Daniel J. Hsu

Department of Computer Science
Columbia University
500 W. 120th St., Mailcode: 0401
New York, NY 10027

Phone: +1 (212) 853-8473
E-mail: djhsu@cs.columbia.edu
WWW: https://www.cs.columbia.edu/~djhsu/

Research Interests

Algorithmic statistics and machine learning.

Academic Positions

Data Science Institute, Columbia University, New York, NY.

Associate Director for Research, September 2025 to present.

Department of Computer Science, Columbia University, New York, NY.

Associate Professor, January 2018 to present. (Tenured in 2020.)

Assistant Professor, July 2013 to December 2017.

Microsoft Research, Cambridge, MA.

Postdoctoral Researcher, June 2011 to July 2013.

Department of Statistics, Rutgers University, Piscataway, NJ.

Postdoctoral Associate, July 2010 to June 2011.

Department of Statistics, University of Pennsylvania, Philadelphia, PA.

Visiting Postdoctoral Scholar, July 2010 to June 2011.

Education

Ph.D. Computer Science, University of California, San Diego, June 2010.

Advisor: Prof. Sanjoy Dasgupta.
Dissertation title: Algorithms for Active Learning.

M.S. Computer Science, University of California, San Diego, June 2007.

B.S. Computer Science and Engineering, University of California, Berkeley, May 2004.

Highest Honors in Electrical Engineering and Computer Sciences.

Publications

Journal Articles

Joan Bruna and Daniel Hsu. Survey on algorithms for multi-index models. *Statistical Science*, 40(3):378–391, 2025.

Gan Yuan, Mingyue Xu, Samory Kpotufe, and Daniel Hsu. Efficient estimation of the central mean subspace via smoothed gradient outer products. *SIAM Journal on Mathematics of Data Science*, 7(3):1241–1264, 2025.

Giannis Karamanolakis, Daniel Hsu, and Luis Gravano. Interactive machine teaching by labeling rules and instances. *Transactions of the Association for Computational Linguistics*, 12:1441–1459, 2024.

Rishabh Dudeja and Daniel Hsu. Statistical-computational trade-offs in tensor PCA and related problems via communication complexity. *Ann. Statist.*, 52(1):131–156, 2024.

Daniel Hsu, Vidya Muthukumar, and Ji Xu. On the proliferation of support vectors in high dimensions. *Journal of Statistical Mechanics: Theory and Experiment*, 2022(11):114011, 2022.

Michał Dereziński, Manfred K. Warmuth, and Daniel Hsu. Unbiased estimators for random design regression. *Journal of Machine Learning Research*, 23(167):1–46, 2022.

Christopher Tosh, Akshay Krishnamurthy, and Daniel Hsu. Contrastive estimation reveals topic posterior information to linear models. *Journal of Machine Learning Research*, 22(281):1–31, 2021.

Vidya Muthukumar, Adhyyan Narang, Vignesh Subramanian, Mikhail Belkin, Daniel Hsu, and Anant Sahai. Classification vs regression in overparameterized regimes: Does the loss function matter? *Journal of Machine Learning Research*, 22(222):1–69, 2021.

Rishabh Dudeja and Daniel Hsu. Statistical query lower bounds for tensor PCA. *Journal of Machine Learning Research*, 15(83):1–51, 2021.

Ji Xu, Arian Maleki, Kamiar Rahnama Rad, and Daniel Hsu. Consistent risk estimation in moderately high-dimensional linear regression. *IEEE Transactions on Information Theory*, 67(9):5997–6030, 2021.

Mikhail Belkin, Daniel Hsu, and Ji Xu. Two models of double descent for weak features. *SIAM Journal on Mathematics of Data Science*, 2(4):1167–1180, 2020.

José Manuel Zorrilla Matilla, Manasi Sharma, Daniel Hsu, and Zoltán Haiman. Interpreting deep learning models for weak lensing. *Phys. Rev. D*, 102:123506, Dec 2020.

Arushi Gupta and Daniel Hsu. Parameter identification in markov chain choice models. *Theoretical Computer Science*, 808:99–107, 2020.

Avner May, Alireza Bagheri Garakani, Zhiyun Lu, Dong Guo, Kuan Liu, Aurélien Bellet, Linxi Fan, Michael Collins, Daniel Hsu, and Brian Kingsbury. Kernel approximation methods for speech recognition. *Journal of Machine Learning Research*, 20(59):1–36, 2019.

Dezső Ribli, Bálint Ármin Pataki, José Manuel Zorrilla Matilla, Daniel Hsu, Zoltán Haiman, and István Csabai. Weak lensing cosmology with convolutional neural networks on noisy data. *Monthly Notices of the Royal Astronomical Society*, 490(2):1843–1860, 2019.

Mikhail Belkin, Daniel Hsu, Siyuan Ma, and Soumik Mandal. Reconciling modern machine learning and the bias-variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32):15849–15854, 2019.

Daniel Hsu, Aryeh Kontorovich, David A. Levin, Yuval Peres, Csaba Szepesvári, and Geoffrey Wolfer. Mixing time estimation in reversible Markov chains from a single sample path. *Ann. Appl. Probab.*, 29(4):2439–2480, 2019.

Chia-Hao Liu, Yunzhe Tao, Daniel Hsu, Qiang Du, and Simon J. L. Billinge. Using a machine learning approach to determine the space group of a structure from the atomic pair distribution function. *Acta Crystallographica Section A*, 75(4):633–643, 2019.

Thomas Effland, Anna Lawson, Sharon Balter, Katelynn Devinney, Vasudha Reddy, HaeNa Waechter, Luis Gravano, and Daniel Hsu. Discovering foodborne illness in online restaurant reviews. *Journal of the American Medical Informatics Association*, 25(12):1586–1592, 2018.

