

# Cynthia Rudin

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

(Secondary Appointments in Statistics and Mathematics at Duke University)  
Berkman 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

Machine learning (particularly interpretable machine learning), statistical learning theory, optimization methods in machine learning, Bayesian hierarchical modeling, causal inference.

Application areas: personalized medicine, computational social science, criminology, energy grid reliability

## Education

- **Ph.D.: Princeton University.** Program in Applied and Computational Mathematics. Advisors: Ingrid Daubechies and Robert Schapire. Additional Committee Members: 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, 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

## Research Highlights

- My colleagues and I design methods and software packages for interpretable machine learning that have a strong theoretical foundation. These new methods aim to replace the state of the art decision tree methods that are widely used in industry for predictive modeling (we provide competitor methods for CART - classification and regression trees, along with C4.5, C5.0, etc.).

I have many other projects on interpretable predictive modeling of various forms, including axis parallel rectangles, prototype models, and recovery curve models. These projects have won several best paper awards and the 2016 INFORMS Innovative Applications in Analytics Award.

- I led a team of scientists in a large scale project to assist with energy grid reliability in with NYC, jointly with power company Con Edison. This was a Big Data (knowledge discovery) project involving large amounts of heterogeneous historical data. This project won the INFORMS Innovative Applications in Analytics Award in 2014, was on the cover of IEEE Computer, and discussed in popular books and media articles about Big Data.
- My student Theja Tulabandhula and I designed a theoretical framework for a new exploratory kind of decision theory for “Machine Learning for Operational Costs.” It bridges a gap between prediction making and decision making.
- My NSF CAREER award is on the topic of learning-to-rank, where the goal is to optimize rank statistics to prioritize items in a better way. The P-Norm Push paper (Rudin 2009) in particular shows how prediction ability trades off with how much the rank statistic concentrates at the top of the ranked list. I also discovered and proved equivalence relationships between algorithms for ranking and for classification.
- My thesis work was on convergence analysis of AdaBoost, which is one of the “top 10” machine learning algorithms. I answered a well-known theoretical problem that was open for almost a decade, which was to determine whether AdaBoost always maximizes the machine learning “margin.” I answered this problem using tools from dynamical systems. Later, some of my joint work answered a published open problem at COLT (Computational Learning Theory), earning a monetary award.
- My colleagues and I designed a “next generation” search engine. It is currently the best search engine for the problem of set completion, as compared with Google Sets and Boo!Wa!. The paper (called “Growing a List”) was featured on Boston Public Radio in an article called “A New Way to Google”.
- I worked with the Cambridge Police Department on detecting patterns of crime. Our work has been featured several times in the local, national, and international press.
- I have published over 60 peer reviewed journal and conference papers.
- I have been the PI on over \$5 million in grant funding.

## *Publications*

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

1. Cynthia Rudin, Prashan Wanigasekara, Jacob Bien. **Elastic Prototype Classification**. Working paper, 2015.
2. Fulton Wang and Cynthia Rudin. **Causal Falling Rule Lists**, Working paper on ArXiv, 2015.
  - Shorter version presented at Fairness, Accountability, and Transparency (FATML), 2017.
3. Siong Thye Goh and Cynthia Rudin. **Cascaded High Dimensional Histograms: An Approach to Interpretable Density Estimation for Categorical Data**, Working paper on ArXiv, 2015.
4. Md. Noor-E-Alam and Cynthia Rudin. **Robust Testing for Causal Inference in Natural Experiments**. Working Paper on Optimization Online, 2015.
5. Md. Noor-E-Alam and Cynthia Rudin. **Robust Nonparametric Testing for Causal Inference in Observational Studies**. Working paper, 2015.
6. Cynthia Rudin and Gah-Yi Vahn. **The Big Data Newsvendor: Practical Insights from Machine Learning Analysis**, MIT Sloan School of Management Working Paper 5036-13. In revision.
7. Fulton Wang, Tyler McCormick, Cynthia Rudin, and John Gore. **Modeling Recovery Curves With Application to Prostatectomy**. In revision, 2016.
  - Winner of Best Poster Award, International Symposium on Business and Industrial Statistics / Conference of the ASA Section on Statistical Learning and Data Mining, 2014. (\$400 award to Fulton)
8. Allison Chang, Cynthia Rudin, and Dimitris Bertsimas. **A Mixed-Integer Optimization Approach to Supervised Ranking**.

