

Harini Suresh

hsuresh@mit.edu
harinisuresh.com

Education

Massachusetts Institute of Technology (MIT) – Cambridge, MA

Bachelor of Science in Computer Science, 2016
Master of Engineering in Computer Science, 2017
PhD in Computer Science, current
GPA: 4.8/5.0

Research

Aug 2017 - Present | MIT, Cambridge, MA

Pursuing a PhD as part of the Clinical and Applied Machine Learning Group, Visualization Group, and Data + Feminism Lab. Areas of focus: human-computer interaction, machine learning, ethical computing.

Aug 2016 - June 2017 | MIT, Cambridge, MA

Completed a Masters of Engineering thesis as part of the MIT Clinical Decision Making Group. Integrated multimodal clinical data and used deep learning models to improve the state-of-the-art for intervention prediction in intensive care units and provide interpretability for model decisions (presented at Machine Learning for Healthcare, 2017). Advised by Peter Szolovits.

Experience

Co-Director, ML Tidbits Inc.

Fall 2018 - Present | Cambridge, MA

- Co-director of ML Tidbits, an educational non-profit intended to empower the public to understand and discuss machine learning concepts and their societal effects
- Writing, filming, illustrating and publishing engaging short videos on YouTube

Research Intern at Google, PAIR Team

Fall 2020 | Remote

- Developed methods for visualizing long narrative text with sentence embeddings
- Produced a backend pipeline for applying methods to new text and an open-sourced interactive visualization

Research SWE Intern at Google, Brain Fairness Team

Summer 2018 | Cambridge, MA

- Quantified “dataset bias” in large text datasets
- Identified features of a dataset that led to bias in resulting word embeddings and downstream tasks
- Implemented a multipart production system to train several embedding models with different datasets

Teaching and Leadership

AI Ethics Reading Group, Lead Organizer

mitaiethics.github.io | October 2018 - May 2020

Co-founded and led an MIT-wide initiative to connect people interested in topics around AI Ethics and facilitate interdisciplinary conversations. Organized bi-weekly meetings with curated readings and discussions.

Intro to Deep Learning, Lead Organizer

introtodeeplearning.com/2017 | January 2017

Led an extensive introduction to the field of deep learning, covering applications to machine translation, image recognition, game playing, image generation and more. Counted as a for-credit course at MIT. Included hands-on labs in TensorFlow and peer brainstorming sessions. 250+ students attended.

Publications

WORKING PAPERS

Harini Suresh, Kathleen M. Lewis, John V. Guttag, Arvind Satyanarayan. [Intuitively Assessing ML Model Reliability through Example-Based Explanations and Editing Model Inputs.](#)

CONFERENCE & JOURNAL PUBLICATIONS

Harini Suresh, John Guttag. [A Framework For Understanding Sources of “Bias” in Machine Learning.](#) *ACM conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (EAAMO '21).

Harini Suresh, Steven R. Gomez, Kevin K. Nam, Arvind Satyanarayan. [Beyond Expertise and Roles: A Framework to Characterize the Stakeholders of Interpretable Machine Learning and their Needs.](#) *CHI Conference on Human Factors in Computing Systems* (CHI '21).

Susanne Gaube*, **Harini Suresh***, Martina Raue, Alexander Merritt, Seth J. Berkowitz, Eva Lerner, Joseph F. Coughlin, John V. Guttag, Errol Colak, Marzyeh Ghassemi. [Do as AI say: susceptibility in deployment of clinical decision-aids.](#) *npj Digital Medicine*, 2021. (* = equal contribution)

Harini Suresh, Natalie Lao, Ilaria Liccardi. [Measuring the Interference of Machine Learning in Human Decision-Making.](#) *ACM Conference on Web Science* (WebSci) 2020. ACM Conference Proceedings.

Harini Suresh^{*}, Jen Gong^{*}, John Guttag. [Learning Tasks for Multitask Learning: Heterogenous Patient Populations in the ICU.](#) *Conference on Knowledge Discovery and Data Mining* (KDD) 2018, London, UK. ACM Conference Proceedings. (* = equal contribution)

Harini Suresh, Nathan Hunt, Alistair Johnson, Leo Anthony Celi, Peter Szolovits, Marzyeh Ghassemi. [Clinical Event Prediction and Understanding using Neural Networks.](#) *Machine Learning for Healthcare Conference* (MLHC) 2017, Boston MA. JMLR Workshop and Conference Track.

Willie Boag, **Harini Suresh**, Leo Celi, Peter Szolovits and Marzyeh Ghassemi. [Racial Disparities and Mistrust in End-of-Life Care.](#) *Machine Learning for Healthcare Conference* (MLHC) 2018, Stanford CA. JMLR Workshop and Conference Track.

WORKSHOPS AND POSTERS

Catherine D'Ignazio, Helena Suarez Val, Silvana Fumega, **Harini Suresh**, Isadora Cruxen, Wonyoung So, Maria De Los Angeles Martinez and Mariel Garcia-Montes. [Femicide & Machine Learning: Detecting Gender-based Violence to Strengthen Civil Sector Activism.](#) *Mechanism Design for Social Good Workshop* (MD4SG) 2020. **(New Horizons Award for Most Inspiring Paper)**

Willie Boag, **Harini Suresh**, Leo Celi, Peter Szolovits and Marzyeh Ghassemi. [Modelling Mistrust in End-of-Life Care.](#) *Fairness, Accountability, and Transparency in Machine Learning Workshop*, ICML 2018, Stockholm.

Harini Suresh*, Divya Shanmugam*, John Guttag. Disparities in the Performance of Natural Language Processing Tools. *Women in Machine Learning (WiML) Workshop*, NIPS 2017, Palm Springs CA. (* = equal contribution)

Harini Suresh, Peter Szolovits, Marzyeh Ghassemi. The Use of Autoencoders for Discovering Patient Phenotypes. *Machine Learning for Healthcare Workshop*, NIPS 2016, Barcelona.

Invited Talks and Panels

UCL-Toronto Ethical Innovation for AI workshop (July 2020). *Understanding and Preventing Unintended Consequences of ML* (talk and panel).

ACM Conference on Web Science (July 2020). *Measuring the Interference of Machine Learning in Human Decision-Making* (talk).

MIT Better World symposium in Atlanta, GA (October 2019). *Trust Issues in Machine Learning* (talk and panel).

Fair ML in Health at Data & Society Research Institute in New York, NY (October 2019). *Deploying decision-aids: real-world considerations* (talk).

Computational Cultures: Uncommon Knowledge at MIT Department of Philosophy (May 2019). *Ethics across disciplines* (lightning talk and panel).

Diversity and Inclusion Symposium by True Blue Inclusion in New York, NY (May 2019). *Tackling Harm and Improving Accountability in the Automation of Talent Management* (talk and panel).

Systems that Learn @ CSAIL Annual Meeting (August 2018). *Bias in Machine Learning and Applications to Healthcare* (talk).

Machine Learning for Healthcare Conference in Boston, MA (August 2017). *Clinical Event Prediction and Understanding using Neural Networks* (lightning talk).