

Clayton H. Sanford

claytonsanford.com ◇ clayton@cs.columbia.edu ◇ +1-831-332-0431

LinkedIn: in/claytonsanford ◇ Github: chsanford

EDUCATION

Columbia University

September 2019 - May 2024 (expected)

Ph.D. Candidate in Computer Science

M.S. (Feb. 2021), M.Phil. (Feb 2023)

New York, NY

- Proposed Thesis: “Neural Network Generalization and Approximation with Intrinsically Low-dimensional Data.”
- Advisors: Rocco Servedio and Daniel Hsu

Brown University

September 2014 - May 2018

Sc.B. with Honors in Applied Mathematics - Computer Science

Providence, RI

- Thesis: “Applying Rademacher-Like Bounds to Combinatorial Samples and Function Selection.”
- Thesis Advisor: Eli Upfal; Concentration Advisor: Caroline Klivans
- Magna Cum Laude

INDUSTRY EXPERIENCE

Applied Sciences Intern

May 2023 - August 2023

Microsoft Research

New York, NY

- Trained transformer models with up to 500 million parameters to learn combinatorial search tasks with behavioral cloning and chain-of-thought reasoning.
- Proved theoretical results about the advantages of transformers over graph neural networks (GNNs) for identifying isomorphisms between different combinatorial problems, to support positive empirical results. (Manuscript in progress.)

Research Intern (PhD)

May 2022 - August 2022

Allen Institute for AI

Seattle, WA

- Improved year-long temperature and humidity predictions of ML-corrected coarse-grid climate models by using novelty detection techniques.
- Work presented at NeurIPS 2022 climate ML workshop and American Meteorological Society conference, and under review of the *Journal of Advances in Modeling Earth Systems*.
- Recognized with the Outstanding Internship award, a cash prize awarded to four AI2 interns.

Software Engineering Intern

April 2019 - August 2019

Lumi Labs

Palo Alto, CA

- Front-end and back-end development with direct ownership of new features core to the product.

Associate Analytics Data Scientist

August 2018 - April 2019

LinkedIn

San Francisco, CA

- Used Hive and SQL to create stable and frequently-used datasets that repopulate daily.
- Performed deep-dive analyses on open questions for the LinkedIn Learning product.
- Co-coordinated a bi-weekly machine learning reading group.

Data Analytics Intern

June 2017 - August 2017

LinkedIn

San Francisco, CA

- Analyzed subscription patterns with LinkedIn Learning team using Pig, HDFS, SQL, and Python.
- Contextualized findings in the Learning business and presented to stakeholders.

PUBLICATIONS

- C. Sanford, D. Hsu, M. Telgarsky. “Representational strengths and limitations of transformers.” *Appearing at Neural Information Processing Systems (NeurIPS) 2023*.
- C. Sanford, A. Kwa, O. Watt-Meyer, S. Clark, N. Brenowitz, J. McGibbon, C. Bretherton. “Improving the predictions of ML-corrected climate models with novelty detection.” *Appearing in Journal of Advances in Modeling Earth Systems (JAMES)*.
- C. Sanford*, N. Ardeshir*, D. Hsu. “Intrinsic dimensionality and generalization properties of the R-norm inductive bias.” *Conference on Learning Theory (COLT)*, 2023.
- A. Bietti*, J. Bruna*, C. Sanford*, M. Song*. “Learning single-index models with shallow neural networks.” *Neural Information Processing Systems (NeurIPS)*, 2022.
- V. Chatziafratis*, I. Panageas*, C. Sanford*, S. Stavroulakis*. “On scrambling phenomena for randomly initialized recurrent networks.” *Neural Information Processing Systems (NeurIPS)*, 2022.
- D. Hsu*, C. Sanford*, R. Servedio*, E.-V. Vlatakis-Gkaragkounis*. “Near-Optimal Statistical Query Lower Bounds for Agnostically Learning Intersections of Halfspaces with Gaussian Marginals.” *Conference on Learning Theory (COLT)*, 2022.
- C. Sanford, V. Chatziafratis. “Expressivity of Neural Networks via Chaotic Itineraries beyond Sharkovsky’s Theorem.” *AISTATS*, 2022.
- T. Chin*, J. Ruth*, C. Sanford*, R. Santorella*, P. Carter*, B. Sandstede*. “Enabling equation-free modeling via diffusion maps.” *Journal of Dynamics and Differential Equations*, 2022.
- N. Ardeshir*, C. Sanford*, D. Hsu. “Support vector machines and linear regression coincide with very high-dimensional features.” *Neural Information Processing Systems (NeurIPS)*, 2021.
- D. Hsu*, C. Sanford*, R. Servedio*, E.-V. Vlatakis-Gkaragkounis*. “On the Approximation Power of Two-Layer Networks of Random ReLUs.” *Conference on Learning Theory (COLT)*, 2021.
- K. Cygan*, C. Sanford*, W. Fairbrother. “Spliceman2 - A Computational Web Server That Predicts Sequence Variations in Pre-mRNA Splicing.” *Bioinformatics* 33 (18), 2017.
- J. Gross*, C. Sanford*, G. Kocks*. “Projected Water Needs and Intervention Strategies in India.” *Undergraduate Mathematics and its Applications* 37 (2), 2016.