- Shorter version: Dimitris Bertsimas, Allison Chang, and Cynthia Rudin, **A Discrete Optimization Approach to Supervised Ranking**, Proceedings of the 5th INFORMS Workshop on Data Mining and Health Informatics (DM-HI 2010) D. Sundaramoorthi, M. Lavieri, H. Zhao, eds, 2010.
  - Shorter version of this work selected as a finalist for *Data Mining Best Student Paper Competition, INFORMS 2011*.
  - This paper accepted in the presentation-only category for a full presentation at the International Conference on Machine Learning (ICML) 2012.
9. Cynthia Rudin and Yining Wang. **On Direct Learning to Rank and Rerank**. Working Paper, 2015.
  10. Theja Tulabandhula and Cynthia Rudin. **Robust Optimization using Machine Learning for Uncertainty Sets**. Work in Progress 2015. ArXiv: <http://arxiv.org/abs/1407.1097>
  11. Stefano Tráca and Cynthia Rudin. **Regulating Greed over Time**. Working Paper on ArXiv, 2015.
    - Winner of best paper award, INFORMS 2016 Data Mining & Decision Analytics (DMDA) Workshop
    - Finalist for 2015 IBM Service Science Student Best Paper Award
  12. Cynthia Rudin and Şeyda Ertekin. **Learning Optimized Lists of Rules**. In revision, 2015.
  13. Ramin Moghaddass and Cynthia Rudin. **Bayesian Patchworks**. In revision, 2016.

## *Journal Publications and Conference Proceedings*

### 2017

14. 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. **A Practical Risk Score for EEG Seizures in Hospitalized Patients**, JAMA Neurology, 2017 (accepted).
15. Elaine Angelino, Nicholas Larus-Stone, Daniel Alabi, Margo Seltzer, and Cynthia Rudin. **Certifiably Optimal Rule Lists for Categorical Data**, Knowledge Discovery in Databases (KDD), 2017.
16. Berk Ustun and Cynthia Rudin. **Learning Optimized Risk Scores from Large-Scale Datasets**. Knowledge Discovery in Databases (KDD), 2017.
17. Hongyu Yang, Cynthia Rudin, and Margo Seltzer. **Scalable Bayesian Rule Lists**. International Conference on Machine Learning (ICML), 2017.
  - Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2016.
18. 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.
19. Himabindu Lakkaraju and Cynthia Rudin. **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.
20. 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), accepted, 2017.

