

# Shashank Singh

## Curriculum Vitae

Updated September 4, 2020

### CONTACT INFORMATION

Google  
6425 Penn Ave.  
Pittsburgh, PA 15213  
Website: <https://sss1.github.io>

Mobile Phone: +1-615-364-7734  
Email: [shashanksi@google.com](mailto:shashanksi@google.com)  
ORCID ID: [0000-0002-7305-673X](https://orcid.org/0000-0002-7305-673X)  
[Google Scholar Profile](#)

### EDUCATION

Ph.D.	Statistics & Machine Learning, Carnegie Mellon University	Aug 2019
	Thesis: <a href="#">Estimating Probability Distributions and Their Properties</a>	
M.S.	Machine Learning, Carnegie Mellon University	Aug 2017
M.S.	Mathematical Sciences, Carnegie Mellon University	May 2014
	Thesis: <a href="#">Concentration Inequalities for Density Functionals</a>	
B.S.	Mathematical Sciences, Carnegie Mellon University	May 2014
B.S.	Computer Science, Carnegie Mellon University	May 2014

### RESEARCH INTERESTS

**Theory:** Statistical Machine Learning (especially nonparametric methods, fundamental limits of learning from data, generative modeling, and causal inference), Information Theory

**Applications:** Computational Biology (genomics, biological systems modeling), Cognitive Science (behavior modeling, cognitive development), Epidemiological Forecasting

### EMPLOYMENT

2019 Sep- Present	Software Engineer at Google, Cloud AI Deployments Team (Pittsburgh, PA) <i>Designed, implemented, and evaluated machine learning systems for industrial and public sector applications (see, e.g., paper "Interpretable COVID-19 Forecasting").</i>
2018 Sep-Dec	Applied Scientist Intern at Amazon Web Services (New York, NY) <i>Developed new algorithms to compress deep neural networks used in computer vision (see paper "DARC: Differentiable ARchitecture Compression").</i>
2015 May-Aug	Software Engineering Intern at Google (Mountain View, CA) <i>Designed, implemented, and tested anomaly detection pipeline for detecting fraudulent activity amongst billions of daily user engagements with online products.</i>

## FELLOWSHIPS AND AWARDS

2019	NeurIPS 2019 Honorable Mention for Outstanding Paper (3/6743 submissions)
2017	Mellon Foundation Presidential Fellow in Life Sciences (1 year graduate funding)
2015	National Science Foundation Graduate Research Fellow (3 years graduate funding)
2012	Undergraduate Research Fellow in Computational Neuroscience, Center for the Neural Basis of Cognition, Carnegie Mellon University. (1 year research stipend)

## PUBLICATIONS

### Submitted Papers (Under Review)

Uppal, A., **Singh, S.** & Póczos, B. [Robust Density Estimation under Besov IPM Losses](#).

Arik, S. O., Li, C.-L., Yoon, J., Sinha, R., Epshteyn, A., Le, L., Menon, V., **Singh, S.**, Zhang, L., Yoder, N., Nikoltchev, M., Sonthalia, Y., Nakhost, H., Kanal, E., & Pfister, T. [Interpretable Covid-19 Forecasting](#).

Khim, J., Xu, Z., & **Singh, S.** [Multiclass Classification via Class-Weighted Nearest Neighbors](#).

**Singh, S.** & Póczos, B. [Minimax Distribution Estimation in Wasserstein Distance](#).

### Journal Articles

2020 Kim, J., **Singh, S.**, Thiessen, E. D., & Fisher, A. V. "A Hidden Markov Model for Analyzing Eye-Tracking of Moving Objects". In *Behavior Research Methods*.

2019 Rashid, S., Long, Z., **Singh, S.**, Kohram, M., Vashistha, H., Navlakha, S., Salman, H., Oltvai, Z. N., Bar-Joseph, Z. "Adjustment in tumbling rates improves bacterial chemotaxis on obstacle-laden terrains". In *Proceedings of the National Academy of Sciences (PNAS)*.

Rashid, S., **Singh, S.**, Navlakha, S., & Bar-Joseph, Z. "A Bacterial based Distributed Gradient Descent Model for Mass Scale Evacuations". In *Swarm and Evolutionary Computation*.

2018 **Singh, S.**, Yang, Y., Póczos, B., & Ma, J. "Predicting Enhancer-Promoter Interaction from Genomic Sequence with Deep Neural Networks". *Quantitative Biology*, Special Issue on Applications of Deep Learning to Biology.