Arushi Gupta, José Manuel Zorrilla Matilla, Daniel Hsu, and Zoltán Haiman. Non-gaussian information from weak lensing data via deep learning. *Phys. Rev. D*, 97:103515, May 2018.

Sasikiran Kandula, Daniel Hsu, and Jeffrey Shaman. Subregional nowcasts of seasonal influenza using search trends. *Journal of Medical Internet Research*, 19(11):e370, 2017.

Cun Mu, Daniel Hsu, and Donald Goldfarb. Greedy approaches to symmetric orthogonal tensor decomposition. *SIAM Journal on Matrix Analysis and Applications*, 38(4):1210–1226, 2017.

Lee H. Dicker, Dean P. Foster, and Daniel Hsu. Kernel ridge vs. principal component regression: Minimax bounds and the qualification of regularization operators. *Electronic Journal of Statistics*, 1(1):1022–1047, 2017.

José Manuel Zorrilla Matilla, Zoltán Haiman, Daniel Hsu, Arushi Gupta, and Andrea Petri. Do dark matter halos explain lensing peaks? *Phys. Rev. D*, 94:083506, Oct 2016.

Karl Stratos, Michael Collins, and Daniel Hsu. Unsupervised part-of-speech tagging with anchor hidden Markov models. *Transactions of the Association for Computational Linguistics*, 4:245–257, 2016.

Daniel Hsu and Sivan Sabato. Loss minimization and parameter estimation with heavy tails. *Journal of Machine Learning Research*, 17(18):1–40, 2016.

Anima Anandkumar, Daniel Hsu, Majid Janzamin, and Sham M. Kakade. When are overcomplete topic models identifiable? Uniqueness of tensor Tucker decompositions with structured sparsity. *Journal of Machine Learning Research*, 16(Dec):2643–2694, 2015.

Cun Mu, Daniel Hsu, and Donald Goldfarb. Successive rank-one approximations for nearly orthogonally decomposable symmetric tensors. *SIAM Journal on Matrix Analysis and Applications*, 36(4):1638–1659, 2015.

Sivan Sabato, Shai Shalev-Shwartz, Nathan Srebro, Daniel Hsu, and Tong Zhang. Learning sparse low-threshold linear classifiers. *Journal of Machine Learning Research*, 16(Jul):1275–1304, 2015.

Anima Anandkumar, Dean P. Foster, Daniel Hsu, Sham M. Kakade, and Yi-Kai Liu. A spectral algorithm for latent Dirichlet allocation. *Algorithmica*, 72(1):193–214, 2015.

Anima Anandkumar, Rong Ge, Daniel Hsu, Sham M. Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15(Aug):2773–2831, 2014.

Anima Anandkumar, Rong Ge, Daniel Hsu, and Sham M. Kakade. A tensor approach to learning mixed membership community models. *Journal of Machine Learning Research*, 15(Jun):2239–2312, 2014.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. Random design analysis of ridge regression. *Foundations of Computational Mathematics*, 14(3):569–600, 2014.

Alekh Agarwal, Dean P. Foster, Daniel Hsu, Sham M. Kakade, and Alexander Rakhlin. Stochastic convex optimization with bandit feedback. *SIAM Journal on Optimization*, 23(1):213–240, 2013.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17(52):1–6, 2012.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. A spectral algorithm for learning hidden Markov models. *Journal of Computer and System Sciences*, 78(5):1460–1480, 2012.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. Tail inequalities for sums of random matrices that depend on the intrinsic dimension. *Electronic Communications in Probability*, 17(14):1–13, 2012.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. Robust matrix decomposition with sparse corruptions. *IEEE Transactions on Information Theory*, 57(11):7221–7234, 2011.

Refereed Conference Proceedings

Sivaraman Balakrishnan, Nika Haghtalab, Daniel Hsu, Brian Lee, and Eric Zhao. Panprediction: optimal predictions for any downstream task and loss. In *Twenty-Ninth International Conference on Artificial Intelligence and Statistics*, 2026.

Avrim Blum, Daniel Hsu, Cyrus Rashtchian, and Donya Saless. Prior knowledge makes it possible: from sublinear graph algorithms to LLM test-time methods. In *Twenty-Ninth International Conference on Artificial Intelligence and Statistics*, 2026.

Saeyoung Rho, Cyrus Illick, Samhitha Narasipura, Alberto Abadie, Daniel Hsu, and Vishal Misra. Time-aware synthetic control. In *Twenty-Ninth International Conference on Artificial Intelligence and Statistics*, 2026.

Navid Ardeshir, Samuel Deng, Daniel Hsu, and Jingwen Liu. Group-realizable multi-group learning by minimizing empirical risk. In *Thirty-Seventh International Conference on Algorithmic Learning Theory*, 2026.

Jingwen Liu, Hantao Yu, Clayton Sanford, Alexandr Andoni, and Daniel Hsu. Fast attention mechanisms: a tale of parallelism. In *Advances in Neural Information Processing Systems 38*, 2025.

Zixuan Wang, Eshaan Nichani, Alberto Bietti, Alex Damian, Daniel Hsu, Jason D. Lee, and Denny Wu. Learning compositional functions with transformers from easy-to-hard data. In *Thirty-Eighth Annual Conference on Learning Theory*, 2025.

Berfin Şimşek, Amire Bendjeddou, and Daniel Hsu. Learning gaussian multi-index models with gradient flow: Time complexity and directional convergence. In *Twenty-Eighth International Conference on Artificial Intelligence and Statistics*, 2025.

Samuel Deng, Daniel Hsu, and Jingwen Liu. Group-wise oracle-efficient algorithms for online multi-group learning. In *Advances in Neural Information Processing Systems 37*, 2024.