* Contributed equally

FELLOWSHIPS AND AWARDS

NSF GRFP Fellowship

March 2021

National Science Foundation

- Competitive fellowship that provides three years of full funding for graduate research.

Paul Charles Michelman Memorial Award

May 2023

Columbia Computer Science

- Given to a PhD student in Computer Science who has performed exemplary service to the department, devoting time and effort beyond the call to further the department’s goals (cash prize).

Department Service Award

May 2020, 2022, 2023

Columbia Computer Science

- Awarded up to 10% of PhD students for their service to the department.

Outstanding Intern Award

December 2022

Allen Institute for AI

- Awarded to four summer interns who went above and beyond as researchers and as colleagues (cash prize).

Computer Science Senior Prize

May 2018

Brown Computer Science

- Awarded to the top students in the computer science department based on academic achievement and department service (cash prize).

Outstanding Winner

April 2016

*Interdisciplinary Contest in Modeling**Consortium for Mathematics and its Applications*

- Designation given to five out of over 3000 teams for mathematical modeling of water scarcity in the ICM contest.

LEADERSHIP AND MENTORSHIP EXPERIENCE

Community Board Member

May 2023 - present

*Manhattan Community Board 9**New York, New York*

- Appointed by the borough president to represent community needs of a district on the west side of Manhattan between 110th and 155th St.
- Serves on the the Economic Development/West Harlem Piers Committee and the LGBTQ Task Force.

PhD Representative

May 2022 - present

*Department of Computer Science**Columbia University*

- Coordinated a well-attended PhD student welcome event to help new students visit.
- Attends faculty meetings to represent student concerns and communicate faculty decisions to student body.
- Personally assisted students ensure that the department is paying them adequately and assisted international students with CPT approval issues.

President

September 2022 - September 2023

*qSTEM (LGBTQ affinity for School of Engineering)**Columbia University*

- Organized coffee hours, happy hours, movie showings, and board game nights to build community among LGBTQ grad students.

President

February 2015 - May 2018

*Applied Math Department of Undergraduates (APMA DUG)**Brown University*

- Hosted well-attended advising panels for students interested in Applied Math courses and research.
- Created problems for and managed a casual math competition every semester.
- Coordinated lectures by Applied Math faculty members for undergrads every semester.
- Welcomed prospective students and new concentrators by planning department-sponsored celebrations.

President

November 2014 - May 2018

*Outing Club**Brown University*

- Led an executive board of forty members that ran trips every weekend of the academic year.
- Managed and apportioned a \$27000 annual budget.
- Recruited, interviewed, and trained new trip leaders.

Peer Advisor

September 2017 - May 2018

Matched Advising Program for Sophomores (MAPS)

Brown University

- Advised two sophomore Applied Math students as they declared their concentrations and decided on coursework and internships.

Peer Advisor

September 2015 - May 2017

Meiklejohn Peer Advisory Program

Brown University

- Advised eleven first year students on adjusting to college life, selecting courses, building connections, and finding their academic paths.

RELEVANT COURSEWORK

Algorithms and Theory: Models of Computation, Analysis and Design of Algorithms, Advanced Algorithms Seminar, Computational Linear Algebra, Intro to Cryptography, Randomized Algorithms, Computation and the Brain

Artificial Intelligence: Machine Learning, Artificial Intelligence, Foundations of Prescriptive Analytics, Independent Study for ML research, Optimization Methods for ML, ML Theory, Algorithmic Game Theory

Probability and Statistics: Probability and Computation, Information Theory, Recent Applications in Probability and Statistics, Probabilistic Methods in Computer Science

Dynamical Systems: Applied Ordinary Differential Equations, Applied Partial Differential Equations I, Topics in Chaotic Dynamics, Independent Study for Dynamical Systems Research

Pure Mathematics: Linear Algebra, Abstract Algebra, Analysis: Functions of One Variable

Non-Technical: Persuasive Communication, Classrooms in Context: Public Education in Providence

TEACHING EXPERIENCE

Graduate Teaching Assistant

September 2023 - December 2023

Columbia University Department of Computer Science

- Provided readings for course syllabus, graded assignments, and mentored student final projects for the Machine Learning and Climate graduate seminar, taught by Alp Kucukelbir.

Graduate Instructor

January 2022 - May 2022

Columbia University Department of Computer Science

- Developed and taught a lab on basics of data science and ML for non-CS students to accompany a then-new class on Natural and Artificial Neural Networks by Christos Papadimitriou.
- Created a series of Colab notebooks and short lectures to accompany each topic for a lab with fifteen students.

Graduate Teaching Assistant

January 2021 - April 2021

Columbia University Department of Computer Science

- Held office hours, graded assignments, and prepared course materials for Introduction to Computational Learning Theory, taught by Rocco Servedio.

