9500 Gilman Drive, Mail Code 0404
La Jolla, CA 92093
geelon@ucsd.edu
https://geelon.github.io
773-621-9265

GEELON SO

INTERESTS   Learning Theory, Theoretical Computer Science

EDUCATION

UNIVERSITY OF CALIFORNIA SAN DIEGO   La Jolla, CA
Ph.D. Student, Computer Science   *Sep 2019—present*
GPA: 3.91/4.0
Relevant courses: Probability Theory, Stochastic Analysis, Computational Statistics, Unsupervised Learning, Continual Learning, Computational Neurobiology, Lattice Algorithms, Cybersecurity

COLUMBIA UNIVERSITY   New York, NY
M.S. Computer Science   *May 2019*
GPA: 3.70/4.0
Relevant courses: 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   *Jun 2017*
GPA: 3.56/4.0
Relevant courses: Algebraic Geometry, Commutative Algebra, Algebraic Number Theory, Topology, Honors Analysis (real & functional analysis, measure theory), Markov Chains, Quantum Mechanics

ACADEMIC
EXPERIENCE

M.S. Thesis   New York, NY
Columbia University   *Jan 2018—Sep 2019*
- 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. *Title:* Active learning with noise, *Advisor:* Daniel Hsu

Reading Seminars   New York, NY
Columbia University   *Sep 2018—May 2019*
- 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
University of Chicago   *Summers 2014, 2015, 2016*
- 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 May
- Attended summer courses, including discrete mathematics and combinatorics, algebraic topology, spectral geometry, random walks, ergodic theory

Summer School in Mathematics   Chicago, IL
University of Chicago   *Summers 2015, 2016*
- Attended two invitational graduate-level math seminars in analysis and topology

TEACHING
EXPERIENCE

Teaching Assistant Fellowship   New York, NY
Columbia University   *Sep 2018—Dec 2018*
- 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
- *Unsupervised Learning* and *Machine Learning*, Summer 2018, Prof. Nakul Verma; *Graph Theory*, Spring 2018, Tim Sun; *Geographic Information Systems*, Fall 2017, Prof. Michael Parrott

WORK EXPERIENCE	**Data Science & Engineering Internship**	Chicago, IL

WORK EXPERIENCE

Data Science & Engineering Internship — Chicago, IL
Home Partners of America — *Jun 2019—Aug 2019*
- 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
Hacking 4 Defense (Department of Defense & Columbia University) — *Sep 2018—Dec 2018*
- 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
AumiPhyte Health (Healthcare Startup) — *Feb 2018—Mar 2019*
- 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
campusCatalyst (Options for Youth & University of Chicago) — *Mar—Jun 2016*
- 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
Volunteer — *Jan 2014—Jul 2017*
- 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
- Dasgupta, Mahajan, So. *Convergence of online k-means* (2021). — *Nov 2021*

PRESENTATIONS

Lectures for Courses/Reading Seminars
- Generalization through differential privacy — *Nov 2019*
- Zero-knowledge proofs — *Apr 2019*
- Geometric gradient descent and lower bounds — *Feb 2019*
- Introduction to homomorphic encryption — *Feb 2019*
- Approximate near neighbor search — *Dec 2018*
- Introduction to tensors for estimation — *Dec 2018*
- Sums-of-squares for robust estimation — *Nov 2018*
- Spectral graph theory — *Oct 2018*
- Sums-of-squares for MAXCUT — *Oct 2018*
- Topological data analysis — *Jul 2018*
- Tensor decomposition for density estimation — *Jul 2018*
- PAC-Bayes for ReLu Neural Networks — *Apr 2018*