Zixuan Wang, Stanley Wei, Daniel Hsu, and Jason D. Lee. Transformers provably learn sparse token selection while fully-connected nets cannot. In *Forty-First International Conference on Machine Learning*, 2024.

Clayton Sanford, Daniel Hsu, and Matus Telgarsky. Transformers, parallel computation, and logarithmic depth. In *Forty-First International Conference on Machine Learning*, 2024.

Samuel Deng and Daniel Hsu. Multi-group learning for hierarchical groups. In *Forty-First International Conference on Machine Learning*, 2024.

Daniel Hsu and Arya Mazumdar. On the sample complexity of parameter estimation in logistic regression with normal design. In *Thirty-Seventh Annual Conference on Learning Theory*, 2024.

Daniel Hsu, Jizhou Huang, and Brendan Juba. Distribution-specific auditing for subgroup fairness. In *Fifth Symposium on Foundations of Responsible Computing*, 2024.

Clayton Sanford, Daniel Hsu, and Matus Telgarsky. Representational strengths and limitations of transformers. In *Advances in Neural Information Processing Systems 36*, 2023.

Navid Ardeshir, Daniel Hsu, and Clayton Sanford. Intrinsic dimensionality and generalization properties of the $\mathcal{R}$-norm inductive bias. In *Thirty-Sixth Annual Conference on Learning Theory*, 2023.

Bingbin Liu, Daniel Hsu, Pradeep Ravikumar, and Andrej Risteski. Masked prediction: a parameter identifiability view. In *Advances in Neural Information Processing Systems 35*, 2022.

Christopher Tosh and Daniel Hsu. Simple and near-optimal algorithms for hidden stratification and multi-group learning. In *Thirty-Ninth International Conference on Machine Learning*, 2022.

Daniel Hsu, Clayton Sanford, Rocco A. Servedio, and Emmanouil-Vasileios Vlatakis-Gkaragkounis. Near-optimal statistical query lower bounds for agnostically learning intersections of halfspaces with Gaussian marginals. In *Thirty-Fifth Annual Conference on Learning Theory*, 2022.

Samuel Deng, Yilin Guo, Daniel Hsu, and Debmalya Mandal. Learning tensor representations for meta-learning. In *Twenty-Fifth International Conference on Artificial Intelligence and Statistics*, 2022.

Max Simchowitz, Christopher Tosh, Akshay Krishnamurthy, Daniel Hsu, Thodoris Lykouris, Miroslav Dudík, and Robert E. Schapire. Bayesian decision-making under misspecified priors with applications to meta-learning. In *Advances in Neural Information Processing Systems 34*, 2021.

Navid Ardeshir, Clayton Sanford, and Daniel Hsu. Support vector machines and linear regression coincide with very high-dimensional features. In *Advances in Neural Information Processing Systems 34*, 2021.

Daniel Hsu, Clayton Sanford, Rocco A. Servedio, and Emmanouil-Vasileios Vlatakis-Gkaragkounis. On the approximation power of two-layer networks of random ReLUs. In *Thirty-Fourth Annual Conference on Learning Theory*, 2021.

Daniel Hsu, Vidya Muthukumar, and Ji Xu. On the proliferation of support vectors in high dimensions. In *Twenty-Fourth International Conference on Artificial Intelligence and Statistics*, 2021.

Daniel Hsu, Ziwei Ji, Matus Telgarsky, and Lan Wang. Generalization bounds via distillation. In *Ninth International Conference on Learning Representations*, 2021.

Christopher Tosh, Akshay Krishnamurthy, and Daniel Hsu. Contrastive learning, multi-view redundancy, and linear models. In *Thirty-Second International Conference on Algorithmic Learning Theory*, 2021.

Giannis Karamanolakis, Daniel Hsu, and Luis Gravano. Cross-lingual text classification with minimal resources by transferring a sparse teacher. In *Conference on Empirical Methods in Natural Language Processing: Findings*, 2020.

Debmalya Mandal, Samuel Deng, Suman Jana, Jeannette M. Wing, and Daniel Hsu. Ensuring fairness beyond the training data. In *Advances in Neural Information Processing Systems 33*, 2020.

Bo Cowgill, Fabrizio Dell'Acqua, Samuel Deng, Daniel Hsu, Nakul Verma, and Augustin Chaintreau. Biased programmers? Or biased data? A field experiment in operationalizing AI ethics. In *Proceedings of the Twenty-First ACM Conference on Economics and Computation*, 2020.

Christopher Tosh and Daniel Hsu. Diameter-based interactive structure discovery. In *Twenty-Third International Conference on Artificial Intelligence and Statistics*, 2020.

Ji Xu and Daniel Hsu. On the number of variables to use in principal component regression. In *Advances in Neural Information Processing Systems 32*, 2019.

Mathias Lecuyer, Riley Spahn, Kiran Vodrahalli, Roxana Geambasu, and Daniel Hsu. Privacy accounting and quality control in the sage differentially private ml platform. In *Twenty-Seventh ACM Symposium on Operating Systems Principles*, 2019.

Giannis Karamanolakis, Daniel Hsu, and Luis Gravano. Leveraging just a few keywords for fine-grained aspect detection through weakly supervised co-training. In *Conference on Empirical Methods in Natural Language Processing*, 2019.

Yucheng Chen, Matus Telgarsky, Chao Zhang, Bolton Bailey, Daniel Hsu, and Jian Peng. A gradual, semi-discrete approach to generative network training via explicit wasserstein minimization. In *Thirty-Sixth International Conference on Machine Learning*, 2019.

Sanjoy Dasgupta, Daniel Hsu, Stefanos Poulis, and Xiaojin Zhu. Teaching a black-box learner. In *Thirty-Sixth International Conference on Machine Learning*, 2019.

