

Cynthia Rudin

Associate Professor of Computer Science, Electrical and Computer Engineering, and Statistical Science

Duke University

Associate Director, SAMSI

Berkman Klein Center for Internet & Society at Harvard University, Faculty Affiliate

D224 LSRC Building, Box 90129, Durham, NC 27708

cynthia@cs.duke.edu <https://users.cs.duke.edu/cynthia/>

Research Interests

My work focuses on interpretable machine learning. This includes the design of algorithms, interpretable policy design, variable importance measures, causal inference methods, new forms of decision theory, ranking methods that assist with prioritization, uncertainty quantification, and methods that can incorporate domain-based constraints and other types of domain knowledge into machine learning. These techniques are applied to critical societal problems in criminology, health-care, energy grid reliability, and other areas. The interpretable machine learning algorithms heavily rely on efficient discrete optimization techniques.

Some accomplishments:

- My collaborators and I developed the first practical code for decision lists/trees that provably optimize accuracy and sparsity, called Certifiably Optimal Rule Lists (CORELS).
- I led a team in the first major effort to maintain an underground electrical distribution network (NYC) using machine learning
- I solved a well-known unsolved theoretical problem in machine learning as a PhD student, which is whether AdaBoost maximizes the margin like SVM.
- I have given invited/keynote talks at several of the top machine learning conferences (KDD, AISTATS, ECML-PKDD, etc).

Education

- **Ph.D.: Princeton University.** Program in Applied and Computational Mathematics. Title: **Boosting, Margins, and Dynamics**. Advisors: Ingrid Daubechies and Robert Schapire. Committee Member: Sanjeev Kulkarni; Outside Reader: Tong Zhang (Rutgers). 2004
- **BS/BA: University at Buffalo (SUNY), Honors Program.** Outstanding Senior Award in the Arts and Sciences (one awarded per year university-wide), separate outstanding senior awards from the Physics Department, Mathematics Department, and Music Department, BS Mathematical Physics, BA Music Theory, Minor in Computer Science, Summa Cum Laude, 1999

Employment History

- **Duke University**, Computer Science Department (50%), Electrical and Computer Engineering Department (50%), Secondary appointment in Statistics, Secondary appointment in Mathematics, Associate Professor, 2016-present
- **Massachusetts Institute of Technology**, MIT Computer Science and Artificial Intelligence Laboratory and Sloan School of Management, Associate Professor, 2013-2016, Assistant Professor, 2009-2013
- **Columbia University**, Center for Computational Learning Systems, Associate Research Scientist, 2007-2009, Adjunct Research Scientist, 2009-present
- **NSF Postdoctoral Research Fellow, New York University**, 2004-2007
- **Princeton University**, Graduate Research Assistant, Program in Applied and Computational Mathematics, September 1999 - June 2004, Postdoctoral Research Fellow, summer 2004

- **University at Buffalo (SUNY)**, “*Constructing a 3d model of Ridge Arrays in Strained Solid Films*,” research assistant to Dr. Brian Spencer, summer 1999

Peer-Reviewed Publications

Papers associated with major awards (winner and finalist)

1. Cynthia Rudin and Berk Ustun. *Finalist for 2017 Daniel H. Wagner Prize for Excellence in Operations Research, Institute for Operations Research and Management Science (INFORMS). **Optimized Scoring Systems: Towards Trust in Machine Learning for Healthcare and Criminal Justice***. Interfaces, In Press, 2018. (Wagner Prize has three rounds of judging)
2. William Souillard-Mandar, Randall Davis, Cynthia Rudin, Rhoda Au, David J. Libon, Rodney Swenson, Catherine C. Price, Melissa Lamar, Dana L. Penney. *2016 INFORMS Innovative Applications in Analytics Award (shared with paper below). **Learning Classification Models of Cognitive Conditions from Subtle Behaviors in the Digital Clock Drawing Test***. Machine Learning, volume 102, number 3, 2016. (Three rounds of judging)
3. Berk Ustun and Cynthia Rudin. *2016 INFORMS Innovative Applications in Analytics Award and also Runner up, Invenia Labs SEE Award 2018 - Supporting Machine Learning Research with a Positive Impact on Social, Economic, or Environmental (SEE) Challenges **Supersparse Linear Integer Models for Optimized Medical Scoring Systems***. Machine Learning, volume 102, number 3, 2016.
4. Cynthia Rudin, Şeyda Ertekin, Rebecca Passonneau, Axinia Radeva, Ashish Tomar, Boyi Xie, Stanley Lewis, Mark Riddle, Debbie Pangsrivini, Tyler McCormick. *INFORMS 2013 Innovative Applications in Analytics Award **Analytics for Power Grid Distribution Reliability in New York City***. Interfaces, volume 44, issue 4, pages 364-383, 2014. (Three rounds of judging)
5. Indraneel Mukherjee, Cynthia Rudin, and Robert E. Schapire. *This paper answered an open question published in COLT 2010. **The Rate of Convergence of AdaBoost***, Journal of Machine Learning Research, volume 14, pages 2315-2347, August 2013.
Preliminary work:
 - Indraneel Mukherjee, Cynthia Rudin, and Robert E. Schapire. **The Rate of Convergence of AdaBoost**, Proceedings of the 24th Annual Conference on Learning Theory (COLT), 2011.

Other award-winning papers

6. Yijie Bei, Alex Damian, Shijia Hu, Sachit Menon, Nikhil Ravi, and Cynthia Rudin. *NTIRE-CVPR 2018 Image Super-Resolution Challenge: winner for Track 1 (classic bicubic), honorable mention for Track 2 (realistic mild adverse conditions). **New Techniques for Preserving Global Structure and Denoising with Low Information Loss in Single-Image Super-Resolution***, New Trends in Image Restoration and Enhancement Workshop and Challenges on Super-Resolution, Dehazing, and Spectral Reconstruction, NTIRE-CVPR, 2018.
7. Cynthia Rudin and Yining Wang. *Finalist for 2017 QSR (Quality, Reliability and Statistics) best refereed paper competition, INFORMS 2017. **On Direct Learning to Rank and Rerank***. Artificial Intelligence and Statistics (AISTATS), 2018.
8. Fulton Wang, Tyler McCormick, Cynthia Rudin, and John Gore. *Best Poster Award, Conference of the ASA Section on Statistical Learning and Data Mining, 2014. **Modeling Recovery Curves With Application to Prostatectomy***. Biostatistics, 2018.
9. Berk Ustun and Cynthia Rudin. *Winner of the 2017 INFORMS Computing Society Student Paper Prize. **Learning Optimized Risk Scores from Large-Scale Datasets***. Knowledge Discovery in Databases (KDD), 2017.
10. Hongyu Yang, Cynthia Rudin, and Margo Seltzer. *Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2016. **Scalable Bayesian Rule Lists***. International Conference on Machine Learning (ICML), 2017.

11. Stefano Tráca and Cynthia Rudin. *Best paper award, INFORMS 2016 Data Mining & Decision Analytics (DMDA) Workshop*, also *Finalist for 2015 IBM Service Science Best Paper Award*. **Regulating Greed over Time**.
12. Himabindu Lakkaraju and Cynthia Rudin. *Finalist for 2017 INFORMS Data Mining Best Paper Competition*. **Learning Cost-Effective and Interpretable Treatment Regimes**. Artificial Intelligence and Statistics (AISTATS), 2017.
 - Shorter versions accepted to workshops: NIPS Machine Learning for Healthcare (ML4HC), NIPS Workshop on Interpretable Machine Learning in Complex Systems, and NIPS Workshop on ML and the Law, NIPS 2016.
13. Jiaming Zeng, Berk Ustun, and Cynthia Rudin. *Winner of 2015 Undergraduate Statistics Research Project Competition (USRESP) sponsored by the American Statistical Association (ASA) and the Consortium for Advancement of Undergraduate Statistics Education (CAUSE)*. **Interpretable Classification Models for Recidivism Prediction**. Journal of the Royal Statistical Society Series A, volume 180, issue 3, 2016.
14. Fulton Wang and Cynthia Rudin. *Winner of Best Student Paper Competition, Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2015*, also *Finalist for INFORMS Data Mining Section Best Student Paper Award, 2015*. **Falling Rule Lists**. Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS), 2015.
15. Tong Wang, Cynthia Rudin, Daniel Wagner and Rich Sevieri. *Second place in INFORMS 2015 Doing Good with Good OR Paper Competition* **Finding Patterns with a Rotten Core: Data Mining for Crime Series with Core Sets**. Big Data. volume 3, issue 1, 2015. (Special issue on Data for Social Good)
16. Benjamin Letham, Cynthia Rudin, Tyler H. McCormick, and David Madigan. *Winner of Data Mining Best Student Paper Competition, INFORMS 2013*, also *Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2014*. **Building Interpretable Classifiers with Rules using Bayesian Analysis: Building a Better Stroke Prediction Model**. Annals of Applied Statistics, volume 9, number 3, pages 1350-1371, 2015.
Shorter version:
 - Benjamin Letham, Cynthia Rudin, Tyler McCormick, and David Madigan. **An interpretable model for stroke prediction using rules and Bayesian analysis**. Proceedings of 2014 KDD Workshop on Data Science for Social Good, 2014
17. Theja Tulabandhula and Cynthia Rudin. *Finalist, Data Mining Best Student Paper Competition, INFORMS 2012*. **Machine Learning with Operational Costs**. Journal of Machine Learning Research (JMLR), volume 14, pages 1989-2028, July 2013. Preliminary work is in the following conference paper.
 - Theja Tulabandhula and Cynthia Rudin. **The Influence of Operational Costs on Estimation**, Proceedings of the International Symposium on Artificial Intelligence and Mathematics (ISAIM), 2012.

Peer-Reviewed Publications (not including those above)

2018

18. Cynthia Rudin and Şeyda Ertekin. **Learning Customized and Optimized Lists of Rules with Mathematical Programming**. Mathematical Programming C (Computation), In press, 2018.
19. Elaine Angelino, Nicholas Larus-Stone, Daniel Alabi, Margo Seltzer, and Cynthia Rudin. **Learning Certifiably Optimal Rule Lists for Categorical Data**, Journal of Machine Learning Research, volume 18, no. 234, pages 1-78, 2018.
20. Nicholas Larus-Stone, Elaine Angelino, Daniel Alabi, Margo Seltzer, and Cynthia Rudin. **Systems Optimizations for Learning Certifiably Optimal Rule Lists**. SysML, 2018.
21. Gah-Yi Ban and Cynthia Rudin. **The Big Data Newsvendor: Practical Insights from Machine Learning**. Operations Research, accepted, 2018.

