Preetum Nakkiran

preetum@cs.harvard.edu | preetum.nakkiran.org
B.S. EECS, UC Berkeley
PhD Student in CS/ML Theory, Harvard

Education

University of California, Berkeley 2012-2016
B.S. in Electrical Engineering and Computer Science (GPA 4.0/4.0)

Harvard University 2016-Present
PhD Student in Computer Science Theory

Advised by Madhu Sudan and Boaz Barak.

Teaching

- Teaching Assistant, Berkeley: CS 70 (Discrete Mathematics) Fall 2014
- Head Teaching Assistant, Berkeley: EE 16A (Linear Algebra, Systems, and Circuits) Fall 2015
- Head Teaching Assistant, Berkeley: EE 121 (Coding for Digital Communication) Spring 2016
- Teaching Assistant, Harvard: CS 221 (Computational Complexity) Spring 2018

Received Berkeley's **Outstanding Graduate Student Instructor Award (2016)** and Harvard's **Distinction and Excellence in Teaching Award (2018)**.

Industry Experience & Academic Visits

Research Intern: Google Brain Summer 2020

Worked with Hanie Sedghi and Behnam Neyshabur at Google Brain, on generalization in deep learning (ongoing research).

Visiting Student: Jacob Steinhardt Spring 2020

Visited Jacob Steinhardt at UC Berkeley and interacted with the Berkeley ML research groups.

Research Intern: OpenAI Summer 2019

Worked with Ilya Sutskever at OpenAI, on independent research in deep learning. Worked on projects involving adversarial examples, the effect of learning-rate on generalization, and "double descent" – culminating in several publications listed below.

Research Intern: Google Speech Team Summer 2014

Worked on deep-learning for detecting "ok google" audio on a phone. Specifically, we explored techniques for improving the noise-robustness and memory footprint of models used in speech recognition. Based on insights into the weight-structure of existing neural-networks for speech, I developed a new "rank-constrained" neural network architecture for speech that reduces the size of neural networks by 75% without degradation in performance. (Work culminated in two research publications and two patents).

Software Intern: Google Ads Team Summer 2013

On the Ads team, worked on regression models for forecasting trends in ads-metrics (eg, click-through-rates over time). In practice, this achieved an improvement on $\sim 90\%$ of cases on customer data, compared to previous forecasting system.

Selected Talks

- China Theory Week ("Algorithmic Polarization") 2018
- IAS Theory of Deep Learning Workshop ("Deep Double Descent") 2019
- Google X Blueshift ("Distributional Generalization") 2020
- Max Planck Institute MiS Math ML Seminar ("Distributional Generalization") 2020

Publications and Manuscripts

Publications in: STOC, COLT, ITCS, ISIT, APPROX, NeurIPS, ICLR, Distill, INTERSPEECH, ICASSP.

- "**Optimal Regularization Can Mitigate Double Descent**"
  Preetum Nakkiran, Prayaag Venkat, Sham Kakade, Tengyu Ma.
  In submission. 2020.
  https://arxiv.org/pdf/2003.01897.pdf

- "**Deep Double Descent: Where Bigger Models and More Data Hurt**"
  Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, Ilya Sutskever.
  ICLR 2020.
  https://arxiv.org/pdf/1912.02292.pdf

- "**SGD on Neural Networks Learns Functions of Increasing Complexity**"
  Preetum Nakkiran, Gal Kaplun, Dimitris Kalimeris, Tristan Yang, Benjamin L Edelman, Fred Zhang, Boaz Barak.
  NeurIPS 2019 Spotlight.
  https://arxiv.org/pdf/1905.11604

- "**Computational Limitations in Robust Classification and Win-Win Results**"
  Akshay Degwekar, Preetum Nakkiran, Vinod Vaikuntanathan.
  COLT 2019.
  https://arxiv.org/pdf/1902.01086

- "**Adversarial Examples are Just Bugs, Too**"
  Preetum Nakkiran.
  Distill 2019.
  https://distill.pub/2019/advex-bugs-discussion/response-5/

- "**Adversarial robustness may be at odds with simplicity**"
  Preetum Nakkiran.
  Preprint.
  https://arxiv.org/pdf/1901.00532

- "**The Generic Holdout: Preventing False-Positives in Adaptive Data Science**"
  Jarosław Błasiok and Preetum Nakkiran.
  Preprint.
  https://arxiv.org/pdf/1809.05596.pdf

- "**Algorithmic Polarization for Hidden Markov Models**"
  Venkatesan Guruswami, Preetum Nakkiran, Madhu Sudan.
  ITCS 2019.
  https://arxiv.org/pdf/1810.01969.pdf

- "**Differentially Private Simultaneous Mechanisms: A New Model and Mechanisms**"
  Rohit Agrawal, Christina Ilvento, Preetum Nakkiran.
  TPDP 2019.