Michał Dereziński, Manfred K. Warmuth, and Daniel Hsu. Correcting the bias in least squares regression with volume-rescaled sampling. In *Twenty-Second International Conference on Artificial Intelligence and Statistics*, 2019.

Alexandr Andoni, Rishabh Dudeja, Daniel Hsu, and Kiran Vodrahalli. Attribute-efficient learning of monomials over highly-correlated variables. In *Thirtieth International Conference on Algorithmic Learning Theory*, 2019.

Mathias Lecuyer, Vaggelis Atlidakis, Roxana Geambasu, Daniel Hsu, and Suman Jana. Certified robustness to adversarial examples with differential privacy. In *IEEE Symposium on Security and Privacy*, 2019.

Mikhail Belkin, Daniel Hsu, and Partha Mitra. Overfitting or perfect fitting? risk bounds for classification and regression rules that interpolate. In *Advances in Neural Information Processing Systems 31*, 2018.

Michał Dereziński, Manfred K. Warmuth, and Daniel Hsu. Leveraged volume sampling for linear regression. In *Advances in Neural Information Processing Systems 31*, 2018.

Ji Xu, Daniel Hsu, and Arian Maleki. Benefits of over-parameterization with EM. In *Advances in Neural Information Processing Systems 31*, 2018.

Rishabh Dudeja and Daniel Hsu. Learning single-index models in Gaussian space. In *Thirty-First Annual Conference on Learning Theory*, 2018.

Daniel Hsu, Kevin Shi, and Xiaorui Sun. Linear regression without correspondence. In *Advances in Neural Information Processing Systems 30*, 2017.

Arushi Gupta and Daniel Hsu. Parameter identification in markov chain choice models. In *Twenty-Eighth International Conference on Algorithmic Learning Theory*, 2017.

Alexandr Andoni, Daniel Hsu, Kevin Shi, and Xiaorui Sun. Correspondence retrieval. In *Thirtieth Annual Conference on Learning Theory*, 2017.

Florian Tramér, Vaggelis Atlidakis, Roxana Geambasu, Daniel Hsu, Jean-Pierre Hubaux, Mathias Humbert, Ari Juels, and Huang Lin. Fairtest: Discovering unwarranted associations in data-driven applications. In *Second IEEE European Symposium on Security and Privacy*, 2017.

Alina Beygelzimer, Daniel Hsu, John Langford, and Chicheng Zhang. Search improves label for active learning. In *Advances in Neural Information Processing Systems 29*, 2016.

Ji Xu, Daniel Hsu, and Arian Maleki. Global analysis of Expectation Maximization for mixtures of two Gaussians. In *Advances in Neural Information Processing Systems 29*, 2016.

Avner May, Michael Collins, Daniel Hsu, and Brian Kingsbury. Compact kernel models for acoustic modeling via random feature selection. In *Forty-First IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016.

Daniel Hsu, Aryeh Kontorovich, and Csaba Szepesvári. Mixing time estimation in reversible Markov chains from a single sample path. In *Advances in Neural Information Processing Systems 28*, 2015.

Tzu-Kuo Huang, Alekh Agarwal, Daniel Hsu, John Langford, and Robert E. Schapire. Efficient and parsimonious agnostic active learning. In *Advances in Neural Information Processing Systems 28*, 2015.

Mathias Lecuyer, Riley Spahn, Yannis Spiliopoulos, Augustin Chaintreau, Roxana Geambasu, and Daniel Hsu. Sunlight: fine-grained targeting detection at scale with statistical confidence. In *Twenty-Second ACM Conference on Computer and Communications Security*, 2015.

Karl Stratos, Michael Collins, and Daniel Hsu. Model-based word embeddings from decompositions of count matrices. In *Fifty-Third Annual Meeting of the Association for Computational Linguistics*, 2015.

Alekh Agarwal, Alina Beygelzimer, Daniel Hsu, John Langford, and Matus Telgarsky. Scalable nonlinear learning with adaptive polynomial expansions. In *Advances in Neural Information Processing Systems 27*, 2014.

Kamalika Chaudhuri, Daniel Hsu, and Shuang Song. The large margin mechanism for differentially private maximization. In *Advances in Neural Information Processing Systems 27*, 2014.

Karl Stratos, Do-kyum Kim, Michael Collins, and Daniel Hsu. A spectral algorithm for learning class-based n-gram models of natural language. In *Thirtieth Conference on Uncertainty in Artificial Intelligence*, 2014.

Alekh Agarwal, Daniel Hsu, Satyen Kale, John Langford, Lihong Li, and Robert E. Schapire. Taming the monster: a fast and simple algorithm for contextual bandits. In *Thirty-First International Conference on Machine Learning*, 2014.

Daniel Hsu and Sivan Sabato. Heavy-tailed regression with a generalized median-of-means. In *Thirty-First International Conference on Machine Learning*, 2014.

James Zou, Daniel Hsu, David Parkes, and Ryan P. Adams. Contrastive learning using spectral methods. In *Advances in Neural Information Processing Systems 26*, 2013.

Anima Anandkumar, Daniel Hsu, Majid Janzamin, and Sham M. Kakade. When are overcomplete topic models identifiable? Uniqueness of tensor Tucker decompositions with structured sparsity. In *Advances in Neural Information Processing Systems 26*, 2013.

Anima Anandkumar, Rong Ge, Daniel Hsu, and Sham M. Kakade. A tensor spectral approach to learning mixed membership community models. In *Twenty-Sixth Annual Conference on Learning Theory*, 2013.

Anima Anandkumar, Daniel Hsu, Adel Javanmard, and Sham M. Kakade. Learning linear Bayesian networks with latent variables. In *Thirtieth International Conference on Machine Learning*, 2013.

