Kai Yuanqing Xiao

32 Vassar Street, G628 ● Cambridge, MA 02139 ● +1.408.828.9796 ● kaix@mit.edu ● https://kaixiao.github.io

EDUCATION

Massachusetts Institute of Technology – Computer Science and Artificial Intelligence Lab — Cambridge, MA
Pursuing a Ph.D. in Computer Science, with a focus on Robustness and Reliability in Machine Learning; GPA: 5.0/5.0 — 2017-Present
Advisor: Aleksander Madry

Massachusetts Institute of Technology — Cambridge, MA
M.Eng. Degree – Master's Thesis on "Cookie Clicker" under the guidance of Erik Demaine; GPA: 5.0/5.0 — 2017-2018
B.S. Degree – Double Major in Computer Science and Mathematics; GPA: 5.0/5.0 — 2013-2017
Coursework: 6.854 (Advanced Algorithms), STAT 210 (Probability Theory), 6.438 (Algorithms for Inference), 6.869 (Computer Vision), 6.840 (Complexity Theory), 6.172 (Performance Engineering)

Oxford University — Oxford, UK
Visiting Student in Mathematics at St. Peter's College — Jan.-June 2016
Coursework: Machine Learning, Networks

RESEARCH

"3DB: A Framework for Debugging Computer Vision Models" 2021
Guillaume Leclerc, Hadi Salman, Andrew Ilyas, Sai Vemprala, Logan Engstrom, Vibhav Vineet, **Kai Xiao**, Pengchuan Zhang, Shibani Santurkar, Greg Yang, Ashish Kapoor, Aleksander Madry. (https://arxiv.org/abs/2106.03805, https://github.com/3db/3db)
- Analyzed the transferability of insights from 3DB to real-world settings
- Tested codebase extensively and wrote documentation describing how to use it

"Noise or Signal: The Role of Image Backgrounds in Object Recognition" 2020
Kai Xiao, Logan Engstrom, Andrew Ilyas, Aleksander Madry. (https://arxiv.org/abs/2006.09994)
Proceedings of the International Conference on Learning Representations (ICLR), 2021.
- Created new toolkit and datasets for investigating the effects of image backgrounds on object recognition models
- Performed extensive evaluation of modern computer vision models' reliance on backgrounds

"Toward Evaluating Robustness of Deep Reinforcement Learning with Continuous Control" 2019
Tsui-Wei Weng, Krisnamurthy (Dj) Dvijotham, Jonathan Uesato, **Kai Xiao**, Sven Gowal, Robert Stanforth, Pushmeet Kohli. (https://openreview.net/forum?id=SylL0krYPS)
Proceedings of the International Conference on Learning Representations (ICLR), 2020.
- Trained dynamics models of various MuJoCo environments
- Helped write code for optimizing attacks against agents

"A Framework for Robustness Certification of Smoothed Classifiers using f-divergences" 2019
Krisnamurthy (Dj) Dvijotham, Jamie Hayes, Borja Balle, Zico Kolter, Chongli Qin, Andras Gyorgy, **Kai Xiao**, Sven Gowal, Pushmeet Kohli. (https://openreview.net/forum?id=SJlKrkSFPH)
Proceedings of the International Conference on Learning Representations (ICLR), 2020.
- Helped proofread and discuss the final results

"Data-Driven Robust Reinforcement Learning for Continuous Control" 2019
Yuanyuan Shi, **Kai Xiao**, Daniel J. Mankowitz, Rae Jeong, Nir Levine, Sven Gowal, Timothy Mann, Todd Hester. (https://sites.google.com/view/neurips19-safe-robust-workshop)
NeurIPS workshop on Safety and Robustness in Decision Making, 2019.
- Trained dynamics models of various MuJoCo environments

"Learning Neural Dynamics Simulators with Adversarial Specification Training" 2019
Kai Xiao, Sven Gowal, Todd Hester, Rae Jeong, Daniel J. Mankowitz, Yuanyuan Shi, Tsui-Wei Weng. (https://sites.google.com/view/neurips19-safe-robust-workshop)
NeurIPS workshop on Safety and Robustness in Decision Making, 2019.
- Used MuJoCo simulators to train dynamics simulators
- Incorporated physics-based specifications during training via adversarial robustness techniques

