Stefani Karp

Research Scientist at Google Research
Final-Year PhD Student in Machine Learning
Carnegie Mellon University
shkarp@cs.cmu.edu

Research Interests:

I'm focused on building the theory of deep learning using a combination of mathematics and experiments. I have studied the optimization, generalization, and feature learning capabilities of various neural network architectures under different data modeling assumptions, and I am currently working on understanding and improving the training of Transformers for language (ranging from mathematical analysis to training industry-scale LLMs). More broadly, I'm motivated by trying to (1) understand the nature of intelligence (in both machines and humans), (2) use this understanding to improve our algorithms, and (3) ultimately unlock human-level (and beyond) machine intelligence.

Peer-Reviewed Publications:

1. Aakash Lahoti, Stefani Karp, Ezra Winston, Aarti Singh, Yuanzhi Li. Role of Locality and Weight Sharing in Image-Based Tasks: A Sample Complexity Separation between CNNs, LCNs, and FCNs. ICLR, 2024 (Spotlight).
2. Kavya Ravichandran, Yatin Dandi, Stefani Karp, Francesca Mignacco. Learning from Setbacks: The Impact of Adversarial Initialization on Generalization Performance. NeurIPS Workshop on Mathematics of Modern Machine Learning, 2023.
3. Stefani Karp, Pranjal Awasthi, Satyen Kale. Provable Gradient-Descent-Based Learning of Decision Lists by Transformers. DeepMath, 2023 (Contributed Talk).
4. Sashank J. Reddi, Sobhan Miryoosefi, Stefani Karp, Shankar Krishnan, Satyen Kale, Seungyeon Kim, and Sanjiv Kumar. Efficient Training of Language Models using Few-Shot Learning. ICML, 2023.
5. Ziwei Ji, Kwangjun Ahn, Pranjal Awasthi, Satyen Kale, and Stefani Karp. Agnostic Learnability of Halfspaces via Logistic Loss. International Conference on Machine Learning, ICML, 2022 (Oral).
6. Stefani Karp, Ezra Winston, Yuanzhi Li, and Aarti Singh. Local Signal Adaptivity: Provable Feature Learning in Neural Networks Beyond Kernels. Neural Information Processing Systems, NeurIPS, 2021.
7. Pranjal Awasthi, Satyen Kale, Stefani Karp, and Mehryar Mohri. PAC-Bayes Learning Bounds for Sample-Dependent Priors. Neural Information Processing Systems, NeurIPS, 2020.

Preprints & Technical Reports:

8. Stefani Karp*, Nikunj Saunshi*, Sobhan Miryoosefi, Sashank J. Reddi, Sanjiv Kumar. Landscape-Aware Growing: The Power of a Little LAG. 2024.
9. Cédric Gerbelot, Avetik Karagulyan, Stefani Karp, Kavya Ravichandran, Menachem Stern, Nathan Srebro. Applying Statistical Learning Theory to Deep Learning. Scribe Notes for Nati Srebro's Lectures at Les Houches Summer School on Statistical Physics & Machine Learning (2022). 2023.
10. Stefani Karp, Behnam Neyshabur, and Mehryar Mohri. On the Algorithmic Stability of SGD in Deep Learning. 2020.

Education:

2018 - Present	PhD in Machine Learning Carnegie Mellon University Advised by Aarti Singh
2011 - 2015	BSE in Computer Science, Summa Cum Laude Princeton University Thesis: A New Examination of Persistent Data Structures, advised by Robert Tarjan

Work Experience:

2024 - Present	Research Scientist, Google Research (part-time)
2019 - 2024	Software Engineer, Google Research (part-time)
2015 - 2018	Software Engineer, Google Worked on search quality (ranging from query understanding to ranking, etc.)

Honors and Awards:

2021	Alan J. Perlis Graduate Student Teaching Award (for "the most outstanding graduate TA in CMU's School of Computer Science")
2021	Student Community Leadership Award (Machine Learning Department, CMU)
2015	Outstanding Computer Science Senior Thesis Prize (Princeton University)
2012	Manfred Pyka Memorial Prize in Physics (Princeton University)

Other: Phi Beta Kappa, Tau Beta Pi Engineering Honor Society, Invited to join Sigma Xi

Teaching:

Carnegie Mellon University

Spring 2021	Introduction to Machine Learning (10-701), Teaching Assistant
Spring 2020	Convex Optimization (10-725), Teaching Assistant

University Service:

Carnegie Mellon University

2019 - Present	Machine Learning Department Wellness Network
2019 - Present	Machine Learning Department Blog Editor
2020, 2021	Machine Learning Department Open House Committee