Daniel Hsu and Sham M. Kakade. Learning mixtures of spherical Gaussians: moment methods and spectral decompositions. In *Fourth Innovations in Theoretical Computer Science*, 2013.

Anima Anandkumar, Dean P. Foster, Daniel Hsu, Sham M. Kakade, and Yi-Kai Liu. A spectral algorithm for latent Dirichlet allocation. In *Advances in Neural Information Processing Systems 25*, 2012.

Anima Anandkumar, Daniel Hsu, Furong Huang, and Sham M. Kakade. Learning mixtures of tree graphical models. In *Advances in Neural Information Processing Systems 25*, 2012.

Daniel Hsu, Sham M. Kakade, and Percy Liang. Identifiability and unmixing of latent parse trees. In *Advances in Neural Information Processing Systems 25*, 2012.

Kamalika Chaudhuri and Daniel Hsu. Convergence rates for differentially private statistical estimation. In *Twenty-Ninth International Conference on Machine Learning*, 2012.

Anima Anandkumar, Daniel Hsu, and Sham M. Kakade. A method of moments for mixture models and hidden Markov models. In *Twenty-Fifth Annual Conference on Learning Theory*, 2012.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. Random design analysis of ridge regression. In *Twenty-Fifth Annual Conference on Learning Theory*, 2012.

Anima Anandkumar, Kamalika Chaudhuri, Daniel Hsu, Sham M. Kakade, Le Song, and Tong Zhang. Spectral methods for learning multivariate latent tree structure. In *Advances in Neural Information Processing Systems 24*, 2011.

Alekh Agarwal, Dean P. Foster, Daniel Hsu, Sham M. Kakade, and Alexander Rakhlin. Stochastic convex optimization with bandit feedback. In *Advances in Neural Information Processing Systems 24*, 2011.

Miroslav Dudík, Daniel Hsu, Satyen Kale, Nikos Karampatziakis, John Langford, Lev Reyzin, and Tong Zhang. Efficient optimal learning for contextual bandits. In *Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, 2011.

Kamalika Chaudhuri and Daniel Hsu. Sample complexity bounds for differentially private learning. In *Twenty-Fourth Annual Conference on Learning Theory*, 2011.

Alina Beygelzimer, Daniel Hsu, John Langford, and Tong Zhang. Agnostic active learning without constraints. In *Advances in Neural Information Processing Systems 23*, 2010.

Kamalika Chaudhuri, Yoav Freund, and Daniel Hsu. An online learning-based framework for tracking. In *Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, 2010.

Daniel Hsu, Sham M. Kakade, John Langford, and Tong Zhang. Multi-label prediction via compressed sensing. In *Advances in Neural Information Processing Systems 22*, 2009.

Kamalika Chaudhuri, Yoav Freund, and Daniel Hsu. A parameter-free hedging algorithm. In *Advances in Neural Information Processing Systems 22*, 2009.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. A spectral algorithm for learning hidden Markov models. In *Twenty-Second Annual Conference on Learning Theory*, 2009.

Sanjoy Dasgupta and Daniel Hsu. Hierarchical sampling for active learning. In *Twenty-Fifth International Conference on Machine Learning*, 2008.

Sanjoy Dasgupta, Daniel Hsu, and Claire Monteleoni. A general agnostic active learning algorithm. In *Advances in Neural Information Processing Systems 20*, 2007.

Sanjoy Dasgupta and Daniel Hsu. On-line estimation with the multivariate Gaussian distribution. In *Twentieth Annual Conference on Learning Theory*, 2007.

Sanjoy Dasgupta, Daniel Hsu, and Nakul Verma. A concentration theorem for projections. In *Twenty-Second Conference on Uncertainty in Artificial Intelligence*, 2006.

Other Articles

Christopher Tosh, Philip Greengard, Ben Goodrich, Andrew Gelman, Aki Vehtari, and Daniel Hsu. The piranha problem: large effects swimming in a small pond. *Notices Amer. Math. Soc.*, 72(1):15–25, 2025.

Daniel Hsu, Nikos Karampatziakis, John Langford, and Alex J. Smola. Parallel online learning. In Ron Bekkerman, Misha Bilenko, and John Langford, editors, *Scaling Up Machine Learning: Parallel and Distributed Approaches*, chapter 14, pages 283–306. Cambridge University Press, 2012.

Advising

Postdoctoral scholars

Daniel Alabi (2022–2025)

Debmalya Mandal (2019–2021)

Christopher Tosh (2018–2021)

Ph.D. students

Navid Ardeshir (Ph.D., 2025)

Noah Bergam (in progress)

Samuel Deng (Ph.D., 2026)

Rishabh Dudeja (Ph.D., 2021)

Giannis Karamanolakis (Ph.D., 2022)
Jingwen Liu (in progress)
Clayton Sanford (Ph.D., 2024)
Kevin Shi (Ph.D., 2020)
Kiran Vodrahalli (Ph.D., 2022)
Ji Xu (Ph.D., 2020)
Keyang Xu

M.S. students

Samuel Deng (M.S., 2021)
Arushi Gupta (M.S., 2018)
Zizhou Liu (M.S., 2022)
Geelon So (M.S., 2019)
Mingyue Xu (M.S., 2022)

Undergraduate students

Alejandro Buendia (B.A., 2019)
Arushi Gupta (B.S., 2016)
Ziyi Liu (B.S., 2021)
Zizhou Liu (B.S., 2021)
Berkan Ottlik (B.S., 2024)
Edward Ri (B.S., 2024)
Manasi Sharma (B.A., 2021)

High school students

Jagdeep Bhatia (High school diploma, 2020)
Alan Chung (High school diploma, 2017)
Bryan Chung (High school diploma, 2024)
Jacob Fisher (High school diploma, 2017)

Funding

NSF award DMS-2502259 (10/2025–09/2028, $475,000), *Collaborative Research: MFAI: Emergence of features in modern machine learning* (PI), 2025.

