Kelly W. Zhang

kellywzhang@seas.harvard.edu
https://kellywzhang.github.io/

Education

Harvard University, School of Engineering and Applied Sciences — Cambridge MA
PhD in Computer Science — September 2018 - May 2023 (expected)
Advisors: Susan A. Murphy and Lucas Janson

New York University, College of Arts and Sciences — New York, NY
BA in Computer Science; *summa cum laude* — May 2018
Advisors: Sam Bowman and Yann LeCun

Research Interests

My research interests lie at the intersection of *adaptive experimentation*, *reinforcement learning*, and *statistical inference*.

Work Experience

Apple — Seattle, WA
HealthAI Research Intern — May 2022 - August 2022
Developed statistical inference methods specific to mobile health problems in large-scale, industry settings.

Facebook AI Research — New York, NY
Research Intern — May 2018 - August 2018
Worked on sparse coding and text generation with Yann LeCun.

eBay — New York, NY
Software Engineering Intern on Recommendations Team — May 2017 - August 2017
Worked on detecting "Not Suitable for Work" content, like nudity, in products sold on eBay.

Honors

Siebel Scholar, Class of 2023 *($35,000 award; given to 100 final year PhD candidates in engineering)*

Institute of Mathematical Statistics Hannan Graduate Student Travel Award, 2022 *($800 award)*

Certificate of Distinction in Teaching awarded by the Harvard Office of Undergraduate Education, based on reviews of my performance as a Teaching Fellow for Susan Murphy's course on *Sequential Decision Making*

National Science Foundation GFRP Fellowship, awarded in 2019 *($147,000 award)*

Computer Science Prize for Academic Excellence in the Honors Program, New York University, 2018

Publications

Paper titles are linked to the online pdf of the papers. * Denotes denotes equal contribution.

Statistical Inference after Adaptive Sampling

Statistical Inference After Adaptive Sampling for Longitudinal Data.
Kelly W. Zhang, Lucas Janson, and Susan A. Murphy.
To be submitted to the Annals of Statistics.

Statistical Inference with M-Estimators on Adaptively Collected Data.
Kelly W. Zhang, Lucas Janson, and Susan A. Murphy.
35th Conference on Neural Information Processing Systems (NeurIPS 2021).

Inference for Batched Bandits.
Kelly W. Zhang, Lucas Janson, and Susan A. Murphy.
34th Conference on Neural Information Processing Systems (NeurIPS 2020).

Designing Reinforcement Learning Algorithms

Reward Design For An Online Reinforcement Learning Algorithm Supporting Oral Self-Care.
Anna L. Trella, Kelly W. Zhang, Inbal Nahum-Shani, Vivek Shetty, Finale Doshi-Velez, Susan A. Murphy.
Thirty-Fifth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-23), 2023.

Designing Reinforcement Learning Algorithms for Digital Interventions: Pre-implementation Guidelines.
Anna L. Trella, Kelly W. Zhang, Inbal Nahum-Shani, Vivek Shetty, Finale Doshi-Velez, Susan A. Murphy.
Algorithms (Special Issue "Algorithms in Decision Support Systems" Vol. 2);
Preliminary version presented at the *5th Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM 2022);* ***(selected for an oral presentation)***.

A Bayesian Approach to Learning Bandit Structure in Markov Decision Processes.
Kelly W. Zhang, Omer Gottesman, and Finale Doshi-Velez.
Challenges of Real-World Reinforcement Learning (NeurIPS 2020 Workshop).

Natural Language Processing

Language Modeling Teaches You More Syntax than Translation Does: Lessons Learned Through Auxiliary Task Analysis.
Kelly W. Zhang and Samuel R. Bowman.
BlackboxNLP 2018 (Workshop at Conference on Empirical Methods in Natural Language Processing).

Adversarially Regularized Autoencoders.
Jake (Junbo) Zhao, Yoon Kim, Kelly Zhang, Alexander M. Rush, and Yann LeCun.
Thirty-fifth International Conference on Machine Learning (ICML 2018).

Additional Forthcoming Work

Optimizing an adaptive Digital Oral Health Intervention for promoting Oral Self Care Behaviors: Micro-Randomized Trial Protocol.
Inbal Nahum-Shani, Zara M. Greer, Anna L. Trella, Kelly W. Zhang, Stephanie Carpenter, David Elashoff, Susan A. Murphy, and Vivek Shetty.
Preparing for submission.

