

GEELON SO

INTERESTS	Learning Theory, Theoretical Computer Science	
EDUCATION	UNIVERSITY OF CALIFORNIA SAN DIEGO	La Jolla, CA
	Ph.D. Student, Computer Science	<i>Sep 2019—present</i>
	<i>Relevant courses:</i> Probability Theory, Stochastic Analysis, Computational Statistics, Unsupervised Learning, Continual Learning, Computational Neurobiology, Lattice Algorithms, Cybersecurity	
	COLUMBIA UNIVERSITY	New York, NY
	M.S. Computer Science	<i>May 2019</i>
	<i>Relevant courses:</i> Machine learning, Unsupervised Learning, Graph Theory, Learning Theory, Information Theory, Algorithms through Geometric Lens, Computer Networks, Privacy-Preserving Technologies	
	THE UNIVERSITY OF CHICAGO	Chicago, IL
	B.S. Mathematics with Honors	<i>Jun 2017</i>
ACADEMIC EXPERIENCE	<i>Relevant courses:</i> Algebraic Geometry, Commutative Algebra, Algebraic Number Theory, Topology, Honors Analysis (real & functional analysis, measure theory), Markov Chains, Quantum Mechanics	
	M.S. Thesis	New York, NY
	<i>Columbia University</i>	<i>Jan 2018—Sep 2019</i>
	<ul style="list-style-type: none">Constructed a general framework to design and analyze active learning algorithms that disentangles the statistical and geometric components of learning; the framework easily extended a previous result to more general noise settings. <i>Title:</i> Active learning with noise, <i>Advisor:</i> Daniel Hsu	
	Reading Seminars	New York, NY
	<i>Columbia University</i>	<i>Sep 2018—May 2019</i>
	<ul style="list-style-type: none">Presented and scribed lectures in a reading group focused on sums-of-squares optimization, generalization theory (Columbia, Jan—May 2018) and type theory (UChicago, Sep 2016—Jun 2017)	
	REUs/Research Internship	Chicago, IL
	<i>University of Chicago</i>	<i>Summers 2014, 2015, 2016</i>
	<ul style="list-style-type: none">Studied foundations of mathematics through intuitionistic type theory and category theory; advised by Prof. Stuart Kurtz. Wrote expository papers on symbolic dynamics and the mathematics of quantum computing; advised by Clark Butler, Tori Akin, and Prof. Peter MayAttended summer courses, including discrete mathematics and combinatorics, algebraic topology, spectral geometry, random walks, ergodic theory	
	Summer School in Mathematics	Chicago, IL
	<i>University of Chicago</i>	<i>Summers 2015, 2016</i>
TEACHING EXPERIENCE	<ul style="list-style-type: none">Attended two invitational graduate-level math seminars in analysis and topology	
	Teaching Assistant Fellowship	New York, NY
	<i>Columbia University</i>	<i>Sep 2018—Dec 2018</i>
	<ul style="list-style-type: none">Presented 15 hours of lectures on unsupervised learning techniques; designed 9 homework problems; taught over 100 hours during office hours and individual meetings; instructor: Prof. Nakul Verma<i>Unsupervised Learning</i> and <i>Machine Learning</i>, Summer 2018, Prof. Nakul Verma; <i>Graph Theory</i>, Spring 2018, Tim Sun; <i>Geographic Information Systems</i>, Fall 2017, Prof. Michael Parrott	

WORK EXPERIENCE	Data Science & Engineering Internship	Chicago, IL
	<i>Home Partners of America</i>	<i>Jun 2019—Aug 2019</i>
	<ul style="list-style-type: none"> Identified, planned and executed a high-impact project given limited resource and time constraints Streamlined ETL for the analytics team by providing API to automatically join, transform and aggregate data, while performing data quality checks with statistical guarantees; eliminated need to understand how the history of business decisions affects the correct way to join tables Designed backend to API to be easily adaptable to future changes in data collection/database 	
	Technology Consulting	New York, NY
	<i>Hacking 4 Defense (Department of Defense & Columbia University)</i>	<i>Sep 2018—Dec 2018</i>
	<ul style="list-style-type: none"> Collaborated with team to perform research, interview subject matter experts, and design system that gives first responders situational awareness of surrounding HAZMATs, while developing project management and teamwork skills; sponsors: Col. Patrick Mahaney, Jason Cahill and Paul Blaer 	
	Machine Learning Consulting	New York, NY
	<i>AumiPhyte Health (Healthcare Startup)</i>	<i>Feb 2018—Mar 2019</i>
	<ul style="list-style-type: none"> Designed user-centric system and tools to analyze and process medical texts, presenting recommendations in a whitepaper that detailed motivation, implementation, rationale, and limitations Implemented rules- and machine-learning based text processing methods using python and spaCy to generate label data for small set of medical texts; estimated 0.9 accuracy with 0.9 confidence Developed GUI utilizing active learning techniques to help user rapidly label data 	
	Nonprofit Consulting	Chicago, IL
	<i>campusCatalyst (Options for Youth & University of Chicago)</i>	<i>Mar—Jun 2016</i>
	<ul style="list-style-type: none"> Conducted quantitative/qualitative impact assessment for a local nonprofit by (i) researching and building a model to estimate the return on investment and (ii) interviewing stakeholders and recording impact; provided actionable best practices for using impact assessment to target donors 	
SERVICE	Friends of Washington Park	Chicago, IL
	<i>Volunteer</i>	<i>Jan 2014—Jul 2017</i>
	<ul style="list-style-type: none"> Tutored 5th–8th grader students in an after-school program in South Side Chicago 	
SKILLS	Technical: Python, TensorFlow, C, Haskell, Bash, Linux, GIS, SQL, technical writing	
AWARDS	2019 Andrew P. Kosoresow Memorial Award for Excellence in Teaching and Service Awarded for outstanding contributions to teaching in the Department of Computer Science at Columbia University and exemplary service to the Department and its mission.	
PUBLICATIONS	Preprints	
	<ul style="list-style-type: none"> Dasgupta, Mahajan, So. <i>Convergence of online k-means</i> (2021). 	<i>Nov 2021</i>
PRESENTATIONS	Lectures for Courses/Reading Seminars	
	<ul style="list-style-type: none"> Generalization through differential privacy 	<i>Nov 2019</i>
	<ul style="list-style-type: none"> Zero-knowledge proofs 	<i>Apr 2019</i>
	<ul style="list-style-type: none"> Geometric gradient descent and lower bounds 	<i>Feb 2019</i>
	<ul style="list-style-type: none"> Introduction to homomorphic encryption 	<i>Feb 2019</i>
	<ul style="list-style-type: none"> Approximate near neighbor search 	<i>Dec 2018</i>
	<ul style="list-style-type: none"> Introduction to tensors for estimation 	<i>Dec 2018</i>
	<ul style="list-style-type: none"> Sums-of-squares for robust estimation 	<i>Nov 2018</i>
	<ul style="list-style-type: none"> Spectral graph theory 	<i>Oct 2018</i>
	<ul style="list-style-type: none"> Sums-of-squares for MAXCUT 	<i>Oct 2018</i>
	<ul style="list-style-type: none"> Topological data analysis 	<i>Jul 2018</i>
	<ul style="list-style-type: none"> Tensor decomposition for density estimation 	<i>Jul 2018</i>
	<ul style="list-style-type: none"> PAC-Bayes for ReLu Neural Networks 	<i>Apr 2018</i>