Head Teaching Assistant

April 2017 - December 2017

Brown University Department of Computer Science

- Led a staff of 14 UTAs through grading assignments, running review sessions, and holding office hours.
- Hired UTAs after interviewing 35 candidates for the job.
- Managed an Algorithms class with 170 students and coordinated interactive grading sessions and exams.
- Taught an supplemental section on NP-hardness to a group of forty students for 90 minutes.
- Brainstormed, wrote-up, and edited problems for homework assignments and exams.

Undergraduate Teaching Assistant

September 2015 - May 2017

Brown University Departments of CS and Applied Math

- Served on the course staffs of four courses: Accelerated Intro to CS, Discrete Structures and Probability, Theory of Computation, Topics in Chaotic Dynamics.
- Created problems for and graded homework assignments and exams.
- Hosted office hours for helping students understand course material and solve homework problems.

Tutor and Volunteer Representative

January 2015 - May 2016

Swearer Tutoring Enrichment in Math and Science (STEMS)

- Tutored math and science in class and after school at a nearby public school in Providence.
- Interviewed potential volunteers and planned meetings to help train tutors.

Tutor

September 2011 - June 2014

Soquel High School

- Tutored math at homework club after school twice a week for three years.

RESEARCH TALKS

- Formal Languages and Neural Networks (FLANN) talk, virtual, November 2023. “Representational Strengths and Limitations of Transformers.”
- Joan Bruna group meeting, NYU, October 2023. “Representational Strengths and Limitations of Transformers.”
- “Theory and Practice of Foundation Models” invited speaker, Yale, October 2023. “Representational Strengths and Limitations of Transformers.”
- Algorithms seminar, Google NYC, July 2023. “Representational Strengths and Limitations of Transformers.”
- Data Science guest lecture, UC San Diego, May 2023. “Representational Strengths and Limitations of Transformers.”
- Engaged Scholars Program Research Symposium, Columbia University, April 2023. “Machine Learning, Neural Networks, and CS Theory.”
- Columbia StatisticalML Symposium, Columbia University, April 2023. “Transformers can learn pairwise—but not three-wise—functions.”
- 22nd Conference on Artificial Intelligence for Environmental Science, 103rd American Meteorological Society Annual Meeting, January 2023. “Improving the Predictions of ML-Corrected Climate Models with Novelty Detection.”
- Tackling Climate Change with Machine Learning workshop, NeurIPS 2022, December 2022. “Improving the Predictions of ML-Corrected Climate Models with Novelty Detection.”
- Vaggos Chatziafratis group meeting, UC Santa Cruz, November 2022. “Why do over-parameterized neural networks work?”
- Kevin Jamieson, Jamie Morgenstern, and Ludwig Schmidt group meeting, University of Washington, July 2022. “Approximation Powers and Limitations of Neural Networks.”

- COLT 2022, July 2022. “Near-Optimal Statistical Query Lower Bounds for Agnostically Learning Intersections of Halfspaces with Gaussian Marginals.”
- Algorithms and Theory Seminar, Boston University, April 2022. “On the approximation power of two-layer networks of random ReLUs.”
- Algorithms and Complexity seminar, MIT, April 2022. “On the approximation power of two-layer networks of random ReLUs.”
- Eli Upfal group meeting, Brown University, April 2022. ““Benign overfitting” and the behavior of high-dimensional linear regression and classification models.”
- AISTATS, March 2022. “Expressivity of Neural Networks via Chaotic Itineraries beyond Sharkovsky’s Theorem.”
- Joan Bruna group meeting, NYU, February 2022. “Near-Optimal Statistical Query Lower Bounds for Agnostically Learning Intersections of Halfspaces with Gaussian Marginals.”
- Data Science Institute virtual poster session Columbia, February 2022. “On the approximation power of two-layer networks of random ReLUs.”
- NeurIPS, December 2021. “Support vector machines and linear regression coincide with very high-dimensional features.”
- COLT 2021, August 2021. “On the approximation power of two-layer networks of random ReLUs.”
- Demystifying the dissertation, Columbia, December 2020. “Opening the Black Box: Mathematical Approaches to Understanding Deep Learning.”
- Demystifying the PhD, Columbia, November 2020.

DEPARTMENT SERVICE

- Organized the CS theory student retreat in Fall 2021 and 2022.
- Ran events at and coordinated the Columbia Visit Day for admitted students in Spring 2020, 2022, and 2023.
- Started the Columbia Theory Student seminar, where students share their research on a weekly basis.
- Advised three cohorts of undergraduate theory seminars on ML and deep learning theory.

ACADEMIC SERVICE

Reviewer

- ICLR (2024), ALT (2024), JMLR (2023), NeurIPS (2023), ICLR climate workshop (2023), SODA (2022), STOC (2022).

TECHNICAL SKILLS AND INTERESTS

Programming Languages

Python, Java, Scala, SQL

Technologies

Pytorch, Tensorflow, Docker, Hadoop, Spark, Git

Spoken Languages

English (native), Spanish (intermediate proficiency)

Other Interests

Backpacking, Running, Climbing, Cooking, New York, Public Transportation