22. Oscar Li, Hao Liu, Chaofan Chen, and Cynthia Rudin. **Deep Learning for Case-based Reasoning through Prototypes: A Neural Network that Explains its Predictions**. Association for the Advancement of Artificial Intelligence (AAAI), 2018.
23. Chaofan Chen and Cynthia Rudin. **An Optimization Approach to Learning Falling Rule Lists**. Artificial Intelligence and Statistics (AISTATS), 2018.

2017

24. Aaron F. Struck, Berk Ustun, Andres Rodriguez Ruiz, Jong Woo Lee, Suzette LaRoche, Lawrence J. Hirsch, Emily J Gilmore, Jan Vlachy, Hiba Arif Haider, Cynthia Rudin, M Brandon Westover. **Association of an Electroencephalography-Based Risk Score With Seizure Probability in Hospitalized Patients**, JAMA Neurology, 74 (12), 1419-1424, 2017.
25. Elaine Angelino, Nicholas Larus-Stone, Daniel Alabi, Margo Seltzer, and Cynthia Rudin. **Certifiably Optimal Rule Lists for Categorical Data**, Knowledge Discovery in Databases (KDD - oral), 2017.
26. Berk Ustun, Lenard A. Adler, Cynthia Rudin, Stephen V. Faraone, Thomas J. Spencer, Patricia Berglund, Michael J. Gruber, Ronald C. Kessler. **The World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for DSM-5**. JAMA Psychiatry, April 2017.
27. Tong Wang, Cynthia Rudin, Finale Doshi, Yimin Liu, Erica Klampfl, and Perry MacNeille. **Bayesian Rule Sets for Interpretable Classification, with Application to Context-Aware Recommender Systems**. Journal of Machine Learning Research (JMLR), volume 18, number 70, pages 137, 2017.

2016

28. Tong Wang, Cynthia Rudin, Finale Doshi, Yimin Liu, Erica Klampfl, and Perry MacNeille. **Bayesian Rule Sets for Interpretable Classification**. IEEE International Conference on Data Mining (ICDM), 2016.
29. Ramin Moghaddass, Cynthia Rudin, and David Madigan. **The Factorized Self-Controlled Case Series Method: An Approach for Estimating the Effects of Many Drugs on Many Outcomes**. Journal of Machine Learning Research, 17(185):124, 2016.
30. Berk Ustun, M. Brandon Westover, Cynthia Rudin, and Matt T. Bianchi. **Clinical Prediction Models for Sleep Apnea: The Importance of Medical History over Symptoms**. Journal of Clinical Sleep Medicine, volume 12, number 2, 2016.
31. Vikas Garg, Cynthia Rudin, and Tommi Jaakola. **CRAFT: ClusteR-specific Assorted Feature selecTION**, Artificial Intelligence and Statistics (AISTATS), 2016.
32. Benjamin Letham, Portia A. Letham, Cynthia Rudin, Edward P. Browne. **Prediction Uncertainty and Optimal Experimental Design for Learning Dynamical Systems**. Chaos, volume 26, number 6, 2016.
33. Benjamin Letham, Lydia M. Letham and Cynthia Rudin. **Bayesian Inference of Arrival Rate and Substitution Behavior from Sales Transaction Data with Stockouts**. Knowledge Discovery in Databases (KDD), 2016
34. Edward P. Browne, Benjamin Letham, and Cynthia Rudin. **A Computational Model of Inhibition of HIV-1 by Interferon-Alpha**. PLoS ONE, vol 11, no 3, pages 1–16, March, 2016.

2015

35. Şeyda Ertekin and Cynthia Rudin. **A Bayesian Approach to Learning Scoring Systems**. Big Data, volume 3, number 4, 2016.
36. Ramin Moghaddass and Cynthia Rudin. **The Latent State Hazard Model, with Application to Wind Turbine Reliability**. Annals of Applied Statistics, volume 9, number 4, pages 1823–1863, 2015.

37. Şeyda Ertekin, Cynthia Rudin, and Tyler McCormick. **Reactive Point Processes: A New Approach to Predicting Power Failures in Underground Electrical Systems**. Annals of Applied Statistics, volume 9, No 1, pages 122–144, 2015.

2014

38. Theja Tulabandhula and Cynthia Rudin. **Tire Changes, Fresh Air and Yellow Flags: Challenges in Predictive Analytics for Professional Racing**. Big Data, vol 2 issue 2, pages 97-112, June 20, 2014.
39. Şeyda Ertekin, Cynthia Rudin and Haym Hirsh. **Approximating the Crowd**. Data Mining and Knowledge Discovery, volume 28 issue 5-6, pages 1189-1221, September 2014.
- Shorter versions appeared at NIPS Workshop on Computational Social Science and the Wisdom of Crowds, 2011, Proceedings of Collective Intelligence (CI) 2012, and Proceedings of the 2012 AAAI Fall Symposium on Machine Aggregation of Human Judgment, MAGG-2012.
40. Been Kim, Cynthia Rudin, and Julie Shah. **The Bayesian Case Model: A Generative Approach for Case-Based Reasoning and Prototype Classification**. Neural Information Processing Systems (NIPS), 2014.
41. Theja Tulabandhula, Cynthia Rudin. **On Combining Machine Learning with Decision Making**. Machine Learning (ECML-PKDD journal track), volume 93, Pages 33-64, 2014
42. Siong Thye Goh and Cynthia Rudin. **Box Drawings for Learning with Imbalanced Data**. Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2014.
43. Theja Tulabandhula and Cynthia Rudin. **Generalization Bounds for Learning with Linear, Polygonal, Quadratic and Conic Side Knowledge**. Machine Learning (ECML-PKDD journal track), December, 2014, pages 1-34.
- Shorter version: Theja Tulabandhula, Cynthia Rudin. **Generalization Bounds for Learning with Linear and Quadratic Side Knowledge**. Proceedings of ISAIM 2014.
44. Jonathan Huggins and Cynthia Rudin. **Towards a Theory of Pattern Discovery** Proceedings of SIAM Conference on Data Mining (SDM) 2014.
45. Been Kim and Cynthia Rudin. **Learning About Meetings**, Data Mining and Knowledge Discovery, (ECML-PKDD Journal track), volume 28 issue 5-6, pages 1134-1157, September 2014.

2013

46. Benjamin Letham, Cynthia Rudin and Katherine Heller. **Growing a List**. Data Mining and Knowledge Discovery (DAMI), ECML-PKDD journal track. volume 27, pages 372-395, 2013.
47. Tong Wang, Cynthia Rudin, Daniel Wagner, Richard Sevieri. **Learning to Detect Patterns of Crime**, Proceedings of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2013.
48. Cynthia Rudin, Benjamin Letham, and David Madigan. **Learning Theory Analysis for Association Rules and Sequential Event Prediction**. Journal of Machine Learning Research (JMLR), volume 14, pages 3385-3436, 2013.
- Shorter version: Cynthia Rudin, Ben Letham, Ansaf Salieb-Aouissi, Eugene Kogan, and David Madigan. **Sequential Event Prediction with Association Rules**, Proceedings of the 24th Annual Conference on Learning Theory (COLT), 2011.
49. Benjamin Letham, Cynthia Rudin and David Madigan. **Sequential Event Prediction**. Machine Learning, volume 93, pages 357-380, 2013

2012

50. Tyler McCormick, Cynthia Rudin, and David Madigan. **Hierarchical Models for Association Rule Mining: A New Approach for Adverse Event Prediction in Clinical Trials**, *Annals of Applied Statistics*, volume 6, No. 2, pages 652–668, 2012.
51. Cynthia Rudin, David Waltz, Roger N. Anderson, Albert Boulanger, Ansaf Salieb-Aouissi, Maggie Chow, Haimonti Dutta, Philip Gross, Bert Huang, Steve Ierome, Delfine Isaac, Arthur Kressner, Rebecca J. Passonneau, Axinia Radeva, Leon Wu. **Machine Learning for the New York City Power Grid**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 34, No 2, February 2012. *Spotlight Paper for the February 2012 Issue*.
52. Allison Chang, Cynthia Rudin, Mike Cavaretta, Robert Thomas and Gloria Chou. **Reverse-Engineering Quality Ratings**, *Machine Learning: volume 88, issue 3*, pages 369-398, 2012.

2011

53. Cynthia Rudin, Rebecca J. Passonneau, Axinia Radeva, Steve Ierome, and Delfina Isaac. **21st-Century Data Miners Meet 19th-Century Electrical Cables**, *IEEE Computer*, volume 44 no. 6, pages 103-105, June 2011. *(One of three articles featured on the cover of the magazine.)*
54. Şeyda Ertekin and Cynthia Rudin. **On Equivalence Relationships Between Classification and Ranking Algorithms**, *Journal of Machine Learning Research*, volume 12, pages 2905–2929, 2011.

2010

55. Cynthia Rudin, Rebecca J. Passonneau, Axinia Radeva, Haimonti Dutta, Steve Ierome, and Delfina Isaac. **A Process for Predicting Manhole Events in Manhattan**. *Machine Learning*, volume 80, pages 1–31, 2010.

- Also oral presentation at ICML 2012

The following conference papers are also related to my projects on grid reliability.

