

Carol (Xuan) Long

✉ carol.long@g.harvard.edu • 🌐 carol-long.github.io

Education

Harvard University

Ph.D. in Applied Mathematics, School of Engineering and Applied Sciences (SEAS)
Advisor: Flavio du Pin Calmon

Cambridge, MA
Sep. 2021 – Expected May. 2026

Harvard University

M.S. in Applied Mathematics, School of Engineering and Applied Sciences (SEAS)
GPA: 3.78/4.0

Cambridge, MA
Sep. 2021 – Sep. 2023

New York University

B.A. in Mathematics and Computer Science, Courant Institute of Mathematical Sciences
GPA: 4.0/4.0, Summa Cum Laude, Phi Beta Kappa Honor Society, Graduated in 3 years

New York, NY
Sep. 2018 – May 2021

Research Interests

Areas: Responsible AI, Information Theory, Statistical Learning Theory

Topics: LLM Watermarking, Algorithmic Fairness, Model Multiplicity

Publications

C. X. Long, W. Alghamdi, A. Glynn, Yixuan Wu, F. P. Calmon, “Kernel Multiaccuracy,” *Under Review*, 2024.

A. Osterling, C. M. Verdun, C. X. Long, A. Glynn, L. M. Paes, S. Vithana, M. Cardone, F. P. Calmon, “Multi-Group Proportional Representation,” *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

J. Watson-Daniels, F. P. Calmon, A. D’Amour, C. X. Long, D. C. Parkes, B. Ustun, “Predictive churn with the Set of Good Models,” *Under Review*, 2024.

C. X. Long, H. Hsu, W. Alghamdi, and F. P. Calmon, “Arbitrariness Lies Beyond the Fairness-Accuracy Frontier,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. **Spotlight Paper.**

L. M. Paes*, C. X. Long*, B. Ustun, and F. P. Calmon, “On the Epistemic Limits of Personalized Prediction,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. *Equal contribution.

Awards, Honors, and Scholarships

Advances in Neural Information Processing Systems (NeurIPS) Travel Award 2023

Harvard University Kao Fellowship 2022
(for exceptional graduate students at Harvard SEAS)

North America School of Information Theory (NASIT) Travel Award 2022

Courant Institute Mathematics Award for Academic Achievement 2021
(presented to graduating seniors for excellence in mathematics)

NYU Alumni Award 2021
(presented to a graduating senior for scholarship and general attainments in Science)

NYU Women in Science Fellowship 2020-2021
(~ 10 students/year)

Courant Institute Summer Undergrad Research Experience Fellowship 2020
(~ 15 students/year)

NYU Dean’s Honors List 2018-2021
(every academic year)

Singapore Ministry of Education Senior-Middle 1 Scholarship 2013-2017
(~ 200 Chinese students/year, full scholarship to attend high school in Singapore)

Professional and Research Experiences

Graduate Research Assistant/ Teaching Fellow

Harvard University, Advisor: Prof. Flavio P. Calmon

Cambridge, MA

Sep. 2021 – Present

- Conduct research on information-theoretic frameworks for responsible AI (e.g., watermarking in LLMs, algorithmic fairness)
- Demonstrate that model arbitrariness is orthogonal to fairness and accuracy in ML models
- Characterize the information-theoretic limit of ensuring “Fair Use” of group attributes in ML models
- Teaching Fellow for Fall 2022 ENI-SCI 250: Information Theory

Quantitative Research Intern

Citadel LLC, Global Quantitative Strategies Team

Hong Kong, China

June 2024 – Sep. 2024

- Apply advanced statistical and quantitative modeling methods to identify opportunities in the equities market
- Develop proprietary research and datasets to yield investment insights that can be adopted into central operations

Software Engineering Intern

Meta Platforms (formerly Facebook), Ads ML Infra Team

Menlo Park, CA

May 2021 – Aug. 2021

- Design a mechanism that improves stability and reliability of the feature pipeline for the Ads ranking ML system
- Implement the design using C++ to optimize the algorithm and ensure efficiency during real-time delivery

Undergraduate Summer Research

NYU Courant SURE Fellowship, PI: Yunan Yang

New York, NY

Jul. 2020 – Oct. 2020

- Modify the K-Means algorithm to enhance its ability to capture geometry in data by replacing Euclidean centroid with Wasserstein barycenter
- Benchmark algorithms for computing Wasserstein distance by investigating regularized and sliced Wasserstein distance against the traditional Kantorovich formulation
- Apply the augmented algorithm on the 99 Shapes Dataset and demonstrate significant visual improvement

Undergraduate Research Assistant

NYU Department of Psychology, PI: Elena Sizikova

New York, NY

Sep. 2019 – May 2020

- Investigate the existence of human vision’s physiological phenomena in CRNN, a deep neural network, to shed light on the root cause of Dyslexia
- Exposed CRNN’s limitations in modeling noise adaptation and crowding in human vision, as well as its low efficiency, as measured by the signal-to-noise ratio of the images
- Generate images similar to human perception by applying Gaussian-noise techniques with Numpy and Matplotlib
- Explore a new method to model reaction time in dyslexia by measuring FLOPs (Floating Point Operations per Second) in networks

Undergraduate Research Volunteer

NYU Langone Health

New York, NY

Oct 2019 – Feb 2020

- Classify attack/non-attack behavior of a mouse using its brain activity data by training a long short-term memory model (LSTM) on High Performance Computing cluster
- Investigate the problem of leveraging the machine learning model to automate the process of labeling mice activity videos

Teaching Experiences

ES 250: Information Theory – Graduate Level Course

Fall 2022

Engineering and Applied Sciences — Harvard University

Teaching Assistant

Hold weekly office hours to address students’ concerns and questions. Guide 10+ students on their final projects.

NYU Courant Undergrad Tutor

2020-2021

Hold weekly office hours on Undergrad Math classes including Discrete Math, Linear Algebra, and Calculus I, II, III.

Professional Service and Activities

Conference Reviewer

- Neural Information Processing Systems (NeurIPS), 2022, 2023
- International Conference on Machine Learning (ICML), 2023
- ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2023, 2024

Conference and Workshop Attendance

- International Conference on Machine Learning (ICML), 2022
- Neural Information Processing Systems (NeurIPS), 2022, 2023, 2024
- North America School of Information Theory (NASIT), including a poster presentation, 2022

Skills and Interests

Technical Skills:

- Programming Languages: Python, C/C++, MATLAB
- Software and Packages: PyTorch, Jupyter Notebook, Matplotlib, Numpy, Pandas, Scikit-Learn
- Other: HPC, terminal, git

Languages: English, Mandarin, Cantonese

Interests: Dancing (ballet, contemporary, hip hop), Learning languages, Long-distance running (half-marathon), Music (classical, jazz, piano, pop), Traveling