2016

21. 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.
22. 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.
23. William Souillard-Mandar, Randall Davis, Cynthia Rudin, Rhoda Au, David J. Libon, Rodney Swenson, Catherine C. Price, Melissa Lamar, Dana L. Penney. **Learning Classification Models of Cognitive Conditions from Subtle Behaviors in the Digital Clock Drawing Test**. Machine Learning, Volume 102, Number 3, 2016.
  - Part of winning entry for 2016 INFORMS Innovative Applications in Analytics Award.
24. Jiaming Zeng, Berk Ustun, and Cynthia Rudin. **Interpretable Classification Models for Recidivism Prediction**. Journal of the Royal Statistical Society Series A, volume 180, issue 3, 2016.
  - This paper won the 2015 Undergraduate Statistics Research Project Competition (USRESP) sponsored by the American Statistical Association (ASA) and the Consortium for Advancement of Undergraduate Statistics Education (CAUSE).
25. Berk Ustun and Cynthia Rudin. **Supersparse Linear Integer Models for Optimized Medical Scoring Systems**. Machine Learning, Volume 102, Number 3, 2016.
  - Part of winning entry for 2016 INFORMS Innovative Applications in Analytics Award.
  - Other versions:
    - Berk Ustun and Cynthia Rudin. **Methods and Models for Interpretable Linear Classification**, 2014. ArXiv: <http://arxiv.org/abs/1405.4047>.
    - Berk Ustun, Stefano Tráca and Cynthia Rudin. **Supersparse Linear Integer Models for Interpretable Classification**, 2013. ArXiV: <http://arxiv.org/pdf/1306.6677v1.pdf> .
    - Berk Ustun, Stefano Tracà, and Cynthia Rudin. **Supersparse Linear Integer Models for Predictive Scoring Systems**, Proceedings of late breaking track, Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), 2013.
26. 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.
27. Vikas Garg, Cynthia Rudin, and Tommi Jaakola. **CRAFT: ClusteR-specific Assorted Feature selectiOn**, Artificial Intelligence and Statistics (AISTATS), 2016.
28. 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.
29. 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
30. 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

31. Fulton Wang and Cynthia Rudin. **Falling Rule Lists**. Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS), 2015.

- Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2015.
  - Finalist for INFORMS Data Mining Section Best Student Paper Award, 2015.
32. Benjamin Letham, Cynthia Rudin, Tyler H. McCormick, and David Madigan. **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.
- Winner of Data Mining Best Student Paper Competition, INFORMS 2013.
  - Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, 2014.
  - Shorter versions of this work:
    - Benjamin Letham, Cynthia Rudin, Tyler McCormick, and David Madigan. **Interpretable Classifiers using Rules and Bayesian Analysis: Building a Better Stroke Prediction Model**. Proceedings of late breaking track, Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), 2013.
    - 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
33. Şeyda Ertekin and Cynthia Rudin. **A Bayesian Approach to Learning Scoring Systems**. *Big Data*, Volume 3, Number 4, 2015.
34. 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.
35. 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.
- Shorter version:
- Şeyda Ertekin, Cynthia Rudin, and Tyler McCormick. **Reactive Point Processes: A New Approach to Predicting Power Failures in Underground Electrical Systems**, Proceedings of late breaking track, Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), 2013.
36. Tong Wang, Cynthia Rudin, Daniel Wagner and Rich Sevieri. **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)
- Second place in INFORMS 2015 Doing Good with Good OR Paper Competition

## 2014

37. 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.
38. Şeyda Ertekin, Cynthia Rudin and Haym Hirsh. **Approximating the Crowd**. *Data Mining and Knowledge Discovery (ECML-PKDD journal track)*, volume 28 issue 5-6, pages 1189-1221, September 2014.
- Şeyda Ertekin, Haym Hirsh, and Cynthia Rudin. **Approximating the Wisdom of the Crowd**, Proceedings of the Second Workshop on Computational Social Science and the Wisdom of Crowds (NIPS) 2011.
  - Şeyda Ertekin, Haym Hirsh, and Cynthia Rudin. **Learning to Predict the Wisdom of Crowds**, Proceedings of Collective Intelligence (CI) 2012.
  - Şeyda Ertekin, Haym Hirsh, and Cynthia Rudin. **Selective Sampling of Labelers for Approximating the Crowd**, Proceedings of the 2012 AAAI Fall Symposium on Machine Aggregation of Human Judgment, MAGG-2012.