- Rebecca J. Passonneau, Cynthia Rudin, Axinia Radeva, Ashish Tomar, Boyi Xie. **Treatment Effect of Repairs to an Electrical Grid: Leveraging a Machine Learned Model of Structure Vulnerability**, *Proceedings of the KDD Applications in Sustainability (SustKDD) Workshop on Data Mining, 17th Annual ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2011.
- Dingquan Wang, Rebecca Passonneau, Michael Collins and Cynthia Rudin. **Modeling Weather Impact on a Secondary Electrical Grid**, *4th International Conference on Sustainable Energy Information Technology (SEIT-2014)*, 2014.
- Leon Wu, Timothy Teräväinen, Gail Kaiser, Roger Anderson, Albert Boulanger, and Cynthia Rudin. **Estimation of System Reliability Using a Semiparametric Model**, *Proceedings of IEEE EnergyTech*, 2011.
- Leon Wu, Gail Kaiser, Cynthia Rudin, and Roger Anderson. **Data Quality Assurance and Performance Measurement of Data Mining for Preventive Maintenance of Power Grid**, *Proceedings of the KDD Workshop on Data Mining for Service and Maintenance (KDD4Service), 17th Annual ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2011.
- Leon Wu, Gail Kaiser, Cynthia Rudin, David Waltz, Roger Anderson, Albert Boulanger, Ansaf Salieb-Aouissi, Haimonti Dutta, and Manoj Poolery. **Evaluating Machine Learning for Improving Power Grid Reliability**, *Proceedings of the ICML 2011 workshop on Machine Learning for Global Challenges, International Conference on Machine Learning*, 2011.
- Axinia Radeva, Cynthia Rudin, Rebecca Passonneau and Delfina Isaac. **Report Cards for Manholes**, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA), 2009. Best Poster Award*.
- Rebecca Passonneau, Cynthia Rudin, Axinia Radeva and Zhi An Liu. **Reducing Noise in Labels and Features for a Real World Dataset: Application of NLP Corpus Annotation Methods**, *Proceedings of the 10th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing), 2009*.

- Haimonti Dutta, Cynthia Rudin, Becky Passonneau, Fred Seibel, Nandini Bhardwaj, Axinia Radeva, Zhi An Liu, Steve Ierome, Delfina Isaac. **Visualization of Manhole and Precursor-Type Events for the Manhattan Electrical Distribution System**, Workshop on GeoVisualization of Dynamics, Movement and Change, 11th AGILE International Conference on Geographic Information Science, 2008.
- Boyi Xie, Rebecca J. Passonneau, Haimonti Dutta, Jing-Yeu Miaw, Axinia Radeva, Ashish Tomar, Cynthia Rudin. **Progressive Clustering with Learned Seeds: An Event Categorization System for Power Grid**. 24th International Conference on Software Engineering and Knowledge Engineering (SEKE 2012). Redwood City, CA. July 1-3, 2012.

2009

56. Cynthia Rudin. **The P-Norm Push: A Simple Convex Ranking Algorithm that Concentrates at the Top of the List**, Journal of Machine Learning Research, volume 10, pages 2233–2271, 2009.
- Shorter version: Cynthia Rudin. **Ranking with a P-Norm Push**. Proceedings of the Nineteenth Annual Conference on Learning Theory (COLT), pages 589 - 604, 2006.

An application of the P-Norm Push is described in this conference paper:

- Heng Ji, Cynthia Rudin, and Ralph Grishman. **Re-ranking Algorithms for Name Tagging**. In Proc. Human Language Technology conference - North American chapter of the Association for Computational Linguistics annual meeting (HLT-NAACL) Workshop on Computationally Hard Problems and Joint Inference in Speech and Language Processing, 2006.
57. Cynthia Rudin and Robert E. Schapire. **Margin-Based Ranking and an Equivalence Between AdaBoost and RankBoost**. Journal of Machine Learning Research, volume 10, pages 2193–2232, 2009.
- Preliminary version: Cynthia Rudin, Corinna Cortes, Mehryar Mohri, and Robert E. Schapire. **Margin Based Ranking Meets Boosting in the Middle**. Proceedings of the Eighteenth Annual Conference on Learning Theory (COLT), pages 63 - 78, 2005.

2008 and before

58. Cynthia Rudin, Robert E. Schapire and Ingrid Daubechies. **Analysis of Boosting Algorithms Using the Smooth Margin Function**. Annals of Statistics, volume 35, number 6, pages 2723-2768, 2007.

Preliminary material:

- Cynthia Rudin, Robert E. Schapire, and Ingrid Daubechies. (2007) **Precise Statements of Convergence for AdaBoost and arc-gv**. In Proc. AMS-IMS-SIAM Joint Summer Research Conference: Machine Learning, Statistics, and Discovery, pages 131-145, 2007.
 - Cynthia Rudin, Robert E. Schapire, and Ingrid Daubechies. **Boosting Based on a Smooth Margin**. Proceedings of the Seventeenth Annual Conference on Computational Learning Theory, (COLT), pages 502-517, 2004.
 - Cynthia Rudin, Ingrid Daubechies, and Robert E. Schapire. **On the Dynamics of Boosting**. Advances in Neural Information Processing Systems (NIPS) 16, 2003.
59. Cynthia Rudin, Ingrid Daubechies, and Robert E. Schapire. **The Dynamics of AdaBoost: Cyclic Behavior and Convergence of Margins**. Journal of Machine Learning Research, 5 (Dec): 1557–1595, 2004.

Preliminary material for this work appears partly within the NIPS paper below, and the open problem in COLT is from the JMLR paper:

- Cynthia Rudin, Ingrid Daubechies, and Robert E. Schapire. **On the Dynamics of Boosting**. Advances in Neural Information Processing Systems (NIPS) 16, 2003.

- Cynthia Rudin, Robert E. Schapire and Ingrid Daubechies. **Does AdaBoost Always Cycle?** JMLR: Workshop and Conference Proceedings, Published as a COLT Open problem, 2012.
- 60. Ryan Roth, Owen Rambow, Nizar Habash, Mona Diab, and Cynthia Rudin. **Arabic Morphological Tagging, Diacritization, and Lemmatization Using Lexeme Models and Feature Ranking**, The 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT), 2008.
- 61. Cynthia Rudin and Brian Spencer. **Equilibrium Ridge Arrays in Strained Solid Films**. Journal of Applied Physics, vol 86, pp 5530-5536, 1999.

Non-Peer-Reviewed Publications

- 62. Cynthia Rudin (with credit to Robin Smith). **Algorithms and Justice: Scrapping the Black Box**. The Crime Report (Blog), 2018.
- 63. Cynthia Rudin. **Can Machine Learning be Useful for Social Science?**. In: The Cities: An essay collection from the Decent City initiative. <http://citiespapers.ssrc.org/essays/>, 2015.
- 64. Cynthia Rudin, David Dunson, Rafael Irizarry, Hongkai Ji, Eric Laber, Jeffrey Leek, Tyler McCormick, Sherri Rose, Chad Schafer, Mark van der Laan, Larry Wasserman, Lingzhou Xue. **Discovery with Data: Leveraging Statistics with Computer Science to Transform Science and Society**. American Statistical Association whitepaper, <http://www.amstat.org/policy/pdfs/BigDataStatisticsJune2014.pdf>, 2014.
- 65. Cynthia Rudin and Kiri Wagstaff. **Machine Learning for Science and Society**, Machine Learning, (Introduction to the Special Issue on Machine Learning for Science and Society), volume 95, issue 1, April 2014, pp 1-9.
- 66. Cynthia Rudin. **Teaching “Prediction: Machine Learning and Statistics”**, Proceedings of the ICML Workshop on Teaching ML, 2012.
- 67. Peter Qian, Yilu Zhou, and Cynthia Rudin, **Proceedings of the 6th INFORMS Workshop on Data Mining and Health Informatics (DM-HI)**, editors, 2011.
- 68. Cynthia Rudin and Miroslav Dudík, **Lecture Notes for the AMS Short Course on Statistical Learning**, editors, includes contributions by Robert E Schapire, Lawrence Saul, Lisa Hellerstein, Adam Tauman-Kalai, and John Lafferty, 2007.

Submitted / In Preparation (Drafts are posted on the Internet for most papers listed below)

- 69. Awa Dieng, Yameng Liu, Sudeepa Roy, Cynthia Rudin, and Alexander Volfovsky. **Collapsing-FLAME: A Fast Almost-Exact Matching Method for Causal Inference**. Working paper on arXiv, 2018.
- 70. Ramin Moghaddass and Cynthia Rudin. **Bayesian Patchworks: An Approach to Case-Based Reasoning**. Working paper on ArXiv, 2018.
- 71. Beau Coker, Cynthia Rudin and Gary King. **A Theory of Statistical Inference for Ensuring the Robustness of Scientific Results**. Working paper on arXiv, 2018.
- 72. Siong Thye Goh and Cynthia Rudin. **A Minimax Surrogate Loss Approach to Conditional Difference Estimation**. Working paper on arXiv, 2018.
- 73. Aaron Fisher, Cynthia Rudin, Francesca Dominici. **Model Class Reliance: Variable Importance Measures for any Machine Learning Model Class, from the “Rashomon” Perspective**. Working paper on ArXiv, 2018.
- 74. Fulton Wang and Cynthia Rudin. **Extreme Dimension Reduction for Handling Covariate Shift**. Working paper on ArXiv, 2018.
- 75. Cynthia Rudin, Prashan Wanigasekara, Jacob Bien. **Elastic Prototype Classification**. Working paper, 2015.
- 76. Fulton Wang and Cynthia Rudin. **Causal Falling Rule Lists**, Working paper on ArXiv, 2015.

- Shorter version presented at Fairness, Accountability, and Transparency (FATML), 2017.
77. Siong Thye Goh and Cynthia Rudin. **Cascaded Histograms: An Approach to Interpretable Density Estimation for Categorical Data**, Working paper on ArXiv, 2015.
 78. Md. Noor-E-Alam and Cynthia Rudin. **Robust Testing for Causal Inference in Natural Experiments**. Working Paper on Optimization Online, 2015.
 79. Md. Noor-E-Alam and Cynthia Rudin. **Robust Nonparametric Testing for Causal Inference in Observational Studies**. Working paper, 2015.
 80. Theja Tulabandhula and Cynthia Rudin. **Robust Optimization using Machine Learning for Uncertainty Sets**. ArXiv: <http://arxiv.org/abs/1407.1097>
 81. Sudeepa Roy, Cynthia Rudin, Alex Volfovsky, and Tianyu Wang. **FLAME: A Fast Large-scale Almost Matching Exactly Approach to Causal Inference**. In progress, 2017.

Note about peer reviewed conference papers: Some machine learning conferences are highly selective. Reviewing is performed by 3-5 referees, an area editor, and the program chairs. Most papers are 6-15 pages, sometimes with a supplement that contains a longer (full length) version of the paper.