## Conference Papers

- 2020 Kim, J., **Singh, S.**, Thiessen, E. D., & Fisher, A. V. *Staying and Returning Dynamics of Sustained Attention in Young Children*. Full paper accepted as Poster Presentation in Annual Meeting of the Cognitive Science Society (CogSci). (63% accept rate)
- 2019 Uppal, A., **Singh, S.** & Póczos, B. *Nonparametric Density Estimation & Convergence of GANs under Besov IPM Losses*. Oral Presentation in Neural Information Processing Systems (NeurIPS). (21.0% accept rate; **Honorable Mention for Outstanding Paper Award**; 0.4% of accepted papers)
- 2018 **Singh, S.**, Uppal, A., Li, B., Li, C.-L., Zaheer, M., & Póczos, B. *Nonparametric Density Estimation under Adversarial Losses*. Accepted as Poster Presentation in Neural Information Processing Systems (NeurIPS). (20.8% accept rate)
- Kim, J., **Singh, S.**, Vande Velde, A., Thiessen, E.D., & Fisher, A.V. *A Hidden Markov Model for Analyzing Eye-Tracking of Moving Objects*. Full paper presented orally in Annual Meeting of the Cognitive Science Society (CogSci). (22.5% accept rate)
- Singh, S.**, Póczos, B., & Ma, J. *Minimax Reconstruction Risk of Convolutional Sparse Dictionary Learning*. International Conference on Artificial Intelligence and Statistics (AISTATS). (33.2% accept rate)
- 2017 **Singh, S.** & Póczos, B. *Nonparanormal Information Estimation*. International Conference on Machine Learning (ICML). (25.5% accept rate)
- Yang, Y., Zhang, R., **Singh, S.**, & Ma, J. *Exploiting sequence-based features for predicting enhancer-promoter interactions*. Full Paper in Intelligent Systems for Molecular Biology (ISMB). (16.5% accept rate; **F1000Prime Recommended**)
- 2016 **Singh, S.**, Du, S., & Póczos, B. *Efficient Nonparametric Smoothness Estimation*. Neural Information Processing Systems (NIPS). (22.7% accept rate)
- Singh, S.** & Póczos, B. *Finite-Sample Analysis of Fixed-k Nearest Neighbor Density Functional Estimators*. Neural Information Processing Systems (NIPS). (22.7% accept rate)
- Singh, S.**, Rashid, S., Navlakha, S., & Bar-Joseph, Z. *Distributed Gradient Descent in Bacterial Food Search*. Full Paper in Research in Computational Molecular Biology (RECOMB). (20.6% accept rate)
- 2014 **Singh, S.** & Póczos, B. *Exponential Concentration of a Density Functional Estimator*. Neural Information Processing Systems (NIPS). (24.7% accept rate)
- Singh, S.** & Póczos, B. *Generalized Exponential Concentration Inequality for Rényi Divergence Estimation*. International Conference on Machine Learning (ICML). (Cycle 1; 14.7% accept rate; **1 of 18 papers nominated for JMLR fast-track**)

## Workshop Papers

- 2020 | Arik, S. O., Li, C.-L., Nikoltchev, M., Sinha, R., Epshteyn, A., Yoon, J., Le, L., Menon, V., **Singh, S.**, Sonthalia, Y., Nakhost, H., Zhang, L., Kanal, E., & Pfister, T. [Interpretable Covid-19 Forecasting](#). ICML Workshop on Healthcare Systems, Population Health, and the role of Health-tech.
- Singh, S.**, Khetan, A., & Karnin, Z. [DARC: Differentiable ARchitecture Compression](#). ICLR Workshop on Neural Architecture Search.
- 2017 | **Singh, S.**, Póczos, B., & Ma, J. *Minimax Reconstruction Risk of Convolutional Sparse Dictionary Learning*. 55th Annual Allerton Conference on Communication, Control, and Computing.
- 2016 | **Singh, S.**, Yang, Y., Póczos, B., and Ma, J. *Predicting Enhancer-Promoter Interaction from Genomic Sequence with Deep Learning*. NIPS Workshop on Machine Learning in Computational Biology.
- 2014 | **Singh, S.**, Navlakha, S., and Bar-Joseph, Z. *Distributed and computationally efficient belief propagation based on swarms of foraging bacteria*.

## Technical Reports

- 2018 | **Singh, S.**, Sriperumbudur, B.K., & Póczos, B., [Nonparametric estimation of Fourier-weighted inner products](#).
- 2016 | **Singh, S.** & Póczos, B. *Analysis of  $k$ -Nearest Neighbor Distances with Application to Entropy Estimation*.

## OTHER PROFESSIONAL ACTIVITIES

### Reviewer for the following journals:

*Journal of Machine Learning Research (JMLR)*

*Bernoulli*

*IEEE Transactions on Information Theory*

*Machine Learning Journal (MLJ)*

*PLOS ONE*

*Communications on Pure and Applied Analysis*

### Reviewer for the following conferences:

Neural Information Processing Systems (NIPS) Conference, 2016–2020

International Conference on Machine Learning (ICML), 2018–2020

– 2019 Top 5% Reviewer (awarded free ICML registration)

Artificial Intelligence and Statistics (AISTATS), 2017–2020

Conference on Learning Theory (COLT), 2018

IEEE International Symposium on Information Theory (ISIT), 2019

International Joint Conference on Artificial Intelligence (IJCAI), 2016

Biological Distributed Algorithms Workshop (BDA), 2016

### Mentoring of PhD Students

Ananya Uppal	PhD in Mathematical Sciences	03/18-Present
	Project: <i>Density Estimation under Besov IPM Losses</i>	

### Mentoring of MS Students

Rui Peng	MS in Machine Learning	08/16-12/17
	Project: <i>Predicting Higher-order Chromatin Interaction from Human Gene Sequence with Deep Neural Networks</i>	
	Member of Data Analysis Project Committee	
Boyue Li	MS in Biotech. Innovation and Computation	08/17-08/18
	Project: <i>Estimating KL Divergence in Exponential Families</i>	

### Teaching Assistantships at Carnegie Mellon University

10-704	Information Processing and Learning	Fall 2016
	<i>Professor Aarti Singh</i>	
36-462	Data Mining	Spring 2015
	<i>Professor Max G'Sell</i>	
36-309	Experimental Design for Behavioral and Social Sciences	Fall 2014
	<i>Professor Howard Seltman</i>	
21-373	Algebraic Structures	Fall 2013
	<i>Professor Rick Statman</i>	
15-251	Great Theoretical Ideas in Computer Science	Spring 2012
	<i>Professors Ryan O'Donnell &amp; Danny Sleator</i>	
15-211	Fundamental Data Structures and Algorithms	Fall 2011
	<i>Professors Jim Morris &amp; Chris Langmead</i>	