"Training for Faster Adversarial Robustness Verification via Inducing ReLU Stability" 2018
Kai Xiao, Vincent Tjeng, Nur Muhammad (Mahi) Shafiullah, Aleksander Madry. (https://arxiv.org/abs/1809.03008)
Proceedings of the International Conference on Learning Representations (ICLR), 2019.
- Explored co-designing neural networks to be both robust and easily verifiable
- Developed regularization technique for encouraging ReLU Stability, allowing for faster verification

"Evaluating Robustness of Neural Networks with Mixed Integer Programming" 2018
Vincent Tjeng, **Kai Xiao**, Russ Tedrake. (https://arxiv.org/abs/1711.07356)
Proceedings of the International Conference on Learning Representations (ICLR), 2019.

- Supported by providing adversarial-training baselines for evaluations of robustness

"Cookie Clicker" - Master's Thesis 2018
Erik Demaine, Hiro Ito, Stefan Langerman, Jayson Lynch, Mikhail Rudoy, **Kai Xiao**. (https://arxiv.org/abs/1808.07540)
Oral Presentation at the 20th Japan Conference on Discrete and Computational Geometry, Graphs, and Games.
- Analyzed optimal strategies for incremental games like Cookie Clicker
- Discovered NP-Hardness results, dynamic programming solutions, and approximation algorithms

Neural Connectivities Analysis (with Shafrira Goldwasser) 2016
- Analyzed neural connectivities dataset using spectral clustering and community graph model

"Online Algorithms Modeled after Mousehunt" - Final Project for 6.854 (Advanced Algorithms) 2014
Jeffrey Ling, **Kai Xiao**, Dai Yang. (https://arxiv.org/abs/1501.01720)
- Studied Markov Decision Processes, randomized online algorithms, and competitive ratios applied to the game

AWARDS

- NDSEG Fellowship Program Award 2019
- NSF Graduate Research Fellowship Program (GRFP) Award 2018
- Top 200 in William Lowell Putnam Mathematical Competition 2014
- Qualified 4 times for USA Math Olympiad; Honorable Mention (top 24 out of over 100,000) in 2012, top 50 in 2011 2010-2013

WORK EXPERIENCES

Teaching Assistant for 6.883 (Data-Driven Decision Making and Society) at MIT Cambridge, MA
- Helping organize logistics and content for first iteration of this course. Spring 2021

Microsoft Research Redmond, WA
Summer Research Intern Summer 2020
- Worked in the Machine Learning Optimization team on a project involving adversarial patches

Google DeepMind London, UK
Summer Research Intern Summer 2019
- Worked with the DeepMind for Google team and Deep Learning Robustness team
- Studied how adversarial robustness could be applied to learning simulators for reinforcement learning

Teaching Assistant for 6.046 (Design and Analysis of Algorithms) at MIT Cambridge, MA
- Taught weekly classes, held twice-a-week office hours, wrote problem set and exam questions for two academic semesters 2016-2017

Citadel Chicago, IL
Summer Quantitative Research Analyst Summer 2016
- Used text mining and sentiment analysis on a unique dataset to construct predictive signal for stock prices
- Improved the data processing pipeline and evaluated changes using characteristic portfolios and simulations

D.E. Shaw & Co. New York City, NY
Quantitative Analyst / Software Development Intern Summer 2015
- Created mathematical models for the behavior of specific types of trades based on market conditions
- Used vectorized operations in NumPy to analyze large amounts of historical data

A9 (Product Search Team) Palo Alto, CA
Software Development Engineer Intern Summer 2014
- Worked with Apache Hadoop and Apache Pig to perform map-reduce tasks
- Generated and logged statistical metrics related to Amazon's product search rankings
- Mined Twitter data for trending music and showed related items available on Amazon (side project)

Jane Street Capital New York City, NY
Assistant Trader January 2014
- Modeled stock market behavior through analysis of historical and recent financial data

Stanford University Chemistry Department; Bianxiao Cui, Ph.D . Palo Alto, CA
Data Analysis Intern July-Aug. 2012
- Processed images of protein movement across axons; traced curves in images using MATLAB program
- Improved functionality of MATLAB curve-tracing program after learning the language from scratch

LEADERSHIP EXPERIENCE

MIT TechX
Director of Corporate Relations
2014-2015
- Leader of student group that communicated with companies to sponsor and exhibit their technologies at MIT's annual xFair
- Worked with other executive board members to run events that expose MIT students to interesting technology