Grants

An Integrated Nonparametric Bayesian and Deep Neural Network Framework for Biologically-Inspired Lifelong Learning (co-PI, with PI Katherine Heller, and other co-PI's Lawrence Carin, David Dunson, Nicolas Brunel, Tamara Broderick, Joshua Tenenbaum, Michael Jordan, Thomas Griffiths), DARPA, 2017-2021. Amount of Award: \$4,388,119 total.

Collaborative Research Framework: Data: HDR: Nanocomposites to Metamaterials: A Knowledge Graph Framework (co-PI, with PI Cate Brinson, and other co-PI's Chiara Daraio, Linda Schadler, Deborah McGuinness, and Wei Chen), NSF, 2018-2023. Amount of Award: \$2,252,971 total.

Lord Foundation "Duke's Super Superresolution Team" (Single PI) 2018-2019. Amount of Award: \$32,000

Duke Energy Initiative "Enabling Better Energy Decisions Through Better Interpretable Causal Inference Methods for Personalized Treatment Effects" (PI, with co-PI's Alex Volfovsky and Sudeepa Roy), 2018-2019. Amount of Award: \$40,500

QuBBB: Collaborative Research: Matching Methods for Causal Inference: Big Data and Networks, 1R01EB025021-01, sponsored by DHHS, PHS, NIH, NIBI&B" (co-Investigator, with PI Alexander Volfovsky, co-Investigators Sudeepa Roy and Allison Aiello). 2017-2020. Amount of Award: \$849K (\$514K to Duke)

Alfred P. Sloan Foundation "Interdisciplinary Energy Data Analytics Ph.D. Fellows Program," (co-PI, with Brian Murray, William Pizer, and Kyle Bradbury). 2018-2020. Amount of Award: \$225,000

Laura and John Arnold Foundation "Interpretable Machine Learning for Pre-Trial Risk Analysis," 2017-2018. (Single PI) Amount of Award: \$87,507

Duke Institute for Health Innovation "Palliatyics: Using analytics to inform a palliative care population health management intervention" (co-PI with Jonathan Fischer, Leslie Alabi, Eugenie Komives from Duke Medical). 2017-2018. Amount of award: \$55,500

2016 Adobe Digital Marketing Research Award, "Compute-intensive causal machine learning models for finding customer segments" (Single PI). 2016. Amount of Award: \$50K

MIT-Lincoln Labs "Adaptable, Interpretable, Machine Learning" (co-PI with Jonathan Su from MIT-LL), 2016-2019. Amount of Subaward: \$542K

Xerox Research “Causal Inference-related Research Threads in the Prediction Analysis Lab, continued” (co-PI with Theja Tulabandhula), 2016. Amount of Award: \$30K

DARPA “Foundations of Sequential Learning” (co-PI with Ron Parr and Kamesh Munagala), April 2016 - Jan 2017. Amount of Award: \$242K

Philips “Algorithms for Interpretable Risk-Scoring” (Single PI), July 2017-December 2017. Amount of Award: \$122K

Xerox Research “Causal Inference-related Research Threads in the Prediction Analysis Lab” (co-PI with Theja Tulabandhula), 2015. Amount of Award: \$30K

Philips “Self-Learning Systems and User-Behavior Modeling” (Single PI), June 2015-May 2016. Amount of Award: \$240K

Army Research Office “Uncertainty Quantification for Unobserved Variables in Dynamical Systems and Optimal Experimental Design” (Single PI), Spring 2015. Amount of Award: \$50K

Big Data Seed Grant, MIT Big Data Initiative “Interpretable, Scalable and Causal Models from Machine Learning” (Single PI), Spring 2015. Amount of award: 1 semester RA

Accenture and MIT Alliance in Business Analytics “Creating a Model of the Usual State of a Machine” (Single PI), March 20 2015 - March 20 2016. Amount of award: \$100,000

Wistron Corporation “Interpretable Predictive Models from Machine Learning” (Single PI), September 4, 2014 - August 31, 2016.
Amount of award \$ 300,000

Accenture and MIT Alliance in Business Analytics “Big Data Analysis for Plant and Commercial Optimization,” June 15 2013 - June 14 2014. Amount of award: \$ 100,000

Ford Racing “Predictive Analytics for Racing” (Single PI),
September 1 2012 - August 31 2013. Amount of award: \$75,000
September 1 2013 - December 31 2014. Amount of award: \$75,000
December 31 2014 - December 31 2015. Amount of award: \$157,000

Siemens Energy, CKI University Innovation Initiative. “CKI Proposal: Augmented Data-Driven Diagnosis using Physical Models” (Single MIT PI with co-PIs at Siemens), June 1, 2013 - May 31, 2016
Amount of award: \$350,000

Siemens Energy, CKI University Innovation Initiative. “CKI Proposal: Incorporating Prediction Analysis into XHQ” (Single MIT PI with co-PIs at Siemens), June 1 2013 - May 31 2015.
Amount of award: 150,000 euro \approx \$200,000

Ford - MIT Alliance. “Develop advanced in-vehicle SYNC advertisement features to target drivers based on their context information, system interaction and past choices” (Single PI), July 1 2012 - August 31 2014.
Amount of award: \$298,000

MIT Sloan Research Fund. “Predictive Models for Highly Imbalanced data,” (Single PI), Amount of award: \$20,000

Solomon Buchsbaum Research Fund. “A New Foundation for Statistical Decision-Making,” June 2 2011-present. (Single PI), Amount of award: \$50,000

NSF-CAREER IIS-1053407. “New Approaches for Ranking in Machine Learning,” September 1 2011 – August 31 2018.
Amount of award: \$480,000

MIT Lincoln Labs. “A Knowledge Discovery Framework for Threat Identification (Phase III)” (co-PI with Shirish Ranjit

from MIT-LL) 7/1/2013-6/30/2014.

Amount of award: \$100,000

MIT Lincoln Labs. “A Knowledge Discovery Framework for Threat Identification (Phase II)” (co-PI with Shirish Ranjit from MIT-LL) 7/1/2012-6/30/2013.

Amount of award: \$100,000

MIT Lincoln Labs. “A Knowledge Discovery Framework for Threat Identification” (co-PI with Shirish Ranjit from MIT-LL and Regina Barzilay from MIT), 6/1/2011- 5/31/2012.

Amount of award: \$83,000

MIT Intelligence Initiative. “Combining human and machine predictions using “boosting” algorithms” (co-PI with Thomas Malone and Patrick Winston) November 1 2010 - October 1 2011.

Amount of award: 0.5 postdoctoral research fellowship

MIT Energy Initiative (MITEI) Seed Fund Program. “A Novel Framework for Electrical Grid Maintenance” (Single PI) May 17 2010 - May 16 2012.

Amount of award: \$150,000

Ford - MIT Alliance. “Achieving Top Quality Ratings with Minimal Cost” (Single PI) July 1 2010 - June 20 2011.

Amount of award: \$106,013

Con Edison Company of New York. “Manhole Events and Secondary System - Machine Learning Project”

Secondary System Project, Manhattan Backbone, December 1, 2011-December 31, 2013, Amount of subaward to MIT (Single PI): \$300,129

Secondary System Project, Manhattan Corollary, May 1, 2011 - May 31, 2012, Amount of subaward to MIT (Single PI): \$93,056

Update Manhattan Consolidated Database and Ranking Model through 2009, and Analyze 2004-2005 Secondary Inspections Data, July 1 2010 - December 31 2010, Amount of subaward to MIT (Single PI): \$81,661

Phase 2, Application to B, Q, and X, December 1 2009 - May 31 2010, Amount of subaward to MIT (Single PI): \$81,899

Phase 1, Application to B, Q, and X, January 2009 - June 2009, Co-PI , Amount: \$464,127

Phase 4, July, 2008 - December, 2008, Co-PI , Amount: \$413,947

Phase 3, January, 2008 - June, 2008, Co-PI , Amount: \$347,104

Phase 2, August, 2007 - December, 2007, Co-PI, Amount: \$486,296

Phase 1, March, 2007 - July, 2007, Co-PI , Amount: \$339,251

National Science Foundation. Postdoctoral Research Fellowship in Biological Informatics, Grant DBI-0434636, March 2005 - February 2007.

Amount of award: \$120,000

Honors (not including awards listed with papers above)

Faculty associate, Berkman Klein Center for Internet and Society at Harvard University, 2015-2019.

2016 Adobe Digital Marketing Research Award, 2016

Named by Businessinsider.com as one of the 12 most impressive professors at MIT in 2015

National Science Foundation CAREER Award, 2011

“Top 40 Under 40” business school professors of *Poets & Quants*, 2015. (Published in Forbes magazine.)

Nominated for Outstanding UROP Mentor Award, UROP (undergraduate research opportunities) Program, MIT, 2012

Nominated for 2012 Sloan Excellence in Teaching Awards, MIT, 2012

Second Place for Phase 1 in the ICML Exploration and Exploitation 3 Challenge, 2012. Goal is to design a recommender system with high click through rates for Yahoo! Front Page News Article Recommendations teammates: Virota Chiraphadhanakul and Edward Su

National Science Foundation Postdoctoral Research Fellowship in Biological Informatics, 2005-2007

University at Buffalo College of Arts and Sciences Outstanding Senior Award in Sciences and Mathematics 1999, one per year at the university (also Department of Physics Outstanding Senior 1998, Department of Mathematics Outstanding Senior 1999, Department of Music Outstanding Senior 1999)

National Science Foundation Graduate Fellowship Honorable Mention, 1999

Barry J. Goldwater Scholarship, 1997-1998

State University of New York Chancellors Award for Student Excellence 1997, 1999

Dr. Stanley T. Sekula Memorial Scholarship, University at Buffalo Physics Department, 1996, 1997

Hildegard F. Shinnars Prize, 1999, Phi Beta Kappa award to recognize mathematics thesis and music thesis

Phi Beta Kappa, inducted 1997

UB Music Department Scholarships, 1994, 1995, 1996

Intellectual Breadth and Liberal Knowledge Award, UB Honors Program, 1999

Media

Crime Data Mining:

“Possible to predict recidivism? Here’s how...,” *The Docket*, MSNBC (Live TV), Tuesday May 19, 2015.

Interviewed for “Mapping the Future. The Power of Algorithms,” Documentary, planned for release in 2015.

