





# GEELON SO

Ph.D. Student

 geelon.github.io  
 1 773 621 9265  
 geelon@ucsd.edu

 9500 Gilman Drive  
Mail Code 0404  
La Jolla, CA 92093

## RESEARCH INTERESTS

Machine learning theory, sequential decision making, stochastic analysis, optimization, geometry

## SKILLS

Python, PyTorch, spaCy

## EDUCATION

2019 - exp. 2025  
GPA: 3.91/4.0

### Ph.D. Computer Science

University of California, San Diego

*Advisors:* Sanjoy Dasgupta, Yian Ma

*Related coursework:* Probability Theory, Stochastic Analysis, Computational Statistics, Unsupervised Learning, Continual Learning, Differential Geometry, Computational Neurobiology, Lattice Algorithms

2017 - 2019  
GPA: 3.70/4.0

### M.S. Computer Science

Columbia University

*Advisor:* Daniel Hsu

*Thesis:* Active Learning with Noise

2013 - 2017  
GPA: 3.59/4.0

### B.S. Mathematics

The University of Chicago

*Advisor:* Stuart Kurtz

## EXPERIENCE

6/2022 - 6/2023 **Seekr Technologies** | Machine Learning Research Intern

- Developed a human-in-the-loop machine learning system for extracting labels for new article clusters.
- Engineered a data exploration tool and a user-friendly GUI for human evaluation of label quality.

6/2019 - 8/2019 **Home Partners of America** | Data Engineering Intern

- Streamlined ETL for the analytics team by providing API to automatically join, transform and aggregate data, while performing data quality checks with statistical guarantees.

## RESEARCH

2019 - present  
La Jolla, CA

### UC San Diego, Department of Computer Science and Engineering | Ph.D. Graduate Researcher

- Research work in algorithmic foundations of machine learning, especially in (i) learning and decision making in the online and strategic/game-theoretic settings, (ii) optimization with theoretical guarantees, and (iii) high-dimensional statistics. Advised by Sanjoy Dasgupta and Yian Ma.
- *Keywords:* theoretical guarantees, stochastic approximation, online/streaming algorithms, active learning,  $k$ -means, nearest neighbor, Langevin dynamics, convex optimization, multi-objective optimization

2017 - 2019  
New York, NY

### Columbia University, Data Science Institute | M.S. Graduate Researcher

- Research work in active learning for the classification setting with noise. Advised by Daniel Hsu.

6/2014 - 9/2016  
Chicago, IL

### University of Chicago REU | Undergraduate Researcher

- Studied foundations of mathematics, intuitionistic type theory. Advised by Stuart Kurtz.
- Studied dynamical systems and quantum computing. Advised by Clark Butler, Tori Akin, and Peter May.

## PUBLICATIONS

Authors in alphabetical order unless otherwise noted.

### Optimization on Pareto sets: On a theory of multi-objective optimization.

Abhishek Roy\*, Geelon So\*, Yian Ma. arXiv, under review, 2023. \*equal contribution

### Online nearest neighbor classification.

Sanjoy Dasgupta and Geelon So. arXiv, under review, 2023.

### Convergence of online $k$ -means.

Sanjoy Dasgupta, Gaurav Mahajan, Geelon So. *International Conference on Artificial Intelligence and Statistics*, 2022.

### Active learning with noise.

Geelon So. *Master's thesis, Columbia University*, 2019.

## ACHIEVEMENTS

- 2020 UC San Diego Changemaker Challenge 1st Place
- 2019 Andrew P. Kosoresow Memorial Award for Excellence in Teaching and Service