39. 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.
40. Theja Tulabandhula, Cynthia Rudin. **On Combining Machine Learning with Decision Making**. Machine Learning (ECML-PKDD journal track), Volume 93, Pages 33-64, 2014

Shorter version (very preliminary):

- Theja Tulabandhula, Cynthia Rudin, and Patrick Jaillet. **The Machine Learning and Traveling Repairman Problem**, Proceedings of the 2nd International Conference on Algorithmic Decision Theory (ADT), 2011.
41. 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.
  42. 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.
43. Theja Tulabandhula, Cynthia Rudin. **Robust Optimization using Machine Learning for Uncertainty Sets**. Shorter version:
    - Proceedings of ISAIM 2014, Special Session on Theory of Machine Learning, 2014.
    - Proceedings of INFORMS Data Analytics workshop, 2014.
  44. Cynthia Rudin, Şeyda Ertekin, Rebecca Passonneau, Axinia Radeva, Ashish Tomar, Boyi Xie, Stanley Lewis, Mark Riddle, Debbie Pangsrivini, Tyler McCormick. **Analytics for Power Grid Distribution Reliability in New York City**. Interfaces, Volume 44, Issue 4, pages 364-383, 2014.

- This work part of the winning entry of the *INFORMS 2013 Innovative Applications in Analytics Competition*.

45. Jonathan Huggins and Cynthia Rudin. **Towards a Theory of Pattern Discovery** Proceedings of SIAM Conference on Data Mining (SDM) 2014.  
also Proceedings of ISAIM 2014, Special Session on Theory of Machine Learning, 2014.  
also: ArXiv: <http://arxiv.org/pdf/1307.0802v1.pdf>
46. 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.

A shorter version of this work is in the following conference paper.

- Been Kim and Cynthia Rudin. **Learning About Meetings**, Proceedings of late breaking track, Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), 2013.

## 2013

47. Benjamin Letham, Cynthia Rudin and Katherine Heller. **Growing a List**. Data Mining and Knowledge Discovery (DAMI), ECML-PKDD 2013 journal track (acceptance rate < 8%). Volume 27, pages 372-395, 2013.
48. 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), 2013.

- Tong Wang, Cynthia Rudin, Daniel Wagner, Richard Sevieri. **Detecting Patterns of Crime with Series Finder**, Proceedings of late breaking track, Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), 2013.
49. 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.  
Preliminary work is in the following conference paper:
- Cynthia Rudin, Ben Letham, Ansaf Salleb-Aouissi, Eugene Kogan, and David Madigan. **Sequential Event Prediction with Association Rules**, Proceedings of the 24th Annual Conference on Learning Theory (COLT), 2011.
50. Theja Tulabandhula and Cynthia Rudin. **Machine Learning with Operational Costs**. Journal of Machine Learning Research (JMLR), Volume 14, pages 1989-2028, July 2013.
- Shorter version of this work selected as a finalist for *Data Mining Best Student Paper Competition, INFORMS 2012*.
- 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.
51. Benjamin Letham, Cynthia Rudin and David Madigan. **Sequential Event Prediction**. Machine Learning, Volume 93, pages 357-380, 2013
52. Indraneel Mukherjee, Cynthia Rudin, and Robert E. Schapire. **The Rate of Convergence of AdaBoost**, Journal of Machine Learning Research, Volume 14, pages 2315-2347, August 2013.
- Preliminary work is in the following conference paper. This paper answered an open question published in COLT 2010 with a monetary award.
- 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.

## 2012

53. 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.  
Shorter version:
- Tyler McCormick, Cynthia Rudin, and David Madigan, **Predicting Medical Conditions with Bayesian Hierarchical Rule Modeling** In Proceedings of the 6th INFORMS Workshop on Data Mining and Health Informatics (DM-HI), 2011.
54. Cynthia Rudin, David Waltz, Roger N. Anderson, Albert Boulanger, Ansaf Salleb-Aouissi, Maggie Chow, Haimonti Dutta, Philip Gross, Bert Huang, Steve Jerome, 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*.
55. 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

56. 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.)*
57. 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

58. 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.
  - This paper accepted in the presentation-only category for a full presentation at the International Conference on Machine Learning (ICML) 2012.

The two journal articles above are related to energy grid reliability. The following conference papers are also related to my projects on grid reliability.