“Crime-Fighting Computer Code from Cambridge Police and MIT,” *WBUR Radio Boston* (National Public Radio), Tuesday August 13 2013

“Cambridge police look at math to solve crimes,” *Boston Globe*, Metro Section, front page, Sunday August 4 2013

“Statistician enlisted to fight crime by numbers,” *The Times of London*, US & Americas section, Tuesday August 6 2013

“Predictive Policing: Using Machine Learning to Detect Patterns of Crime,” *WIRED Innovation Insights*, August 22, 2013

Meetings Analysis:

“How to sound like a boss at your next work meeting,” Interview on Fox & Friends, *Fox News*, national television, 6:45AM July 3 2013

“At Work: Just Say ‘Yeah’,” *Wall Street Journal*, Business section, on page B8 in the U.S. edition, June 19, 2013

“How to be effective at meetings? Say ‘yeah’,” *Toronto Star*, Business section, June 28, 2013

“So I was thinking...” *Entrepreneur Magazine*, pages 30-31, June 2014.

“‘Yeah’ is the most persuasive word to use in a meeting: study says,” *NY Daily News*, living section, June 26, 2013

“Most persuasive four-letter-word in meetings (it’s not what you think),” *Yahoo! Shine*, work-money section, June 25, 2013

“5 Words That Get Results Business Meetings,” *Yahoo! News*, (also *LiveScience* and *Business News Daily*), June 21, 2013

“Researchers discover the key to persuasion,” *ABC News*, consumer report blog / business, June 24, 2013

“MIT Sloan Professor’s Research Identifies Top Persuasive Words in Meetings,” *CNBC.com*, June 20, 2013

“5 Important words to say in business meetings,” *Business News Daily*, June 20, 2013

“How To Win Over Co-Workers And Influence Meetings: Use These 3 Words,” *Forbes.com*, Pharma & healthcare section, June 23, 2013

“The most effective words to use at a meeting,” *Fox Nation*, June 20, 2013

“Eight most persuasive words you can use in a meeting,” *Fox 9 News* (TV), June 25, 2013

“Five persuasive words for more successful meetings,” *Fox 8 News* (TV), June 24, 2013

Other mentions of this work: “The most persuasive word you can use in a meeting is ‘yeah’,” *Quartz.com*, *Yahoo! Finance*, June 24, 2013, “Want to get ahead in meetings, say ‘yeah’ ... a lot” *The Telegraph* (UK), Finance section, June 26, 2013, ““Be successful at meetings with a ‘yeah’,” *Business Standard*, human-interest-society section, June 25, 2013”, “Could ‘yeah’ be the most persuasive word to use in a meeting?: The Science Behind Persuasion,” *Medical Daily*, Science/tech section, June 26, 2013, “Dialogue analysis: Most persuasive words for business meetings” WLS 980 AM - Don Wade and Roma (Audio) June 24 2013, “Want a good raise? Five words that can help you.” *New Delhi Television Limited*, June 25, 2013, “Yeah named most persuasive word,” *The Times of India*, Science section, June 26 2013, “Yeah, yeah, yeah: These words will get you ahead in business meetings, yeah?” *The Independent* (UK), Business news section, June 27, 2013

Energy Grid:

- “Unexpected Consequences.” *Civil Engineering: The Magazine of the American Society of Civil Engineers*, pages 72-77, December, 2015
- “Why Manhole Explosions Happen in the Summer,” *NBC News*, Business/Energy section, August 19, 2015
- “New York’s Exploding Manhole Covers Pose Unexpected Winter Hazard,” *Reuters*, appeared in *New York Times*, and *MSN.com*, February 28, 2015
- Interviewed for Chasing NJ, March 10, 2015.
- Interviewed for History Channel show on manhole accidents, March 11, 2015
- My work discussed in book Big Data: A Revolution that Will Transform how we Live, Work, and Think, by Victor Mayer-Schönberger and Kenneth Cukier, Houghton Mifflin Harcourt Publishing Company, 2013
- Analytics Magazine*, INFORMS. Headlines: Innovative Applications in Analytics Award, April 18, 2013
- “Don’t blow your top,” *CIO magazine “finish” section*, Sept 1 issue, 2010
- “Ranking Of Manholes Prioritizes Repairs And Maintenance,” *Energy Daily*, Nov 4 2009
- Interview with Indianapolis star “IPL: City will be safe for the Super Bowl: But fixing manhole problems goes beyond big game,” November 22, 2011
- Interview with Band News FM, Journalismo Radio, São Paulo, Brazil, July 8, 2011
- “Machine Versus Manhole,” *ScienceNews*, *U.S. News and World Report*, *WIRED Science*, *Slashdot*, *Discovery News / Discovery Channel*, and others, July 8-9 2010

Information Retrieval / Building New Search Engines:

- Radio segment about my work on Growing a List. “A New Way to Google,” *Boston Public Radio*, show on innovation at 12:45pm-1pm, hosted by Kara Miller, October 9, 2012

Health and Interpretable Predictive Models:

- “Do You Zone Out? Procrastinate? Might Be Adult ADHD,” by Rebecca Hersher, NPR, April 5 2017
- “Good News for Screening for Adult Attention-Deficit/Hyperactivity Disorder,” by Philip Shaw, Kwangmi Ahn, and Judith L. Rapoport, invited commentary in *JAMA Psychiatry*, 2017
- “Algorithms Learn From Us, and We’ve Been Bad Parents,” by Bahar Gholipour, Mach Technology, NBC News, March 10 2017, 2:17 PM ET
- “Big-Data or Slim-Data: Predictive Analytics Will Rule with World,” by Daniel Combs, Safal Shetty, Sairam Parthasarathy, commentary in *Journal of Clinical Sleep Medicine*, volume 12, issue 02, 2016.
- “Could a pen change how we diagnose brain function? Team from MIT’s Computer Science and Artificial Intelligence Lab detects dementia using AI and a digital pen.” MIT News. August, 2015.
- “Could a Pen Change How We Diagnose Brain Function?” ACM TechNews, August 17 2015 (top article)
- “New Computer Tool Can Predict Dementia From Your Simple Drawings” *Popular Science*, August 13, 2015
- “Digital Pen is Better Dementia-Prediction Tool than a Doctor” *WIRED Magazine*, August 17, 2015
- “Computers that teach by example: New system enables pattern-recognition systems to convey what they learn to humans.” MIT News (also front page of MIT main website), December 5-10, 2014
- Blog entry about my work on recovery curves after prostatectomy on theheathcareblog.com: “Using Big Data to Manage Healthcare Expectations,” July 28, 2014
- “Data-driven tool enables patients to balance recovery expectations,” *FierceHealthIT*, July 2, 2014
- “New Statistical Model Lets Patient’s Past Forecast Future Ailments,” *Science News* section, *Science Daily*, June 9, 2012
- Blog entry about this work: “Using Data to Predict Your Future Health,” *Huffington Post*, August 1 2012

Other topics:

- Data Informed* blog entry: *For Analytics to Have an Impact, Keep it Simple*, discussing the session at the Joint Statistical Meetings that I organized. August 3, 2015 (Authors: Tyler H. McCormick, Cynthia Rudin, Dmitry Malioutov, and Kush Varshney)
- Article discussing the whitepaper effort I led: AmstatNews: The Membership Magazine of the American Statistical Association News. Cover Story, *Statistical Scientists Advance Federal Research Initiatives*, July 1, 2014.

- Article about whitepaper: “ASA Calls for Interdisciplinary Approach to Big Data Research,” *Data Informed*, September 17 2014
- Article about whitepaper: “4 high impact zones in statistical discovery with big data” *Fierce Big Data*, September 22 2014 (Authors: Sherri Rose, David Dunson, Tyler McCormick, and Cynthia Rudin)
- Article about my work: “How to Improve Product Rankings,” *Businessweek*, B-School Research Briefs section, March 9, 2012
- Quoted in article “From movies you’ll love to drugs you’ll take” about machine learning methods for ranking, *ScienceNews*, April 20 2010

Professional Societies and Government Committees

- Associate Director, Statistical and Applied Mathematical Sciences Institute (SAMSI)**, 2018-current
- Chair, Section on Statistical Learning and Data Mining, American Statistical Association**, 2017-2018.
- Member of Committee on Applied and Theoretical Statistics (CATS), National Academies of Sciences, Engineering, and Medicine** (2016-2018)
- Member of Committee on Law and Justice (CLAJ), National Academies of Sciences, Engineering, and Medicine** (2017-2019)
- Councilor of the AAI**. (2017-2020)
- Chair, INFORMS Data Mining Section**, 2015-2016 (Vice Chair for 2014-2015, Council Member 2017-2018, Council Member, 2011- 2013).
- Member of Committee on Analytical Research Foundations for the Next-Generation Electric Grid, and author of consensus report: “Analytical Foundations for the Next Generation Electric Grid”**, National Academy of Sciences, Engineering and Medicine, 2014-2016.
- Bureau of Justice Assistance Criminal Justice Technology Forecasting Group (BJA CJTFG)**, United States Department of Justice, 2014-2016.
- DARPA Information Science and Technology (ISAT) study group** (faculty advisory board of DARPA), 2014-2018.
- American Statistical Association Committee on Funded Research**, 2015-2018.

Activities

Events Organized

- Conference co-Chair, Conference on Statistical Learning and Data Science / Nonparametric Statistics**, co-Chair with Annie Qu, American Statistical Association, June 4-6, NYC, 2018
- Triangle Machine Learning Day**, co-organized with Jade Vinson, Richard Lucic, and Kirsten Shaw, April 3, 2018.
- Member of planning committee, SAMSI Program on Statistical, Mathematical, and Computational Methods for Precision Medicine** (2016-2018)
- TAMALE: Toolkit of Algorithms for Machine Learning**, DARPA ISAT workshop, co-organized with Margo Seltzer, March 2018.
- Judge for INFORMS Data Mining Best Paper competitions**, INFORMS, 2018.
- Member of organizing committee, Conference on Statistical Learning and Data Science**, UNC Chapel Hill. 2016.
- What if: Machine Learning Models for Causal Inference**, DARPA ISAT workshop, co-organized with Dustin Tingley, February 2016.
- The Cassandra Problem: Building Trust in Predictive Models**, DARPA ISAT workshop, co-organized with Carla Brodley and Stephen Boyd, April 2015.
- Invited Session: The Fifth V in “Big Data” is *Variables***, co-organizer with Tyler McCormick, Joint Statistical Meetings, 2015.
- Topic Contributed Session: Predictive Policing**, organizer, Joint Statistical Meetings, 2014.
- Judge for Statistical Learning and Data Mining Best Student Paper competition**, American Statistical Association, 2015.