Invited Talks

Statistical Inference for Longitudinal Data After Adaptive Sampling

Modern Experimental Design and Active Learning in the Real World, Online seminar organized by Ilija Bogunovic, Willie Neiswanger, and Mojmir Mutny, January 2022

University of Copenhagen, *Center for Social Data Science*, January 2023

INFORMS Annual Meeting 2022, Session on *Efficient Learning via Adaptive Experimentation* organized by Daniel Russo, October 2022

University of Amsterdam, *Statistics Department Seminar*, September 2022

Joint Statistical Meeting 2022, Session on *Prediction and Inference in Statistical Machine Learning* organized by Tracy Ke, August 2022

Institute of Mathematical Sciences Annual Meeting 2022, Session on *Inference Methods for Adaptively Collected Data* as a speaker and the session organizer, June 2022

University of Toronto, *Intelligent Adaptive Interventions Lab*, June 2022

Apple Machine Learning Research, *Health AI Team*, June 2022

Statistical Inference with M-Estimators on Bandit Data

École Polytechnique Fédérale de Lausanne (EPFL), *Statistics Seminar*, December 2021 (`https://tube.switch.ch/videos/pCYwBwSYh6`)

Pennsylvania State University, *Statistics Department Colloquium*, November 2021

INFORMS Annual Meeting 2021, Session on *Advances in Causal Inference and Reinforcement Learning for the Online Service Industry* organized by Tim Keaton

London School of Hygiene and Tropical Medicine, *Health Data Science Seminar Series*, June 2021

Cambridge University, *Medical Research Council Biostatistics Unit Seminar*, April 2021 (`https://www.youtube.com/watch?v=jflR7KOrqNA&t=2s`)

Oralytics: A Mobile Health Study to Improve Oral Health Behaviors

Pennsylvania State University, *QuantDev Methodology Brown-Bag Seminar*, November 2021

Inference for Batched Bandits

Bernoulli World One Symposium, Session on *Statistical Methods in Machine Learning*, August 2020

Teaching

Guest Lecturer, **Harvard University**. Susan Murphy's course on *Sequential Decision Making*, STAT 234, Spring 2022.

Teaching Fellow, **Harvard University**. Susan Murphy's course on *Sequential Decision Making*, STAT 234, Spring 2021.

Grader, **New York University**. Sam Bowman and Kyunghyun Cho's course on *Natural Language Processing with Representation Learning*, DS-GA 1011, Fall 2017.

Service

Workshops and Sessions

Organized Invited Session at 2022 Institute of Mathematical Sciences Annual Meeting on *Inference Methods for Adaptively Collected Data.* The speakers were Nathan Kallus, Koulik Khamaru, Evan Munro, and myself. Joseph Jay Williams and Nina Deliu chaired the session.

Assisted in Organizing the Harvard Radcliffe 2022 Exploratory Seminar on *Ethical Considerations in the Use of Big Data, AI, and Real-Time Information for Prediction of Behavioral Health Outcomes*, which was led by Jordan Smoller and Matthew Nock.

Co-Organized NeurIPS 2021 Workshop on *Causal Inference Challenges in Sequential Decision Making: Bridging Theory and Practice* with Aurelien Bibaut, Maria Dimakopoulou, Nathan Kallus, Xinkun Nie, and Masatoshi Uehara.

Co-Organized NeurIPS 2020 Workshop on *Machine Learning for Mobile Health* with Walter Dempsey, Nick Foti, Joseph Futoma, Yian Ma, Marianne Njifon, and Jieru Shi.

Assisted in a practical workshop on Online Learning and Experimentation Algorithms in Mobile Health organized by Walter Dempsey as a part of the AI4Health Winter School 2021.

Reviewing

AISTATS 2023

NeurIPS 2020, 2021

Conference on Causal Learning and Reasoning 2022

NeurIPS 2020 Workshop on Challenges of Real World Reinforcement Learning

NeurIPS 2022 Workshop on Causal Machine Learning for Real-World Impact

Mentoring

Harvard Women in STEM Mentorship Program, Mentor for undergraduate students, 2020-2022.

Summer Institute on Just-in-Time Adaptive Interventions via Micro-Randomized Trials run by the d3center at UMichigan, Mentor, 2021

Undergraduate Student Research Projects

- I mentored an undergraduate student from the University of Science and Technology of China, Zeyang Jia, for the summer and fall of 2020 on *Weighting Methods to Maximize the Power of Hypothesis Tests on Bandit Data*. He presented this work at the Session on Statistical methods in Machine Learning at the Bernoulli-IMS World One Symposium 2020. He is now a PhD student in statistics at Harvard.
- I mentored a Harvard undergraduate student, Raymond Feng, for the winter and summer of 2020 on *Predicting User Engagement in Large-scale Mobile Health Studies*.