



Education	New York University , PhD in Data Science	Sep 2020 - present
	Advisor: <i>Andrew Gordon Wilson</i>	
	<ul style="list-style-type: none"> Center for Data Science Fellowship, 2020-2025 	
	New York University , MS in Computer Science	Sep 2017 - May 2019
	Advisor: <i>Joan Bruna</i>	
	<ul style="list-style-type: none"> Masters Thesis Fellowship, Courant Institute, 2018 	
	IIT Hyderabad , B.Tech in Computer Science	Aug 2012 - May 2016
	<ul style="list-style-type: none"> TODAI Scholarship, University of Tokyo, 2013 Academic Excellence Award, 2012 	
Industry Experience	NVIDIA , Applied Research Intern, USA	Jun 2024 - Aug 2024
	<ul style="list-style-type: none"> Research on in-context learning and large language model (LLM) robustness. 	
	Netflix , Research Intern, USA	Jun 2022 - Aug 2022
	<ul style="list-style-type: none"> Research in probabilistic recommender systems. 	
	Amazon , Applied Science Intern, Germany	Jul 2021 - Sep 2021
	<ul style="list-style-type: none"> Research in multi-fidelity Bayesian optimization. 	
	Uber , AI Resident, USA	Aug 2019 - Jul 2020
	<ul style="list-style-type: none"> Less than 1% acceptance rate; research in approximate Bayesian inference. 	
	Google , Software Engineering Intern, USA	May 2018 - Aug 2018
<ul style="list-style-type: none"> Natural language code search on Kubeflow at KubeCon North America 2018. 		
Headout , Software Engineer, India	Dec 2016 - Jul 2017	
<ul style="list-style-type: none"> Led internal developer tooling; slashed deployment/rollback downtime by 99%. 		
StoryXpress , Co-Founder, India	May 2013 - Aug 2016	
<ul style="list-style-type: none"> Designed the in-house OpenGL video engine for creation at scale. 		
Publications	S. Kapoor* , N. Gruver*, M. Roberts, K. M. Collins, A. Pal, U. Bhatt, A. Weller, S. Dooley, M. Goldblum, and A. G. Wilson. Large Language Models Must Be Taught to Know What They Don't Know. In <i>NeurIPS</i> , 2024	
	S. Qiu*, T. R. Rudner*, S. Kapoor* , and A. G. Wilson. Should We Learn Most Likely Functions or Parameters? In <i>NeurIPS</i> , 2023	
	T. R. Rudner, S. Kapoor , S. Qiu, and A. G. Wilson. Function-Space Regularization in Neural Networks: A Probabilistic Perspective. In <i>ICML</i> , 2023	
	S. Lotfi*, M. Finzi*, S. Kapoor* , A. Potapczynski*, M. Goldblum, and A. G. Wilson. PAC-Bayes Compression Bounds So Tight That They Can Explain Generalization. In <i>NeurIPS</i> , 2022	
	R. Shwartz-Ziv*, M. Goldblum*, H. Sourì, S. Kapoor , C. Zhu, Y. LeCun, and A. G. Wilson. Pre-Train Your Loss: Easy Bayesian Transfer Learning with Informative Priors. In <i>NeurIPS</i> , 2022	
	S. Kapoor* , W. Maddox*, P. Izmailov*, and A. G. Wilson. On Uncertainty, Tempering, and Data Augmentation in Bayesian Classification. In <i>NeurIPS</i> , 2022	

W. J. Maddox, **S. Kapoor**, and A. G. Wilson. When are Iterative Gaussian Processes Reliably Accurate? In *ICML Workshop*, 2021

S. Kapoor and Valerio Perrone. A Simple and Fast Baseline for Tuning Large XGBoost Models, 2021. *Technical report*

N. Gruver, **S. Kapoor**, M. Cranmer, and A. G. Wilson. Epistemic Uncertainty in Learning Chaotic Dynamical Systems. In *ICML Workshop*, 2021

S. Kapoor, M. Finzi, A. Wang, and A. G. Wilson. SKIing on Simplices: Kernel Interpolation on the Permutohedral Lattice for Scalable Gaussian Processes. In *ICML*, 2021a. (**Oral, Top 3%**)

S. Kapoor, T. Karaletsos, and T. D. Bui. Variational Auto-Regressive Gaussian Processes for Continual Learning. In *ICML*, 2021b

T. Moskovitz, R. Wang, J. Lan, **S. Kapoor**, T. Miconi, J. Yosinski, and A. Rawal. First-Order Preconditioning via Hypergradient Descent. In *NeurIPS Workshop*, 2019

S. Kapoor. Leveraging Communication for Efficient Sampling, 2019. *Masters thesis*

C. Resnick*, R. Raileanu*, **S. Kapoor**, A. Peysakhovich, K. Cho, and J. Bruna. Backplay: “Man muss immer umkehren”. In *AAAI Workshop*, 2019

S. Kapoor. Multi-Agent Reinforcement Learning: A Report on Challenges and Approaches, 2018. *Technical report*

Technical Skills

Languages: Python, Node, Javascript, Bash, C, C++, Java

Technologies: PyTorch, JAX, TensorFlow, Pyro PPL, CUDA, MySQL, React, Docker, Ansible, OpenGL

Honors & Awards

StackOverflow Top Contributor: Reputation 6.5k (top 6% overall as of Oct 2024); answers reached ~3.1 million people, 2024

NASSCOM Emerge 50: *StoryXpress* among top startups from 500+ across India for innovation impact, 2015

HYSEA Best Software Product, Student Innovation: *StoryXpress* winner among 100+ startups, 2015

Microsoft Build the Shield: First Runner up among 280 teams across India, 2015

ACM ICPC Amritapuri Regionals: Finalist among 1500+ teams, 2013

Joint Entrance Exam (JEE): Top 0.1% among 0.5 million students across India for undergraduate admissions, 2012

Teaching Experience

Section Leader, *Optimization and Computational Linear Algebra*, NYU Fall 2023

Teaching Assistant, *Introduction to Machine Learning*, NYU Spring 2021

Head Grader, *Machine Learning*, NYU Spring 2019

Teaching Assistant, *Introduction to Machine Learning*, NYU Spring 2019

Section Leader, *Inference and Representation*, NYU Fall 2018

Grader, *Introduction to Machine Learning*, NYU Fall 2018

Recitation Leader, *Data Structures*, NYU Spring 2018

Grader, *Machine Learning*, NYU Spring 2018

Outreach & Services

Reviewer: ICML (2021); NeurIPS (2021,2022,2023); BDL (2021); ICLR (2022,2024)

Instructor: CDS Undergraduate Research Program (2021); NYU AI School (2022)