59. 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.
60. 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.
61. 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.
62. 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.
63. Leon Wu, Gail Kaiser, Cynthia Rudin, David Waltz, Roger Anderson, Albert Boulanger, Ansaif Salleb-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.
64. 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.*
65. 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.
66. 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.



67. 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

68. 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.

Preliminary material for this work appears partly within this conference paper:

- 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.

69. Raphael Pelossof, Michael Jones, Ilia Vovsha, Cynthia Rudin. **Online coordinate boosting**, Proceedings of IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), page 1354-1361, 2009.

70. 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 material for this work appears partly within this conference paper:

- 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

71. 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 for this work appears partly within these shorter works:

- 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.

72. 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.
73. 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.
74. Cynthia Rudin and Brian Spencer. **Equilibrium Ridge Arrays in Strained Solid Films**. Journal of Applied Physics, vol 86, pp 5530-5536, 1999.

### *Other Publications (not listed above)*

74. 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.
75. 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.
76. 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.
77. Cynthia Rudin. **Teaching “Prediction: Machine Learning and Statistics”**, Proceedings of the ICML Workshop on Teaching ML, 2012.
78. Peter Qian, Yilu Zhou, and Cynthia Rudin, **Proceedings of the 6th INFORMS Workshop on Data Mining and Health Informatics (DM-HI)**, editors, 2011.
79. 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.
80. Cynthia Rudin. **A Different Kind of Convergence for Statistical Learning Algorithms**. Math arXiv: cs.LG/0502016, 2003.

### *Grants*

**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). Amount of award: \$55,500

**2016 Adobe Digital Marketing Research Award**, 2016. Amount of Award: \$50K

**MIT-Lincoln Labs** “Adaptable, Interpretable, Machine Learning” (co-PI with Jonathan Su from MIT-LL), 2016. Amount of Subaward: \$145K

**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

**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” (co-PI with postdoc Şeyda Ertekin), 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 2016.  
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

2016 Adobe Digital Marketing Research Award, 2016

Winner of 2016 INFORMS Innovative Applications in Analytics Award, 2016.

Berkman faculty associate, Berkman Center for Internet and Society at Harvard University, 2015-2017.

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

National Science Foundation CAREER Award, 2011

Winner of 2013 INFORMS Innovative Applications in Analytics Award, 2013.

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

Paper “Regulating Greed Over Time” Winner of Best Paper Award, Data Mining & Decision Analytics (DMDA) Workshop, INFORMS 2016.

Paper “Scalable Bayesian Rule Lists” Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, Awarded to Hongyu Yang, 2016.

Paper “Finding Patterns with a Rotten Core: Data Mining for Crime Series with Core Sets”, Second place in Doing Good with OR Competition, Awarded to Tong Wang, INFORMS 2015.

Paper “Regulating Greed Over Time” was finalist for IBM Best Service Science Best Student Paper Award, 2015.

- Paper “Interpretable Classification Models for Recidivism Prediction” won first place in the 2015 Undergraduate Statistics Research Project Competition (USRESP) sponsored by the American Statistical Association (ASA) and the Consortium for Advancement of Undergraduate Statistics Education (CAUSE), Awarded to Jiaming Zeng, 2015.
- Paper “Falling Rule Lists” Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, Awarded to Fulton Wang, 2015.
- Paper “An Interpretable Stroke Prediction Model using Rules and Bayesian Analysis,” Winner of INFORMS Data Mining Best Student Paper Competition, Awarded to Ben Letham, 2013.
- Paper “An Interpretable Stroke Prediction Model using Rules and Bayesian Analysis,” Winner of Student Paper Competition sponsored by the Statistical Learning and Data Mining section (SLDM) of the American Statistical Association, Awarded to Ben Letham, 2014.
- Nominated for Outstanding UROP Mentor Award, UROP (undergraduate research opportunities) Program, MIT, 2012
- Work with Indraneel Mukherjee and Robert Schapire answered an open problem from COLT 2010, with a monetary award (\$100).
- Best Poster Award, International Symposium on Business and Industrial Statistics / Conference of the ASA Section on Statistical Learning and Data Mining, Awarded to Fulton Wang, 2014.
- Best Poster Award, International Conference of Machine Learning and Applications (ICMLA) 2009
- 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
- Nominated for 2012 Sloan Excellence in Teaching Awards, MIT, 2012