Judge for 2015 INFORMS Innovative Applications in Analytics Award, 2014-2015.

Workshop on Data Analytics: Challenges in Big Data for Data Mining, Machine Learning and Statistics organizer, MIT CSAIL Big Data, March 26, 2014.

The ISBIS (International Society of Business and Industrial Statistics) 2014 and SLDM (Statistical Learning and Data Mining section of the American Statistical Association) Meeting on Data Mining in Business and Industry, Program Committee, June 9-11, 2014.

Statistical Analysis and Data Mining (a journal of the American Statistical Association), committee to choose the next editor-in-chief, 2014-2015.

ECML/PKDD 2013 Workshop on “DARE: Data Analytics for Renewable Energy Integration”, Technical program committee member, 2013.

Workshop on Hospital Readmission Prediction and Clinical Risk Management (HRPCRM) at IEEE International Conference on Healthcare Informatics (ICHI) 2013, program committee member, organizers are John Cromwell and Si-Chi Chin

Session on Smart Grid Data Analytics (SGDA) at International Conference on Smart Grid and Clean Energy Technologies (ICSGCE) 2013, co-chair with Zeyar Aung

MIT Energy Initiative, Energy Education Task Force, representative from Sloan, 2013-2014.

Dagstuhl Seminar on Preference Learning, co-organizer, 5 day seminar, 45 participants, Germany, March 3th to March 7th, 2014.

IMS/ASA Spring Research Conference, organizer of Machine Learning session, Harvard University, June 14th, 2012.

New England Machine Learning Day, Co-organizer, Microsoft Research New England, May 16th, 2012.

New England Statistics Symposium (NESS), organizer of Machine Learning session, Boston University, April 21, 2012.

Collective Intelligence 2012 (CI 2012), Local Arrangements Chair, 2012

INFORMS Data Mining and Health Informatics Workshop (DH-MI), co-organizer, 2011

MIT Energy Initiative, Organizing Committee for the MITEI Seminar Series, member, Fall 2010- Spring 2012

AMS Short Course on Aspects of Statistical Learning, organizer, 2007 AMS joint math meetings, New Orleans, January 3-4, 2007

AMS/AWM/MAA Special Session on Mathematical Results and Challenges in Learning Theory, session organizer, AMS joint math meetings, San Antonio Texas, January 12-15, 2006

Program for Women in Mathematics, Institute for Advanced Study, Program Committee Member, 2003-2006. Women in Science Seminar organizer, 2004, 2005, 2006. TA for the Wavelets course in 2002. Discussion group organizer, 2001. Research seminar speaker 2003. Poster session organizer 2003, Panel discussion organizer 2003, Panelist 2007

PACM Conference Princeton University, organizer, 2002-2004, speaker 2005

Editorial Responsibilities

Associate Editor for Management Science, in the Big Data Analytics department, 2018 - present

Action editor: 2013 - 2017, (Editorial board member: June 30 2010 - June 30 2013), Machine Learning Journal

Action editor: Statistical Analysis and Data Mining (SAM, a journal of the American Statistical Association), 2012-2017

Editor for Special Issue on Sports Analytics for Statistical Analysis and Data Mining (SAM, a journal of the American Statistical Association), co-editor with Theja Tulabandhula, 2015

Editorial Board, Journal of Artificial Intelligence Research (JAIR), July 2014 - June 2017

Member of Guest Editorial Board for ECMLPKDD 2014 journal track

Editor for Special Issue on Machine Learning for Science and Society, co-editor with Kiri Wagstaff, 2012-2013

Editorial board member: Journal of Machine Learning Research, 2012 - present

Program Committee Memberships

NIPS Workshop on Machine Learning for Healthcare ML4HC (2016, 2017, 2018), ECML-PKDD Workshop on Social Good (2016), ICML Workshop on Human Interpretability for Machine Learning (2016), ECML-PKDD Workshop on Data Science for Social Good (2016), SDM senior pc (2016), ICDM area chair (2015), AAAI senior pc (2016), NIPS area chair (2015), Visual Data Science (2015), IJCAI senior pc (2015), ICML area chair (2015) ICDM area chair (2014), IEEE Big

Data (2013), NIPS area chair (2013), AAAI (2013), ICML area chair (2013), NIPS area chair (2012), ECML-PKDD (2012), ICML (2012), COLT (2011), IJCAI (2011), ECML-PKDD (2010), ICML (2009), ICML (2006), AAAI (2006), KDD (2005)

Referee Panels

NSF (IIS - Intelligent Systems) panel in 2015, NSF (IIS - Intelligent Systems) panel in 2010, MIT International Science and Technology Initiatives (MISTI), grant review panel 2010, 2011, 2012, 2013.

Referee Assignments

Journal Reviewing: Journal of Artificial Intelligence Research (JAIR), Journal of Quantitative Criminology (JOQC), Data Mining and Knowledge Discovery (DAMI), Journal of Machine Learning Research (JMLR), Machine Learning Journal, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Mathematical Programming, Journal of Computational Statistics and Data Analysis, Communications on Pure and Applied Mathematics (CPAM), Management Science, International Journal of Renewable Energy Research (IJRER), Annals of Operations Research, Recent Patents on Computer Science

Conference and Other Reviewing: NIPS (Neural Information Processing Systems) 2018, NIPS 2017, NIPS 2011, NIPS 2010, NIPS 2009, NIPS 2008, NIPS 2007, ICDM (International Conference on Data Mining) 2011, Grant Proposal Review for Natural Sciences and Engineering Research Council of Canada 2016, Book Proposal for Springer 2013, COLT (Conference on Learning Theory) 2012, COLT 2010, COLT 2005, AISTATS (Conference on Artificial Intelligence and Statistics) 2012, AISTATS 2010, ICMLA (International Conference on Machine Learning and Applications) 2011, Book Proposal for Manning Publications 2010, ALT (Algorithmic Learning Theory) 2010, ICML (International Conference on Machine Learning) 2010, NIPS Ranking Workshop 2009, ACM SIGKDD (Conference on Knowledge Discovery and Data Mining) 2009, Applied and Computational Harmonic Analysis Journal, 2006, Conference on Machine and Statistical Learning: Prediction and Discovery 2006

Presentations

Keynotes/Invited for Notable Conferences

The European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), Dublin, September 11, 2018

Females Excelling More in Math, Engineering, and Science (FEMMES), keynote for capstone event, 200 girls grades 4-6, Duke University, February 17, 2018

International Symposium of Artificial Intelligence and Mathematics (ISAIM), January 3, 2018

Machine Learning for Healthcare (MLHC), August 18, 2017

Artificial Intelligence and Statistics (AISTATS), April 22, 2017

The Frontiers of Machine Learning, Sackler Forum on Machine Learning, National Academy of Sciences and The Royal Society, Jan 31-Feb 1, 2017

Fairness, Accountability and Transparency in Machine Learning (FAT-ML), November 18, 2016

Predictive Applications and APIs (PAPIs), October 2016

Discovery Science, Banff, Canada. October 5, 2015.

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), August 24-27, 2014.

Weathering the Data Storm: The Promise and Challenges of Data Science, Third Annual Symposium on the Future of Computation in Science and Engineering, Harvard University, Jan 26, 2014.

Other Conference Invited Presentations

(This list does not include regular or invited talks at annual meetings such as JSM, INFORMS, SLDM/SLDS, American Society for Criminology, Atlantic Causal Inference Conference, which I participate in every year.)

Nicholas Institute, Duke University, invited commenter on deep learning session, October 5, 2018

3rd International Workshop on Biomedical Informatics with Optimization and Machine Learning (BOOM), invited talk, IJCAI workshop, June 13, 2018

- Workshop on Human Interpretability in Machine Learning (WHI)**, invited talk, ICML workshop, June 17, 2018
- CVPR NTIRE 2018 New Trends in Image Restoration and Enhancement workshop and challenges on super-resolution, dehazing, and spectral reconstruction**, invited talk, representing our competition entry on superresolution, June 18, 2018
- Triangle Machine Learning Day**, invited talk, April 3, 2018
- 4th Annual Morgan Stanley Quantitative Equity Research Conference**, invited talk, November 17, 2016
- Japan-America Frontiers of Engineering Symposium, National Academy of Engineering**, invited talk, Beckman Center, Irvine CA, June 16, 2016
- Criminal Justice in the Age of Big Data**, panelist. Harvard Kennedy School. November 13, 2015
- Duke University Workshop on Sensing and Analysis of High-Dimensional Data**, Duke University, July 27-29, 2015
- Operations, Technology and Information Management Research Camp**, Johnson School of Management, Cornell University. June 23, 2015
- MIT Big Data Workshop – “From Data to Insights”**. May 21, 2015.
- MIT Media Lab – DARPA ISAT Workshop on Computational Health**. April 20, 2015
- Workshop on Big Data and Statistical Machine Learning**. (Part of the Thematic Program on Statistical Inference, Learning, and Models for Big Data.) Fields Institute, Toronto, January 29, 2015
- Bringing Social Science Back In: The ‘Big Data’ Revolution and Urban Theory**, Radcliffe Institute, Harvard University, December 15-16, 2014
- Understanding and Improving Cities: Policy/Research Partnerships in the Digital Age**, Invited Speaker/Panelist, District Hall, Boston, December 12, 2014
- New England Machine Learning Day**, May 1, 2013
- Exploration and Exploitation 3**, Presentation for Second Place team in Data Mining Contest, ICML Workshop, June 2012
- IMA Workshop on User-Centered Modeling**, May 2012
- IMA Workshop on Machine Learning**, March 2012
- ICML Workshop on Global Challenges**, International Conference on Machine Learning, July 2, 2011
- American Institute of Mathematics (AIM) workshop on the Mathematics of Ranking**, August 16-20, 2010
- DIMACS/CCICADA Workshop on Statistical Issues in Analyzing Information from Diverse Sources**, Rutgers University, May 6-7, 2010
- New England Statistics Symposium (NESS)**, Harvard University, April 17, 2010
- International Utility Working Group: Workshop on Computer-Aided Lean Management (CALM)**, Columbia University, April 16, 2008
- Conference on Machine and Statistical Learning: Prediction and Discovery**, AMS-IMS-SIAM Summer Research Conferences in the Mathematical Sciences, invited by organizers Joe Verducci, Xiaotong Shen, and John Lafferty, Snowbird, Utah, June 25-29th, 2006
- FOCM 2005 Foundations of Computational Mathematics, Workshop 4 on Learning Theory**, Santander Spain. Invited by organizer Steve Smale, 2005
- Machine Learning Summer School, Workshop on the Dynamics of Learning**, TTI-C, Chicago, invited by organizer Steve Smale, May 16-26, 2005

