clara wong-fannjiang

clarafy@berkeley.edu | clarafy.github.io

education

University of California, Berkeley 2018 - 2023 (expected)

Ph.D., Electrical Engineering & Computer Sciences
Advised by Jennifer Listgarten and Michael I. Jordan

Stanford University 2012 - 2016

B.S., Computer Science with Honors and Distinction

publications

Asterisks denote equal contribution.
(α-β) denotes alphabetical ordering.

preprints

1. (α-β) Anastasios N. Angelopoulos, Stephen Bates, **Clara Fannjiang**, Michael I. Jordan, and Tijana Zrnic. Prediction-powered inference. 2023. DOI: 10.48550/arXiv:2301.09633. (arXiv | code)
2. Danqing Zhu*, David H. Brookes*, Akosua Busia*, Ana Carneiro, **Clara Fannjiang**, Galina Popova, David Shin, Tomasz J. Nowakowski, Jennifer Listgarten, and David. V. Schaffer. Optimal trade-off control in machine learning-based library design, with application to adeno-associated virus (AAV) for gene therapy. 2021. DOI: 10.1101/2021.11.02.467003. (bioRxiv)
3. Akosua Busia, George E. Dahl, **Clara Fannjiang**, David H. Alexander, Elizabeth Dorfman, Ryan Poplin, Cory Y. McLean, Pi-Chuan Chang, and Mark DePristo. A deep learning approach to pattern recognition for short DNA. 2019. DOI: 10.1101/353474. (bioRxiv)

refereed conferences

1. Ghassen Jerfel*, Serena Wang*, **Clara Fannjiang**, Katherine Heller, Yian Ma, Michael Jordan. Variational refinement for importance sampling using the forward Kullback-Leibler divergence. 2021. In C. P. de Campos, M. H. Maathuis, and E. Quaeghebeur (Eds.), *Proc. of the 37th Conference on Uncertainty in Artificial Intelligence (UAI)*, PMLR 161:1819-1829. A previous version of this work appeared in *Advances in Approximate Bayesian Inference (AABI) 2021*. (arXiv)
2. **Clara Fannjiang** and Jennifer Listgarten. Autofocused oracles for model-based design. 2020. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), *Adv. in Neural Information Processing Systems (NeurIPS) 33*. (arXiv | proceedings | code)
3. David H. Brookes, Akosua Busia, **Clara Fannjiang**, Kevin Murphy, and Jennifer Listgarten. A view of estimation of distribution algorithms through the lens of expectation-maximization. 2020. In *Proc. of the 2020 Genetic and Evolutionary Computation Conference (GECCO)*, 189-190. (arXiv (extended version) | proceedings)

journals

1. **Clara Fannjiang**, Stephen Bates, Anastasios Angelopoulos, Jennifer Listgarten, and Michael I. Jordan. Conformal prediction under feedback covariate shift for biomolecular design. 2022. *Proceedings of the National Academy of Sciences*, 119(43), e2204569119. DOI: 10.1073/pnas.2204569119. (arXiv | publication | code | talk)

2. Chloe Hsu, Hunter Nisonoff, **Clara Fannjiang**, Jennifer Listgarten. Learning protein fitness models from evolutionary and assay-labelled data. 2022. *Nature Biotechnology*, 40, 1114–1122. DOI: 10.1038/s41587-021-01146-5. (PDF | publication)
3. I. Masmitja, J. Navarro, S. Gomariz, J. Aguzzi, B. Kieft, T. O'Reilly, K. Katija, P. J. Bouvet, **C. Fannjiang**, M. Vigo, P. Puig, A. Alcocer, G. Vallicrosa, N. Palomeras, M. Carreras, J. Del-Rio, J. B. Company. 2020. Mobile robotic platforms for the acoustic tracking of deep-sea demersal fishery resources. Science Robotics, 5(48), eabc3701. DOI: 10.1126/scirobotics.abc3701. (PDF | publication)
4. **Clara Fannjiang**, T. Aran Mooney, Seth Cones, David Mann, K. Alex Shorter, and Kakani Katija. 2019. Augmenting biologging with supervised machine learning to study *in situ* behavior of the medusa *Chrysaora fuscescens*. *Journal of Experimental Biology*, 222, jeb207654. DOI: 10.1242/jeb.207654. (PDF | publication | code)
5. **Clara Fannjiang**. 2013. Optimal arrays for compressed sensing in snapshot-mode radio interferometry. *Astronomy & Astrophysics*, 559, A73. DOI: 10.1051/0004-6361/201321079. (PDF | publication)

