Divya Shanmugam

divyas@cornell.edu, dmshanmugam.github.io

Academic Positions and Degrees

Cornell University Current
Postdoctoral Researcher, Department of Computer Science

Massachusetts Institute of Technology 05/2024
Ph.D, Electrical Engineering and Computer Science
Thesis: *Advancing Equity and Reliability in Machine Learning*

Massachusetts Institute of Technology 05/2018
Master of Engineering, Electrical Engineering and Computer Science
Thesis: *Representation Learning for Improved Distance and Risk Metrics*

Massachusetts Institute of Technology 05/2017
B.S., Electrical Engineering and Computer Science

Publications

* denotes equal contribution.

1. **D. Shanmugam**, M. Johnson, D. Meyers, J. Wiens, E. Pierson. "The Trillion Dollar Algorithm: Lessons from Machine Learning for Medicare Advantage Risk Adjustment". *(under review at NEJM AI)*
2. **D. Shanmugam**, B. Hardy, A. Wang, S. Divikaran, E. Pierson**, M. Barnett**. "Machine learning reveals hidden diagnoses among underserved patients". *(under review at PNAS)*
3. A. Wang, **D. Shanmugam**, S. Divikaran, E. Pierson, M. Barnett. "Identifying Mechanisms of Disparities within Cascades of Cardiovascular Care after an Emergency Department Visit". *(under review)*
4. M. Krishnamoorthy, **D. Shanmugam**, D. Tjandra, A.E. Kowalski, J. Wiens. "Survival Analysis with Limited Overlap and Censoring Distribution Shift". *(under review)*
5. **D. Shanmugam***, S. Sadhuka*, M. Raghavan, J. Guttag, B. Berger**, E. Pierson**. "Evaluating multiple models using labeled and unlabeled data". *NeurIPS 2025*
6. E. Chiang, **D. Shanmugam**, A. Beecy, G. Sayer, N. Uriel, D. Estrin, N. Garg, E. Pierson. "Learning Disease Progression Models That Capture Health Disparities". *CHIL 2025.*
7. **D. Shanmugam**, H. Lu, S. Swaminarayan, J. Guttag. "Test-time augmentation improves efficiency in conformal prediction". *CVPR 2025.*
8. **D. Shanmugam**, M. Agrawal, R. Movva, I.Y. Chen, M. Ghassemi, M. Jacobs, E. Pierson. "Generative Artificial Intelligence in Medicine". *Annual Review of Biomedical Data Science* (2025).
9. E. Pierson*, **D. Shanmugam***, R. Movva*, J. Kleinberg*, et al. "Use large language models to promote health equity" *NEJM AI 2025.*

10. **D. Shanmugam**, K. Hou, E. Pierson. "Quantifying disparities in intimate partner violence: a machine learning method to correct for underreporting". *npj Women's Health 2024.*
11. J. Gassen, S. Mengelkoch, **D. Shanmugam**, J. Pierson, A. van Lamsweerde, E. Benhar, S.E. Hill. "Longitudinal Changes in Desire and Attraction Among Women Who Have Discontinued Hormonal Contraceptives: A Preregistered Study of 1,596 Natural Cycles Users" *Hormones and Behavior 2024.*
12. R. Movva*, **D. Shanmugam***, K. Hou, P. Pathak, J. Guttag, N. Garg, E. Pierson. "Coarse race data conceals disparities in clinical risk score performance". *MLHC 2023.*
13. **D. Shanmugam**, M. Espinosa, J. Gassen, A. van Lamsweerde, J. Pearson, E. Behar, S.E. Hill. "The Relationship Between Photoperiod and Ovulation Rate: A Multi-Site Study Using NaturalCycles data". *Scientific Reports 2023.*
14. H. Suresh, **D. Shanmugam**, A. Bryan, T. Chen, A. D'Amour, J. Guttag, A. Satyanarayan. "Kaleidoscope: Semantically-grounded, Context-specific Model Evaluation". *ACM CHI 2023*
15. A. Raghu, **D. Shanmugam**, E. Pomerantsev, J. Guttag, C. Stultz. "Data Augmentation for Electrocardiograms". *ACM CHIL 2022*
16. **D. Shanmugam**, S. Shabanian, F. Diaz, M. Finck, A. Biega. "Learning to Limit Data via Scaling Laws: A Computational Interpretation for the Legal Principle of Data Minimization". *ACM FAccT 2022*
17. **D. Shanmugam**, D. Blalock, G. Balakrishnan, J. Guttag. "Better Aggregation in Test-Time Augmentation". *ICCV 2021* (oral, top 3%)
18. **D. Shanmugam**, D. Blalock, J. Guttag. "Multiple Instance Learning for ECG Risk Stratification". *MLHC 2019* (oral)