Invited Presentations at Universities, Societies, and Research Labs

- NSF Webinar: Statistics at a Crossroads: Challenges and Opportunities in the Data Science Era, October 2, 2018
- University of North Carolina Health Care and NC Women in Machine Learning and Data Science MeetUp, June 2, 2018
- University of North Carolina, Chapel Hill, Causal Inference Research Group Seminar, April 6, 2018
- North Carolina State University, Bioinformatics Seminar, March 29, 2018
- University of Maryland, Distinguished Speaker Series, Computer Science Department, December 1, 2017
- Temple University, Fox School of Business, Statistics Department Colloquium, October 6, 2017
- Duke University Algorithms Seminar, September 28, 2017
- University of Toronto Law School, Law and Economics Colloquium, September 26, 2017
- Microsoft Research, New York City, Sept 19, 2017

University at Buffalo, CSE, Distinguished Speaker Series, May 11, 2017
Columbia University, IEOR-DRO Seminar, May 9, 2017
Statistics C. V. Starr Lectureship Series, Biostatistics Department, Brown University, April 24 2017
Computational Social Science and Public Policy Colloquium, Harris School of Public Policy, University of Chicago, April 14, 2017
Statistics Department Seminar, University of Chicago, April 11, 2017
Duke Network Analysis Center Seminar, Duke University, February 6, 2017
Applied Mathematics Seminar, Duke University, October 3, 2016
Statistics and Complex Systems Seminar, University of Michigan, March 5, 2016
One Day University, New York, April 29, 2016
Urban Social Processes Workshop and Quantitative Methods Workshop, Harvard University, March 10, 2016
Center for Statistics and Machine Learning Seminar Series, Princeton University, October 20, 2015
Applied Statistics Workshop, Harvard University, October 14, 2015.
Brainstorming Session on the Next Generation of Search Engines, Berkman Center for Internet and Society, Harvard University, September 22, 2015
Machine Learning Seminar, Gatsby Unit, University College London. September 16, 2015
MIT Conversation Series, Accenture. July 10, 2015
ENAR webinar, (Eastern North American Region, International Biometric Society), May 8, 2015
Oracle Labs, Research Seminar, April 29, 2015
NYU-Poly Center for Urban Science and Progress, Research Seminar, March 12, 2015
American Express (NYC), Decision Science Seminar, March 11, 2015
Columbia University, IEOR-DRO Distinguished Seminar Series, March 10, 2015
University of Washington, Statistics Seminar, March 2, 2015
Duke University, Machine Learning Seminar, February 23, 2015
Harvard Business School, Technology and Operations Management Seminar, December 19, 2014
NYU Stern School, Department of Information, Operations & Management Sciences, IOMS Colloquium Series, December 3, 2014
UC Berkeley, Seminar in Computer Science Department, November 12, 2014
MIT Lincoln Labs, Seminar, November 4, 2014.
Brown University, Pattern Theory Group Seminar, October 22, 2014
University of Washington, Data Science Seminar, October 15, 2014
University of Alberta, Operations and Information Systems Seminar, October 3, 2014.
Carnegie Mellon University, ECE Seminar / Machine Learning Special Seminar, September 18, 2014
Harvard University, Computer Science Seminar, September 3, 2014
IBM TJ Watson Research Center, KDD Speaker Day, invited talk, August 28, 2014
MIT CSAIL CAP Meeting, May 30, 2014
MIT Lincoln Labs, Seminar, May 27, 2014
Stanford University, Operations Management Seminar, May 6, 2014
Stanford University, Institute for Research in the Social Sciences, Data Science and Inference Seminar, May 5, 2014
UMass Amherst, Machine Learning and Friends Lunch Seminar, April 29, 2014
MIT International Liaison Program Conference, Plenary Speaker, April 23, 2014
Schlumberger-Doll Research Center Seminar, April 14, 2014
Cornell University Operations Research and Industrial Engineering Seminar, April 8, 2014
Liberty Mutual Research Center Seminar, March 14, 2014
University of Pennsylvania Criminology Seminar, March 7, 2014
MIT Theory of Computation Seminar, March 4, 2014
Harvard Applied Statistics Workshop, October 9, 2013
Harvard/MIT Econometrics Workshop, September 12, 2013

MIT Lincoln Laboratory, Seminar, June 4, 2013
 Massachusetts General Hospital (MGH), Quantitative Medicine Seminar, April 29, 2013
 UT Austin McCombs School of Business, Research Seminar, April 5, 2013
 Laboratory for Information Decision Systems (LIDS) lunchtime seminar, February 26, 2013
 North Carolina State University, Statistics Department Seminar, February 28, 2013
 Yale Statistics Department Seminar, January 14, 2013
 Harvard High Dimensional and Correlated Data Seminar, December 17, 2012
 MIT Operations Research Center Seminar, December 11, 2012
 Yale Computer Science Department Seminar, December 6, 2012
 Columbia University, Statistics Department Seminar, October 22, 2012
 Robert H. Smith School of Business, University of Maryland, DO&IT Seminar Series, October 12, 2012
 MIT Center for Collective Intelligence, July 17, 2012
 Notre Dame Computer Science Department Seminar, November 3, 2011
 Rutgers Statistics Department Seminar, October 26, 2011
 Wharton Statistics Department Seminar, University of Pennsylvania, September 21, 2011
 MIT Energy Initiative External Board Meeting Speaker, October 15, 2010
 MIT Energy Initiative Seminar Series, October 12, 2010
 Boston University, Probability and Statistics Seminar, October 7, 2010
 Microsoft Research New England, Machine Learning Seminar, October 4, 2010
 Tufts, Computer Science Seminar, September 30, 2010
 ABB (Asea Brown Boveri Ltd) Corporate Research Center - United States, Lunch time Seminar, September 1, 2010.
 Harvard Statistics Colloquium, April 5, 2010
 MIT Operations Research Center Seminar, February 11, 2010
 MIT Imaging Seminar, October 22, 2009
 University of Chicago, Statistics Department Seminar, February 23, 2009
 Ohio State University, Computer Science Department Seminar, February 19, 2009
 Brown University, Applied Mathematics Seminar, February 17, 2009
 Indiana University, Computer Science Department Seminar, March 5, 2009
 University of Houston, Mathematics Department Seminar, Spring 2009
 Polytechnic University, Brooklyn, Computer Science Colloquium, April 30, 2007
 Columbia University, Applied Math Seminar, April 3, 2007
 New York University, Theory Seminar, Computer Science Department, November 9, 2005
 IBM Yorktown Heights, April 5, 2005
 Rensselaer Polytechnic Institute, March 8, 2005
 CCR (IDA Center for Communications Research), Princeton NJ, March 2, 2005
 Institute for Advanced Study, Computer Science/Discrete Math Seminar, Princeton, February 14, 2005
 Columbia University CCLS Center, February 4, 2005
 Google Labs Inc., Research Seminar, October 29, 2004
 New York University, Harmonic Analysis and Signal Processing Seminar, October 20, 2004
 SUNY at Buffalo, Math Department Seminar, May 2004
 NYU, Workshop on Computational and Biological Learning, January 16, 2004

Teaching and Mentoring

- **Machine Learning Summer School, Duke**, 2018 (2 days)
- **CS 290, Data Science Competition, Duke**, undergraduate course, 2018 (Spring)
- **CS 571 / STA 561 Machine Learning, Duke**, graduate and undergraduate machine learning, 2018 (Spring)

- **CS 571 / STA 561 Machine Learning, Duke**, graduate and undergraduate machine learning, 2016 (Fall)
- **Microsoft-DAT203x, Data Science and Machine Learning Essentials**, co-taught with Stephen Elston, free online course, edX, 2015. Over 17,500 students registered.
- **Microsoft-DAT203.2x Principals of Machine Learning**, co-taught with Stephen Elston, free online course, edX, 2016. Over 14,500 students registered.
- **Microsoft-DAT203.3x, Applied Machine Learning**, co-taught with Stephen Elston, free online course, edX, 2016. Over 8,000 students registered.
- **15.060 Data Models and Decisions, MIT**, MBA course (core course), Fall 2014, Instructor
- **15.075 Statistical Thinking and Data Analysis, MIT**, undergraduate course, Fall 2009, Fall 2010, Fall 2011, Spring 2013, Instructor.
- **15.097 Prediction: Machine Learning and Statistics, MIT**, graduate course, Spring 2012, Instructor. Course materials available on MIT Open Courseware.
- **15.060 Data Models and Decisions, MIT** MBA course (core course), Fall 2012, Instructor
- **15.064 Probability and Statistics, MIT**, Summer 2010, Summer 2011. masters student course (Leaders for Global Operations Program), Co-Instructor, 2010 and 2011
- **COMS 4771 Machine Learning, Computer Science Department, Columbia University**, Spring 2008, 4 lectures on regression, boosting, logistic regression, and ranking. Other lecturers were Alina Beygelzimer, John Langford, Tony Jebara.
- **Math 103 Calculus, Princeton**, Fall 2002, Fall 2001, Instructor
 - My lectures were videotaped and placed online. I was the first instructor at Princeton in the sciences to have their lectures videotaped. Class average was over 10 percentage points higher than the average of the other sections on a shared final exam that was worth 50% of their grade; this class was the top scoring class, and it scored 5 percentage points above the second highest class.
- **Math 199 Math Alive, Princeton**, Fall 2003, Teaching Assistant, responsible for the cryptography section, taught by Dr. Ingrid Daubechies
- **Wavelets Course, Program for Women in Mathematics, Institute for Advanced Study**, Summer 2002, Teaching Assistant, taught by Dr. Ingrid Daubechies
- **Physics Classes, Buffalo Seminary Women's High School**, Substitute Teacher, part-time during winter and spring, 1999, taught physics classes daily to freshmen (conceptual physics) and seniors (physics and advanced physics).