ONR award N00014-24-1-2700 (09/2024–08/2027, $534,617), *Multi-group machine learning: theory and algorithms* (PI), 2024.

NSF award CCF-1740833 (09/2017–08/2020, $1,500,000), *TRIPODS: From Foundations to Practice of Data Science and Back* (co-PI with Wright, Du, Blei, and Andoni), 2017.

NSF award IIS-1563785 (09/2016–08/2020, $1,196,617), *III: Medium: Adaptive Information Extraction from Social Media for Actionable Inferences in Public Health* (co-PI with Gravano), 2016.

NSF award DMREF-1534910 (10/2015–9/2018, $982,786), *DMREF: Deblurring our View of Atomic Arrangements in Complex Materials for Advanced Technologies* (co-PI with Billinge and Du), 2015.

Institute for Data Sciences and Engineering Research Opportunities and Approaches to Data Science Grant, *Learning the Nature of Dark Energy from Weak Gravitational Lensing with Computer Vision and Machine Learning* (co-PI with Haiman, Peek, and Nayar), 2014.

Awards and Honors

Google.org Scientific Progress Award, 2025.

Best Student Paper Award, Symposium on Foundations of Responsible Computing, 2024.

National Academy of Sciences Kavli Fellow, 2017.

Alfred P. Sloan Research Fellow in Computer Science, 2016.

IEEE Intelligent Systems "AI's 10 to Watch", 2016.

Yahoo Academic Career Enhancement Award, 2014.

U.C. San Diego Departmental Dissertation Award, 2010.

Honorable Mention for Outstanding Student Paper Award at NeurIPS, 2009.

Invited Talks

Simons Institute for the Theory of Computing, Berkeley, CA, June 2026.

Department of Statistics and Data Science Seminar, Yale University, New Haven, CT, March 2026.

Foundations of Optimization, Learning, and Data Science Seminar, University of Pennsylvania, Philadelphia, February 2026.

Simons Institute for the Theory of Computing, Berkeley, CA, January 2026.

Columbia University Statistics Student Seminar, New York, NY, October 2025.

Institute for the Foundations of Data Science Workshop, University of Washington, Seattle, August 2025.

Mathematics and Data Seminar, New York University, New York, April 2025.

Workshop on Theoretical Perspectives on Large Language Models, University of California, San Diego, March 2025.

Econometrics and Statistics Colloquium, University of Chicago Booth School of Business, Chicago, IL, October 2024.

Computer Science Seminar, University of Rochester, Rochester, NY, October 2024.

Workshop on Transformers as a Computational Model, Simons Institute for the Theory of Computing, University of California, Berkeley, September 2024.

Special Year on Large Language Models and Transformers Boot Camp, Simons Institute for the Theory of Computing, University of California, Berkeley, September 2024.

Seminars on Formal Languages and Neural Networks (virtual), May 2024.

ICERM Workshop on Connecting Higher-Order Statistics and Symmetric Tensors, Brown University, Providence, RI, January 2024.

IFML Workshop on Generative AI, The University of Texas at Austin, Austin, TX, December 2023.

Department of Statistics and Data Science Seminar, Yale University, New Haven, CT, November 2023.

Oberwolfach Mini-Workshop on Interpolation and Over-parameterization in Statistics and Machine Learning, Oberwolfach, Germany, September 2023.

Halıcıoğlu Data Science Institute, University of California San Diego, La Jolla, CA, May 2023.

Workshop on Multigroup Fairness and the Validity of Statistical Judgment, Simons Institute for the Theory of Computing, Berkeley, CA, April 2023.

Theory Seminar, University of California San Diego, La Jolla, CA, April 2023.

Data Sciences and Operations Statistics Seminar, University of Southern California, Los Angeles, CA, March 2023.

Information Theory and Applications Workshop, San Diego, CA, February 2023.

Harvard Probabilitas Seminar, Harvard University, December 2022.

Oberwolfach Workshop on Mathematical Foundations of Robust and Generalizable Learning, Oberwolfach, Germany, October 2022.

Workshop on the Theory of Overparameterized Machine Learning, April 2022 (virtual).

2nd Workshop on Seeking Low-dimensionality in Deep Neural Networks, November 2021 (virtual).

Joint Statistics Seminar of Center for Research in Economics and Statistics (École Nationale de la Statistique et de l'Administration Économique) and Centre de Mathématiques Appliquées (École Polytechnique), Paris, France, November 2021 (virtual).

Operations Research and Financial Engineering Department Colloquium, Princeton University, NJ, October 2021.

Yale Institute for Network Science, Yale University, New Haven, CT, October 2021 (virtual).

Workshop on Mathematical Foundations and Applications of Deep Learning; Institute for Mathematics and its Applications, Purdue University, IN, August 2021 (virtual).

Workshop on Theory of Deep Learning; Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, August 2021 (virtual).

Mixtures, Hidden Markov Models, and Clustering (MHC 2021); Institut de Mathématique d'Orsay, Orsay, France, June 2021 (virtual).

Workshop on Efficient Tensor Representations for Learning and Computational Complexity, Institute for Pure & Applied Mathematics, Los Angeles, CA, May 2021 (virtual).

Workshop on the Multifaceted Complexity of Machine Learning, Institute for Mathematical and Statistical Innovation, Chicago, IL, April 2021 (virtual).

Oberwolfach Workshop on Mathematical Foundations of Machine Learning, Oberwolfach, Germany, March 2021 (virtual).