## 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, Journalism 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 (note that title of article is misleading)
- “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 thehealthcareblog.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 Activities*

### *Events Organized, Professional Society Responsibilities, and Government Committees*

- Chair-Elect for 2017 (chair in 2018), Section on Statistical Learning and Data Mining, American Statistical Association.**
- Member of Committee on Applied and Theoretical Statistics (CATS), National Academies of Sciences, Engineering, and Medicine (2016-2017)**
- Member of Committee on Law and Justice (CLAJ), National Academies of Sciences, Engineering, and Medicine (2017-2019)**
- Chair for 2015-2016 (Vice Chair for 2014-2015) INFORMS Data Mining Section.**
- Member of organizing committee, Conference on Statistical Learning and Data Science, UNC Chapel Hill. 2016.**
- Member of committee on “Analytical Foundations for the Next Generation Electric Grid” appointed by The Board on Mathematical Sciences and Their Application (BMSA) of the National Research Council (National Academy of Sciences), 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), 2014-2017.
- 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.
- American Statistical Association Committee on Funded Research**, 2015-2018.
- 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.
- INFORMS Data Mining Section**, Council Member, (2011- 2013)
- Collective Intelligence 2012 (CI 2012)**, Local Arrangements Chair, 2012
- INFORMS Data Mining and Health Informatics Workshop (DH-MI)**, co-organizer, 2011
- MIT International Science and Technology Initiatives (MISTI) grant reviewer** (9 proposals), 2012
- MIT International Science and Technology Initiatives (MISTI) grant reviewer** (11 proposals), 2011
- MIT International Science and Technology Initiatives (MISTI) grant reviewer** (6 proposals), 2010
- 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*

Action Editor for Journal of Machine Learning Research, starting 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

Committee to choose the next editor in chief of Statistical Analysis and Data Mining (an American Statistical Association Journal), 2014-2015

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

Action editor: 2013 - present, (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-present

### *Program Committee Memberships*

Machine Learning for Healthcare (ML4HC), Senior PC, 2017

Workshop on Social Good, ECML-PKDD, 2016

Workshop on Human Interpretability for Machine Learning at ICML, 2016

Workshop on Data Science for Social Good, ECML-PKDD 2016

SDM (SIAM International Conference on Data Mining) 2016 (senior PC member)

ICDM (International Conference on Data Mining) 2015 (area chair)

AAAI (AAAI Conference on Artificial Intelligence) 2016 (senior program committee)

NIPS (Neural Information Processing Systems) 2015 (area chair)

Visual Data Science (VDS) 2015

IJCAI (International Joint Conferences on Artificial Intelligence) 2015 (senior program committee)

ICML (International Conference on Machine Learning) 2015 (area chair)

ICDM (International Conference on Data Mining) 2014 (area chair)

IEEE Big Data 2013

NIPS (Neural Information Processing Systems) 2013 (area chair)

AAAI (Assoc. for the Advancement of Artificial Intelligence) 2013

ICML (International Conference on Machine Learning) 2013 (area chair)

NIPS (Neural Information Processing Systems), 2012 (area chair)

ECML-PKDD (European Conference on Machine Learning, and Principles and Practice of Knowledge Discovery in Databases) 2012

ICML (International Conference on Machine Learning) 2012

COLT (Conference on Learning Theory) 2011

IJCAI (International Joint Conferences on Artificial Intelligence) 2011

ECML-PKDD (European Conference on Machine Learning, and Principles and Practice of Knowledge Discovery in Databases) 2010