- "**Tracking the ℓ_2 Norm with Constant Update Time**"
  Chi-Ning Chou, Zhixian Lei, Preetum Nakkiran.
  APPROX-RANDOM 2019.
  https://arxiv.org/pdf/1807.06479.pdf

- "**General Strong Polarization**"
  Jarosław Błasiok, Venkatesan Guruswami, Preetum Nakkiran, Atri Rudra, Madhu Sudan.
  STOC 2018.
  https://arxiv.org/pdf/1802.02718.pdf

- "**Predicting Positive and Negative Links with Noisy Queries: Theory & Practice**"
  Charalampos E. Tsourakakis, Michael Mitzenmacher, Kasper Green Larsen, Jarosław Błasiok, Ben Lawson, Preetum Nakkiran, Vasileios Nakos.
  Allerton 2018.
  https://arxiv.org/pdf/1709.07308.pdf

- "**Near-Optimal UGC-hardness of Approximating Max k-CSP$_R$**"
  Pasin Manurangsi, Preetum Nakkiran, and Luca Trevisan.
  APPROX 2016.
  http://arxiv.org/pdf/1511.06558v1

- "**Optimal Systematic Distributed Storage Codes with Fast Encoding**"
  Preetum Nakkiran, K.V. Rashmi, and Kannan Ramchandran.
  ISIT 2016.
  http://arxiv.org/pdf/1509.01858

- "**Having Your Cake and Eating It Too: Jointly Optimal Erasure Codes for I/O, Storage, and Network-bandwidth**"
  K.V. Rashmi, Preetum Nakkiran, Jingyan Wang, Nihar B. Shah, Kannan Ramchandran.

USENIX Conference on File and Storage Technologies (FAST) 2015.
https://www.usenix.org/system/files/conference/fast15/fast15-paper-rashmi.pdf

- "**Compressing Deep Neural Networks using a Rank-Constrained Topology**"
  Preetum Nakkiran, Raziel Alvarez, Rohit Prabhavalkar, and Carolina Parada.
  Conference of the International Speech Communication Association (Interspeech) 2015.
  http://research.google.com/pubs/archive/43813.pdf

- "**Automatic Gain Control and Multi-style Training for Robust Small-Footprint Keyword Spotting with Deep Neural Networks**"
  Rohit Prabhavalkar, Raziel Alvarez, Carolina Parada, Preetum Nakkiran, and Tara Sainath.
  International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2015.
  http://research.google.com/pubs/archive/43289.pdf

- "**Fundamental limits on communication for oblivious updates in storage networks**"
  Preetum Nakkiran, Nihar B Shah, and K.V. Rashmi.
  IEEE Global Communications Conference (GLOBECOM) 2014.
  http://arxiv.org/pdf/1409.1666

- "**Iterative Hard Thresholding for Keyword Extraction from Large Text Corpora**"
  Steve Yadlowsky, Preetum Nakkiran, Jingyan Wang, Rishi Sharma, and Laurent El Ghaoui.
  International Conference on Machine Learning and Applications (ICMLA) 2014.
  http://www.eecs.berkeley.edu/~elghaoui/Pubs/IhtSummarizationIMCLA14.pdf

Notable Awards

- Google PhD Fellowship Recipient 2020
- NSF Graduate Research Fellowship Recipient 2016
- Patents (with Google): US9767410B1, US9842608B2
- Berkeley Outstanding Graduate Student Instructor Award 2016
- U.S. Physics Team - Top 20 students in US Physics Olympiad 2012

Service

Reviewer: STOC, CRYPTO, ITCS, IEEE Transactions on Information Theory, NeurIPS, JMLR.

Co-founder and co-organizer of Harvard ML Theory Seminar.