refereed workshops

1. **Clara Fannjiang**, Micah Olivas, Eric R. Greene, Craig J. Markin, Bram Wallace, Ben Krause, Margaux M. Pinney, James Fraser, Polly Fordyce, Ali Madani, Nikhil Naik. Designing active and thermostable enzymes with sequence-only predictive models. Workshop on Learning Meaningful Representations of Life (LMRL) at Neural Information Processing Systems (NeurIPS) 2022.
2. Katherine Lee, Orhan Firat, Ashish Agarwal, **Clara Fannjiang**, and David Sussillo. Hallucinations in neural machine translation. Workshop on Interpretability and Robustness for Audio, Speech, and Language at Neural Information Processing Systems (NeurIPS) 2018. (PDF)

other conferences

1. Ivan Masmitja, Spartacus Gomariz, Joaquin Del Rio, Brian Kieft, Tom O'Reilly, Jacobo Aguzzi, Pierre-Jean Bouvet, **Clara Fannjiang**, and Kakani Katija. Area-only method for underwater object tracking using autonomous vehicles. IEEE OCEANS 2019. (PDF | proceedings)
2. **Clara Fannjiang** and Kakani Katija. Using supervised machine learning to understand fine-scale *in situ* behavior of *Chrysaora fuscescens*. Society for Integrative and Comparative Biology (SICB) 2019 Annual Meeting. Oral, Marlene Zuk Best Student Paper Finalist.
3. **Clara Fannjiang**, Marius Cătălin Iordan, Diane M. Beck, and Fei-Fei Li. Pushing the boundaries of fine-grained object classification with fMRI decoding in human occipito-temporal cortex. Vision Sciences Society (VSS) 2015 Annual Meeting. Poster. DOI: 10.1167/15.12.1167. (abstract)

talks

seminars

1. Conformal prediction for biomolecular design. Physics of Living Systems Seminar Series (EPFL), November 18, 2022.
2. Conformal prediction for the design problem. Machine Learning for Protein Engineering Seminar Series. October 18, 2022. (recording)
3. Conformal prediction for the design problem. AIDD Summer School on Advanced Machine Learning for Drug Discovery, May 13, 2022.
4. Augmenting biologging with supervised machine learning to study *in situ* behavior of the medusa *Chrysaora fuscescens*. Monterey Bay Aquarium Research Institute Seminar, 2018.

external lab meetings

1. A machine learning-guided tour of thermostability and activity in two enzyme families. James Fraser Lab (UCSF), October 31, 2022.
2. Autofocused oracles for model-based design. Debora Marks Lab (Harvard), August 2020.

professional experience

Salesforce Research, Research Intern — Summer 2022

Developed method leveraging protein language models and thermal proteome profiling data to design thermostable enzymes. Currently collaborating with the Fordyce, Pinney, and Fraser labs at Stanford and UCSF to experimentally characterize thedesigned enzymes.

Monterey Bay Aquarium Research Institute, Research Assistant — 2018 - 2019

Designed and conducted biologging field experiments on jellyfish in Monterey Bay, CA. Developed supervised learning methods to characterize novel fine-scale *in situ* behavioral and movement patterns.

Google, Google Brain Resident — 2016 - 2017

Developed regularization schemes for suppressing chaotic dynamics in recurrent neural networks. Contributed to deep learning approach for taxonomic identification of genetic reads.

teaching

University of California, Berkeley

STAT 102: Data, Inference, and Decisions (Graduate Student Instructor)	Fall 2020
STAT 102: Data, Inference, and Decisions (Graduate Student Instructor & Guest Lecturer)	Spring 2020

Stanford University

EE 364A: Convex Optimization (Teaching Assistant)	Winter 2016
EE 103: Introduction to Matrix Methods (Teaching Assistant)	Fall 2015

leadership & outreach

Mentor for Berkeley AI Research Undergraduate Mentoring Program	2019 - present
Coordinator for UC Berkeley WICSE	2019 - 2020
Docent at Jasper Ridge Biological Preserve	2016 - present
Co-Instructor for Stanford Splash! M4053: For the Love of Optimization	2014
Co-Instructor for Stanford Splash! B4329: The Biology of Vision and Perception	2015
Organizer for Stanford SAILORS (now Stanford AI4ALL)	2014

honors & awards

NSF Graduate Research Fellowship	2019
SICB Marlene Zuk Best Student Paper Finalist	2019
UC Berkeley Allen D. Wilson Fellowship (*declined*)	2017
UC Berkeley EECS Excellence Award (*declined*)	2017
Tau Beta Pi National Honor Society for Engineering	2016
Stanford University President's Award for Excellence in the Freshman Year	2013