Service to Duke

Lead organizer for the Machine Learning seminar, 2016-present

Lead organizer of Triangle Machine Learning Day, 2018

Chair of faculty search committee, 2017-2018

Bass connections reviewing, 2016

Grad student admissions reviewing 2015-present

Committee for a faculty member's reappointment 2017-2018

Committee of Guillermo Sapiro, Vince Conitzer, and I, appointed by Provost Kornbluth to write "Computing For Humanity," 2017

Committee of 7 CS/ECE faculty members led by Carlo Tomasi to write a document similar to the above on AI, 2018

Review committee for a dean's reappointment, 2018

CS graduate awards committee, 2018

Member of numerous RIP, prelim, PhD, MS thesis, and undergraduate honors thesis committees, 2017-2018

Outreach such as hosting Duke Conversations, giving keynote for FEMMES at Duke, meeting with Visiting Committee, etc., 2018.

Supervision

Postdocs

Dr. Aaron Fisher, co-advised with Francesca Dominici, Harvard, 2016-present.
 Dr. Keivan Sadeghzadeh, MIT Sloan, 2016.
 Dr. Berk Ustun, Harvard CS, 2017-present.
 Dr. Noor-E-Alam, MIT Sloan, 2014-2015. (Now assistant professor at Northeastern University)
 Dr. Ramin Moghaddass, MIT Sloan, 2013-2015 (Now assistant professor at University of Miami).
 Dr. Şeyda Ertekin, MIT Sloan, 2010-2014. (Now assistant professor at Middle East Technical University).

Graduate Students

MS student Chunxiao Li, Duke Statistics student, 2018-present
 MS student Weiyu Yan, Duke ECE student, 2018-present
 PhD student Usaid Awan, Duke Econ student, 2018-present
 PhD student Zhi Chen, Duke CS PhD student, 2018-present
 PhD student Jiayun Dong, Duke Economics student, 2018-present
 MS student Kangcheng Lin, Duke Statistics student, 2018-present
 MS student Yang Bao, Duke Statistics student, 2018-present
 MS student Sijia Wang, Duke ECE student, 2018-present
 MS student Lei Chen, Duke ECE student, 2018-present
 MS student Xiyang Hu, Duke Statistics student, 2018-present
 PhD Student Alina Barnett, Duke CS student, 2017-present
 PhD Student Harsh Parikh, Duke CS student, 2018-present
 MS Student Yameng Liu, Duke Computer Science student, 2017-present
 PhD Student Lesia Semenova, Duke CS student, 2016-present
 PhD Student Chaofan Chen, Duke CS student, 2016-present
 PhD Student Tianyu Wang, Duke CS student, 2016-present
 MS Student Beau Coker, Duke Statistics student, 2017-2018
 PhD Student Marco Morucci, Duke Political Science student, 2017-present
 PhD Student Hongyu Yang, MIT EECS student, 2014-present
 MS Student Peter Alexander Lee, MIT ORC student, 2015-2016
 MS Student Prashan Wanigasekara, MIT EECS student, 2014-2016
 MS Student Christopher Choo, Engineering and Management, 2014-2015 (now at SUTD and Singapore Grand Prix)
 PhD Student Vikas Garg, MIT EECS student, 2014-2016 (co-advised with Tommi Jaakola)
 PhD Student Fulton Wang, MIT EECS student, 2013-2018. (Now at Sandia National Labs)
 PhD Student Berk Ustun, MIT EECS student, 2012-2017 (Now Postdoc at Harvard)
 PhD Student Stefano Traca, MIT ORC student, 2012-2018 (now working at Disney Research)
 PhD Student Siong Thye Goh, MIT ORC student, 2012-2018
 PhD Student Tong Wang, MIT EECS student, 2012-2016. (Now assistant professor at University of Iowa)
 Project Student Ashia Wilson, MIT Sloan, 2012. (5 months before starting a PhD program at Berkeley)
 PhD Student Theja Tulabandhula, MIT EECS student, 2010-2014. (Now senior lecturer at University of Sydney Business School)
 PhD Student Ben Letham, MIT ORC student, 2010-2015. (Now at Facebook)
 PhD Student Allison Chang, MIT ORC student, co-supervised with Dimitris Bertsimas, 2009-2012 (now at MIT Lincoln Labs).
 Masters Student William Harris, MIT ORC Student, co-advised with Michael Ricard, 2014-2015 (now in the US military)

MS Student Oscar Moll, MIT CSAIL student, 2010-2011.

MS Graduate Research Assistant, Nandini Bhardwaj, Columbia & Con Edison Secondary Events Project, 2008.

Masters Project Course, Jawwad Sultan, Columbia & Con Edison Secondary Events Project, Fall 2007.

Summer Students, Supervision of 2 masters students and 1 undergraduate. Columbia & Con Edison Secondary Events Project, Summer 2007.

Undergraduate Students

Duke undergraduate, Jerry Chia Rui Chang, 2018-present.

Duke undergraduate, Wilson Zhang, 2018-present

Duke undergraduate, Divya Koyalagunta, 2018-present.

Duke undergraduate, Anna Sun, 2018-present.

Duke undergraduate, Peter Hase, 2018-present.

Duke undergraduate, Daniel Tau, 2017-present.

Duke undergraduate, Caroline Wang, 2017-present.

Duke undergraduate, Hao Liu, summer 2017.

Duke undergraduate, Oscar Li, 2017-present.

MIT undergraduate, Chelsea Ge, summer 2014.

MIT undergraduate, Jeffrey Chan, spring-fall 2014.

MIT undergraduate, Jiaming Zeng, fall-spring 2015.

MIT undergraduate, Shawn Qian, summer-fall 2012.

Undergraduate exchange student, Yining Wang, spring 2013.

PhD student at MIT and previously undergraduate from Arizona State University, Lydia Letham, summer 2012, summer 2014.

MIT undergraduate project courses, three students (Kang Zhang, Arash Delijani, Kevin Pang) 2011-2012.

Undergraduate Visiting Student from Ecole Centrale Paris (through MISTI), Fabrice Vegetti, 2012.

Undergraduate Visiting Student from Ecole Centrale Paris (through MISTI), Adel Basli, 2011.

Undergraduate project course on Collaborative Filtering, Association Rules and Information Retrieval, Eugene Kogan, Columbia University, co-supervision with Dr. Ansaf Salleb-Aouissi, Spring 2008.

Undergraduate thesis advisement at Princeton, Krysta Svore, entitled "Multiscale Image Processing Using Single and Double Gaussian Techniques, and Hidden Markov Models," 2001-2002.

Tutoring

Assisted a high school student with preparation for her AP Physics exam, 1998-1999

Assisted a student with preparation for the physics section of her MCAT, 1998-1999

Mentoring

Mentor Advocacy Partnership, mentor for an MIT undergraduate minority student, 2011-present

Girls' Angle, A Math Club for Girls, lectured to ~20 girls ages 8-14, Cambridge MA, April 1 2010

Program for Women in Mathematics, Institute for Advanced Study, annually 2000-2008

Program Committee 2003-2008

Women in Science Seminar, organizer 2004, 2005, 2006

TA for the course on wavelets taught by Ingrid Daubechies, 2002

panel member 2007, discussion group organizer 2001, poster session organizer 2003, panel discussion organizer 2003, research seminar speaker "ERM, SRM, VC Dim, and SVM" 2002

WICS, Columbia (society for women in computer science), 2007-2009

Noetherian Ring, Princeton (society for women in mathematics), 1999-2004, officer 2002-2004

Thesis/Prelim Committees (not including my students)

Prelim committee for PhD student Paidamoyo Chapfuwa, ECE PhD student, 2018
Research initiation committee and prelim committee for Duke CS PhD student Andrew Lee, 2017, 2018
Graduation with honors committee for Duke undergraduate Wuming Zhang, 2018
Graduation with honors committee for Duke undergraduate Tianlin Duan, 2018
Prelim committee for Duke PhD student Xiaonan Hu, 2017, 2018
Prelim committee for Duke PhD student Greg Spell, 2018
Prelim committee for Duke PhD student Rachel Draelos, 2017
Prelim committee for Duke PhD student Stavros Sintos, 2017
Prelim committee for Duke CS PhD student Zilong Tan, 2017
Thesis committee for Stanford CS student Himabindu Lakkaraju, PhD in spring 2018
Research Proposal committee for Duke CS PhD student Xiaonan Hu, 2017
Research Proposal committee for Duke CS PhD student Shuzhi Yu, 2017
Research Proposal committee for Duke CS PhD student Abe Frandsen, 2017, 2018
Thesis committee for Duke CS undergraduate with distinction Aditya Mukund, 2017
Thesis committee for Duke CS masters student Guan-Wun Hao, 2017
Thesis committee for Duke CS masters student Mona Prakash, 2016
Thesis committee for Duke statistics masters student Emily Shao, 2017
Thesis committee for Duke statistics masters student Sanjay Harihanan, 2017
Thesis committee for Duke ECE Phd student Jordan Hashemi, 2017
Thesis committee for Duke ECE Phd student Zhuoqing Chang, 2017, 2018
Thesis reader for Harvard CS undergraduate Nicholas Larus-Stone, 2017
Thesis committee for Duke PhD student Shan Shan, 2017
Thesis committee for Duke PhD student Narayanan Rengaswamy, 2017.
Thesis committee for MIT PhD student Yingxiang Yang, 2015.
Thesis committee for MIT PhD student Been Kim, PhD in spring 2015.
Thesis committee for MIT PhD student Anima Singh, PhD in spring 2015.
Thesis committee for Pannaga Shivaswamy at Columbia University CS, PhD in spring 2009.

Society Memberships

- INFORMS
- International Machine Learning Society
- American Statistical Association (ASA)
- Institute of Mathematical Statistics (IMS)
- Association for the Advancement of Artificial Intelligence (AAAI)
- Association for Computing Machinery (ACM)