ICML (International Conference on Machine Learning) 2009

ICML (International Conference on Machine Learning) 2006

AAAI (Association for the Advancement of Artificial Intelligence) 2006

ACM-SIGKDD Research Track 2005

### *Referee Panels*

NSF (IIS - Intelligent Systems) panel in 2015  
 MIT International Science and Technology Initiatives (MISTI) panel in 2013  
 NSF (IIS - Intelligent Systems) panel in 2010

### *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

#### Conferences and Proposal Reviewing:

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

### *Presentations*

#### *Keynotes/Invited*

**Artificial Intelligence and Statistics (AISTATS), April 22, 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)**, (one of 7 talks), 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, Plenary lecture, Harvard University, Jan 26, 2014.

**New England Machine Learning Day**, Plenary lecture, May 1, 2013.

### *Other Conference/Workshop Invited Presentations*

(Not including annual meetings such as JSM, INFORMS, SLDM/SLDS, American Society for Criminology, Atlantic Causal Inference Conference, etc.)

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

**4th Annual Morgan Stanley Quantitative Equity Research Conference**, invited talk, November 17, 2016

**Statistical Learning and Data Science (SLDS)**, American Statistical Association, University of North Carolina at Chapel Hill, June 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.

**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 and Research Labs*

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  
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*

- **CS 570 / STA 560 Machine Learning, Duke**, graduate and undergraduate machine learning, 2016
- **Applied Machine Learning**, co-taught with Stephen Elston, free online course, Edx 2016
- **Data Science Essentials**, co-taught with Stephen Elston, free online course, Edx 2016

- **Introduction to Machine Learning**, co-taught with Stephen Elston, free online course, Edx 2016
- **Edx Microsoft-DAT203x, Data Science and Machine Learning Essentials**, co-taught with Stephen Elston, free online course, Fall 2015. Over 35,000 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. Course materials available on MIT Open Courseware.
- **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
- **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).

## *Supervision*

### *Postdocs*

- Dr. Keivan Sadeghzadeh, MIT Sloan, 2016.  
 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*

- PhD Student Hongyu Yang, EECS student, 2014-present  
 Masters Student Peter Alexander Lee, ORC student, 2015-2016  
 Masters Student Prashan Wanigasekara, EECS student, 2014-2016  
 Masters Student Christopher Choo, Engineering and Management, 2014-2015 (now at SUTD and Singapore Grand Prix)  
 PhD Student Vikas Garg, EECS student, 2014-2016 (co-advised with Tommi Jaakola)  
 PhD Student Fulton Wang, EECS student, 2013-present  
 PhD Student Berk Ustun, EECS student, 2012-present  
 PhD Student Stefano Traca, ORC student, 2012-present  
 PhD Student Siong Thye Goh, ORC student, 2012-present  
 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 assistant professor at University of Illinois at Chicago)

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)

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

Masters 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*

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, co-supervision with Dr. Ansaf Salieb-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)*

Thesis committee for Stanford CS student Himabindu Lakkaraju, 2017

Thesis committee for Duke CS PhD student Xiaonan Hu, 2017

Thesis committee for Duke CS PhD student Shuzhi Yu, 2017  
Thesis committee for Duke CS PhD student Abe Frandsen, 2017  
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  
Thesis reader for Harvard CS undergraduate Nicholas Larus-Stone, 2017  
Thesis committee for Shan Shan at Duke, 2017  
Thesis committee for Narayanan Rengaswamy at Duke, 2017.  
Thesis committee for Yingxiang Yang at MIT, 2015.  
Thesis committee for Been Kim at MIT Aero/Astro, PhD in spring 2015.  
Thesis committee for Anima Singh at MIT EECS, PhD in spring 2015.  
Thesis committee for Pannaga Shivaswamy at Columbia University Computer Science Department, PhD in spring 2009.

## *Society Memberships*

- INFORMS
- International Machine Learning Society
- American Statistical Association (ASA)
- Institute of Mathematical Statistics (IMS)
- American Society of Criminology (ASC)
- Association for Computing Machinery (ACM)