Selected talks

1. *Mind the Gap: Proxies in Public Health.* ICSDS, Vancouver June 2025
2. *Mind the Gap: Methods to Evaluate the Quality of Clinical Care.* CPH, UC Berkeley May 2025
3. *Mind the Gap: Methods to Evaluate the Quality of Clinical Care.* UVA Data Science, Virginia April 2025
4. *Measuring Human Behavior in Health Data: Hidden & Manipulated Diagnoses.* Healthy ML Group, MIT April 2025
5. *The Medicare Advantage algorithm.* AI & Public Policy Seminar, Cornell University April 2025
6. *Advancing Equity & Reliability in Machine Learning.* Statistical Methods for Health Equity, Data Science for Health Equity June 2024
7. *All the Data We Cannot See.* AI and Medicine Workshop, Mt. Sinai April 2024
8. *All the Data We Cannot See.* Computational Healthcare for Equity and iNclusion Lab, UC Berkeley November 2023
9. *Quantifying Inequality in Underreported Conditions.* MIT Algorithmic Fairness Reading Group, October 2023
10. *At the Intersection of Conceptual Art and Deep Learning: The End of Signature* List Center, October 2022.
11. *Quantifying Inequality in Underreported Conditions.* MIT AI Ethics Seminar, March 2022
12. *Quantifying Inequality in Underreported Conditions.* University of Chicago Crime and Education Lab, Virtual, December 2021

13. *Quantifying Inequality in Underreported Conditions.* Cornell Information Sciences Seminar, Virtual, November 2021
14. *Estimating the Relative Prevalence of Underreported Medical Conditions.* Microsoft Research New England, Virtual, July 2021
15. *Learning to Limit Data Collection using Scaling Laws.* Microsoft Research Montreal, Virtual, August 2020
16. *Machine Learning, Data Collection, and Women's Health.* Texas Christian University, Virtual, June 2020

Awards

Rising Stars in EECS, 2025
Best Paper Award, CHIL 2025
Best Findings Paper Award (Honorable Mention), ML4H 2023
Jane Street Fellowship (Honorable Mention), 2022
Best Talk, GW6 Research Summit 2021
NSF Graduate Research Fellowship, 2017

Professional Service

Mentorship

Helen Lu, MEng	2021-2024
William Hou, UROP	2022-2023
Anna Bryan, UROP	2021-2022
Tiffany Chen, UROP	2021-2022
Angela Zhang, UROP	2021-2022
Neha Hulkund, UROP	2020-2021
Roshni Sahoo, SuperUROP	2018-2020
Skylar Gordon, AI Mentee	2018-2019
Xinyi Guo, AI Mentee	2018-2019

Reviewing

Fairness, Accountability and Transparency Conference	2022, 2023, 2024, 2025
Conference on Health, Inference, and Learning	2020, 2022
Machine Learning for Healthcare Conference	2020, 2021, 2022, 2023
Machine Learning for Health Symposium	2019, 2020, 2022
International Conference on Machine Learning	2021, 2022
Computer Vision and Pattern Recognition Conference	2021
Women in Machine Learning NeurIPS workshop	2018, 2020
Neural and Information Processing Systems	2020

Panels

AI and Geriatrics Workshop (Johns Hopkins)	2025
AI and Society Panel (Olin College)	2023
Career Mentorship Panel (MIT Undergraduate Research Technology Conference)	2021
Graduate Student Panel (McCormick Hall)	2020
Graduate Student Panel (MIT Women in EECS)	2019
Lightning Talks (MIT Women in EECS)	2017

Service

Undergraduate Mentorship Improvement Initiative	2019-2020
GW6 Event Coordinator	2018-2019
MIT AI Mentorship Program Coordinator	2018-2020

Workshops and poster sessions

1. *At the Intersection of Conceptual Art and Deep Learning*
   Broadening Research Collaborations Workshop (NeurIPS 2022). New Orleans, LA.
2. *Improved Test Classification using Test-Time Augmentation*
   UpdatableML Workshop (ICML 2022). Virtual.
3. *Day Length Predicts Ovulatory Shifts in Women's Reproductive Function*
   Society for Personality and Social Psychology Conference (SPSP 2022). San Francisco, CA.
4. *Learning to Limit Data Collection for Data Minimization Compliance*
   Women in Machine Learning Workshop, NeuRIPS 2020
5. *Unsupervised Domain Adaptation in the Absence of Source Data*
   Uncertainty & Robustness in Deep Learning Workshop, ICML 2020
6. *Image Segmentation of Liver Stage Malaria Infection with Spatial Uncertainty Sampling*
   Workshop on Computational Biology, ICML 2019
7. *Multiple Instance Learning for Cardiac Risk Stratification.*
   Women in Machine Learning Workshop, NeurIPS 2018 (oral presentation)
8. *Multiple Instance Learning for ECG Risk Stratification.*
   Machine Learning for Health Workshop, NeurIPS 2018
9. *ECG Risk Stratification Using Multiple Instance Learning.*
   MIT DSAIL 2018
10. *Jiffy: A Convolutional Approach to Learning Time Series Similarity.*
    MIT MasterWorks 2018
11. *A Convolutional Approach to Learning Time Series Similarity.*
    Women in Machine Learning Workshop, NeurIPS 2017
12. *Identifying and Accounting for Disparities in Language Due to Economic Class.*
    Women in Machine Learning Workshop, NeurIPS 2017
13. *Compressive Metagenomics*
    MIT Microbiome Center Symposium 2016