, Michael Xueyuan Han, Mark Goldstein, Thomas Moyer, David Eyers, Margo Seltzer, Jean Bacon. Practical Whole-System Provenance Capture.
Proceedings of the ACM Symposium on Cloud Computing (SoCC) 2017.

Thomas Pasquier, Michael Xueyuan Han, Mark Goldstein, Margo Seltzer, David Eyers, Jean Bacon. *Practical Provenance Capture in the Linux Operating System.*
Poster at USENIX ATC. 2017.

Workshop Papers

Mark Goldstein, Jörn-Henrik Jacobsen, Olina Chau, Adriel Saporta, Aahlad Puli, Rajesh Ranganath, Andrew C. Miller. Learning Invariant Representations with Missing Data.

DistShift Workshop @ NeurIPS 2021.

Lily H. Zhang, Mark Goldstein, Rajesh Ranganath. Understanding Out-of-Distribution Detection with Deep Generative Models.
RobustML Workshop @ ICLR 2021.

Michael Xueyuan Han, Thomas Pasquier, Tanvi Ranjan, Mark Goldstein, Margo Seltzer. FRAPpuccino: Fault-detection through Runtime Analysis of Provenance.
HotCloud Workshop @ USENIX ATC 2017.

Reviewing

Since 2020, I have been reviewing for machine learning venues such as NeurIPS, AISTATS, ICML, AAAI, ICLR, CLeaR, JMLR, and miscellaneous workshops at these conferences.

Misc.

Coding: Python (Jax, PyTorch, etc).

Languages: English (fluent), Russian (fluent), Arabic (a bit)

Instruments: oud, guitar

Music.

Studied composition with Stratis Minakakis, Anthony Coleman, Pierluigi Billone, Franck Bedrossian, Wolfgang Rihm, and Ran Blake

Compositions performed at The Stone (NYC), Summer Institute for Contemporary Performance Practice (Boston), Composit Festival (Italy), Etchings Festival (France), and International Summer Academy at the Mozarteum (Austria).

Compositions performed by members of Ensemble InterContemporain (Paris), Klangforum Wien (Vienna), and Ictus Ensemble (Brussels).

Additional Mentoring / Teaching in music

Volunteered teaching after school trumpet lessons at Mission Hill Elementary School in Jamaica Plain, MA (a neighborhood of Boston) as well as occasionally substituting for their band director. 2013.

Volunteered teaching guitar lessons on Saturdays at Boston Public Charter Prep in Hyde Park, MA (a neighborhood of Boston). 2011-2012.

REFERENCES

1. Rajesh Ranganath, NYU, rajeshr@cims.nyu.edu **2**. Finale Doshi-Velez, Harvard, finale@seas.harvard.edu.