Flatiron Computational Methods and Data Science Journal Club, Flatiron Institute, New York, NY, January 2021 (virtual).

The Institute for Data, Econometrics, Algorithms, and Learning Seminar, Northwestern University, Chicago, IL, October 2020 (virtual).

Mathematical Institute for Data Science Seminar, Johns Hopkins University, Baltimore, MD, October 2020 (virtual).

Yahoo Research, Sunnyvale, CA, August 2020 (virtual).

Joint Statistical Meeting (discussant, session on "Theoretical Advances in Deep Learning"), Seattle, WA, August 2020 (virtual).

Machine Learning Seminar, Carnegie Mellon University, Pittsburgh, PA, July 2020 (virtual).

Department of Computer & Information Science, University of Pennsylvania, Philadelphia, PA, February 2020.

Two Sigma Research Seminar, January 2020.

Columbia DSI/TRIPODS Deep Learning Workshop, Columbia University, New York, NY, November 2019.

Department of Statistics, University of Chicago, Chicago, IL, November 2019.

University of Maryland, College Park, MD, October 2019.

Microsoft Research AI Institute, Redmond, WA, August 2019.

Simons Institute for the Theory of Computing, Berkeley, CA, June 2019.

Simons Symposium on New Directions in Theoretical Machine Learning, Elmau, Germany, May 2019.

Google Research Machine Learning Seminar, February 2019.

Information Theory and Applications Workshop, San Diego, CA, February 2019.

Department of Computational Mathematics, Science and Engineering Colloquium, Michigan State University, November 2018.

TRIPODS/DIMACS Workshop on Optimization in Machine Learning, Lehigh University, August 2018.

Simons Institute for the Theory of Computing, University of California, Berkeley, June 2018.

Workshop on Modern Challenges of Learning Theory, Centre de Recherches Mathematiques, University of Montreal, April 2018.

Joint Statistics/CSE Seminar, Ohio State University, November 2017.

Machine Learning Seminar, New York University, October 2017.

1st Japanese-American-German Kavli Frontiers of Science Symposium, National Academy of Sciences, Bad Neuenahr, Germany, September 2017.

Theory Lunch Seminar, Princeton University, May 2017.

Stochastics and Statistics Seminar, Massachusetts Institute of Technology, April 2017.

Foundations of Machine Learning Boot Camp, Simons Institute for the Theory of Computing, University of California, Berkeley, January 2017.

Machine Learning Seminar, University of Illinois, Urbana-Champaign, October 2016.

The Gatsby Computational Neuroscience Unit, University College London, July 2016.

Janelia Research Campus, Ashburn, VA, June 2016.

DIMACS Theoretical Computer Science Seminar, Rutgers University, Piscataway, NJ, April 2016.

Wilks Statistics Seminar, Princeton University, Princeton, NJ, April 2016.

Bloomberg Labs, New York, NY, March 2016.

Conference on Information Sciences and Systems, Princeton, NJ, March 2016.

Information Theory and Applications Workshop, La Jolla, CA, February 2016.

PrivacyCon (Federal Trade Commission), Washington, DC, January 2016.

New York University Machine Learning Seminar, November 2015.

Program in Algorithmic and Combinatorial Thinking, Princeton University, June 2015.

D.E. Shaw, New York, NY, April 2015.

Two Sigma research seminar, January 2015.

Columbia University Statistics Student Seminar, New York, NY, November 2014.

Wisconsin Institute of Discovery, University of Wisconsin–Madison, October 2014.

Google Research Machine Learning Seminar, October 2014.

Wireless Networking & Communications Group Seminar, University of Texas at Austin, September 2014.

Yahoo! Labs, New York, September 2014.

AAAI Conference on Artificial Intelligence (Tutorial) 2014, Québec City, Québec, July 2014.

Workshop on Mathematical Foundations of Learning Theory, Barcelona, June 2014.

Workshop on Overcoming the Intractability Bottleneck in Unsupervised Machine Learning (co-located with 46th Annual Symposium on the Theory of Computing), New York, NY, May 2014.

New York University Machine Learning Seminar, April 2014.

Information Theory and Applications Workshop, San Diego, CA, February 2014.

Columbia University Statistics Student Seminar, New York, NY, September 2013.

International Conference on Machine Learning (Tutorial) 2013, Atlanta, GA, June 2013.

Second Joint Symposium on Natural Language Processing and Machine Learning, Atlanta, GA, June 2013.

Department of Computer Science, University of California, Los Angeles, April 2013.

Department of Computer Science, Cornell University, April 2013.

Department of Computer Science, Columbia University, April 2013.

Department of Computer Science, Yale University, April 2013.

Department of Computer Science, University of Toronto, April 2013.

School of Computer Science, University of Waterloo, April 2013.

Department of Computer Science, Princeton University, March 2013.

Department of Computer and Information Science, University of Pennsylvania, March 2013.

Department of Computer Science, University of Maryland, College Park, March 2013.

Department of Electrical and Computer Engineering, Boston University, March 2013.

Department of Computer Science and Engineering, Washington University in St. Louis, March 2013.

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, March 2013.

Microsoft Research, Mountain View, CA, March 2013.

Department of Computer Science, University of Illinois at Urbana-Champaign, March 2013.

Department of Computer Science, University of Chicago, February 2013.

Machine Learning Department, Carnegie Mellon University, February 2013.

Department of Statistics, Rutgers University, February 2013.

Department of Computer Science, Johns Hopkins University, February 2013.

Information Theory and Applications Workshop, San Diego, CA, February 2013.

The Gatsby Computational Neuroscience Unit, University College London, February 2013.

School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, February 2013.

NeurIPS 2012 Workshop on Modern Nonparametric Methods in Machine Learning, Tahoe, NV, December 2012.

NICTA Machine Learning Seminar, Canberra, Australia, April 2012.

Department of Statistics, Rutgers University, Piscataway, NJ, November 2011.

Microsoft Research, Redmond, WA, October 2011.

Department of Computer Science and Technology, Nanjing University, Nanjing, China, August 2011.

School of Computer Science, Fudan University, Shanghai, China, August 2011.

ID Analytics, San Diego, CA, May 2010.

IBM Research, Yorktown Heights, NY, October 2009.

Service

Journal Associate/Action Editor

ACM Transactions on Algorithms, January 2017 to present.

SIAM Journal on Mathematics of Data Science, January 2021 to present.

Journal of Machine Learning Research, January 2022 to August 2024.

Foundations and Trends in Machine Learning, April 2024 to present.

Conference Program Chair

Conference on Learning Theory (COLT), 2019 (with Alina Beygelzimer).

International Conference on Algorithmic Learning Theory (ALT), 2024 (with Claire Vernade).

International Conference on Machine Learning (ICML), 2025 (with Virginia Smith, Simon Lacoste-Julien, Maryam Fazel).

Conference Senior Area Chair

Conference on Neural Information Processing Systems (NeurIPS), 2019, 2021, 2023.

Conference Senior Program Committee/Area Chair

Conference on Learning Theory (COLT), 2020, 2021, 2022, 2023, 2024, 2026.

International Conference on Machine Learning (ICML), 2012, 2013, 2015, 2016, 2017.

Conference on Neural Information Processing Systems (NeurIPS), 2012, 2013, 2017.

International Conference on Artificial Intelligence and Statistics (AISTATS), 2016, 2017, 2019.

International Conference on Algorithmic Learning Theory (ALT), 2022, 2023.

ACM-IMS Foundations of Data Science Conference, 2020.

Conference Program Committee/Reviewer

AAAI Conference on Artificial Intelligence (AAAI), 2014.

Conference on Learning Theory (COLT), 2011, 2013, 2015, 2016, 2017, 2018.

International Conference on Algorithmic Learning Theory (ALT), 2017, 2018, 2019, 2021.

International Conference on Artificial Intelligence and Statistics (AISTATS), 2010, 2011.

International Conference on Machine Learning (ICML), 2010, 2011.

Conference on Neural Information Processing Systems (NeurIPS), 2008, 2011, 2014, 2015, 2020, 2026.

Conference on Uncertainty in Artificial Intelligence (UAI), 2010.

IEEE Symposium on Foundations of Computer Science (FOCS), 2021.

Conference External Reviewing

Conference on Algorithmic Learning Theory (ALT)

Conference on Learning Theory (COLT)

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

IEEE Symposium on Foundations of Computer Science (FOCS)

ACM-SIAM Symposium on Discrete Algorithms (SODA)

ACM Symposium on Theory of Computing (STOC)

Journal Reviewing

Algorithmica

Annals of Statistics

Artificial Intelligence Journal

Data Mining and Knowledge Discovery

Electronic Journal of Statistics

IEEE Transactions on Information Theory

IEEE Transactions on Pattern Analysis and Machine Intelligence

Journal of the Association for Computing Machinery

Journal of Machine Learning Research

Journal of the Royal Statistical Society: Series B

Machine Learning Journal

Mathematical Statistics and Learning

Neural Computation

Transactions of the Association for Computational Linguistics

Funding Proposal Reviewing

National Science Foundation, 2014, 2018, 2019, 2022.

Israel Science Foundation (external), 2026.

Swiss National Science Foundation (external), 2013.

German-Israeli Foundation for Scientific Research and Development (external), 2010.

Computing Innovation Fellows, 2020.

Seminar Organization

Machine Learning and AI Seminar Series, Columbia University, Fall 2025, Spring 2026.

Columbia Year of Statistical Machine Learning, Columbia University, Fall 2019, Spring 2020.

Computer Science Distinguished Lecture Series, Columbia University, Fall 2018, 2019, 2020.

Foundations of Data Science Seminar Series, Columbia University, Fall 2015.

Workshop Organization

Frontiers in Data Science and AI: (Fun)damental Uses of Math in AI, 2026.

Columbia Foundations of Data Science Center Tutorials and Workshop, 2025.

Columbia Foundations of Data Science Center Workshop, 2024.

Columbia Statistical Machine Learning Symposium, 2023.

Workshop on Machine Learning, co-located with IEEE Symposium on Foundations of Computer Science (FOCS), 2021.

Workshop on Method of Moments and Spectral Learning, co-located with International Conference on Machine Learning (ICML), 2014.

DIMACS/CCICADA Workshop on Systems and Analytics of Big Data, 2014.

Workshop on Spectral Learning, co-located with International Conference on Machine Learning (ICML), 2013.

Workshop on Spectral Learning, co-located with Conference on Neural Information Processing Systems (NeurIPS), 2013.

Conference Organization

Publicity chair, International Conference on Artificial Intelligence and Statistics (AISTATS), 2019.

Local arrangements chair (with Satyen Kale), Conference on Learning Theory (COLT), 2016.

Virtual arrangements chair (with Peter Grünwald, Benjamin Guedj, Satyen Kale, and Wouter Koolen), Conference on Learning Theory (COLT), 2020.

Publications chair (with Yingyu Liang), International Conference on Machine Learning (ICML), 2021.

Program Organization

Modern Paradigms in Generalization, Simons Institute for the Theory of Computing, University of California, Berkeley, August–December, 2024.

Board of Directors

Association for Computational Learning (ACL), 2017–2020.

Information Officer

Association for Computational Learning (ACL), 2019–present.

Board of Advisors

TheoretiCS, 2021–2024.

Columbia-Dream Sports AI Innovation Center, 2024–present.