

Yixiao HUANG

yixiaoh@berkeley.edu | homepage | Last updated on September 18, 2025

EDUCATION

University of California, Berkeley

Ph.D in Computer Science Advisor: Prof. Somayeh Sojoudi

Sept, 2024 - now

City University of Hong Kong

B.E. in Information Engineering CGPA: 3.98 / 4.3 (rank 1/91)

Sept, 2019 – Jun, 2023

SELECTED PUBLICATIONS

8. Yixiao Huang*, Hanlin Zhu*, Tianyu Guo*, Jiantao Jiao, Somayeh Sojoudi, Michael I. Jordan, Stuart Russell, Song Mei, “**Generalization or Hallucination? Understanding Out-of-Context Reasoning in Transformers**”. The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS), 2025. [pdf]

7. Ziheng Cheng*, Yixiao Huang*, Hui Xu, Somayeh Sojoudi, Xuandong Zhao, Dawn Song, and Song Mei, “**OVERT: A Benchmark for Over-Refusal Evaluation on Text-to-Image Model**”. The Thirty-ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS D&B), 2025. [pdf]

6. Xuyang Zhong, Yixiao Huang, Chen Liu, “**Understanding and Improving Fast Adversarial Training against ℓ_0 Bounded Perturbations**”. The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS), 2025. [pdf]

5. Mingchen Li, Xuechen Zhang, Yixiao Huang, and Samet Oymak. “**On the Power of Convolution Augmented Transformer.**”. The 39th Annual AAAI Conference on Artificial Intelligence (oral presentation), 2025. [pdf]

4. Muhammed Emrullah Ildiz, Yixiao Huang, Yingcong Li, Ankit Singh Rawat, and Samet Oymak. “**From Self-Attention to Markov models: Unveiling the Dynamics of Generative Transformers.**”. The 41st International Conference on Machine Learning (ICML), 2024. [pdf]

3. Xuyang Zhong, Yixiao Huang, Chen Liu, “**Towards Efficient Training and Evaluation of Robust Models against ℓ_0 Bounded Adversarial Perturbations**”. The 41st International Conference on Machine Learning (ICML), 2024. [pdf]

2. Yingcong Li*, Yixiao Huang*, Muhammed Emrullah Ildiz, Ankit Singh Rawat, Samet Oymak, “**Mechanics of Next Token Prediction with Self-Attention**”. The 27th International Conference on Artificial Intelligence and Statistics (AISTATS), 2024. (* equal contribution) [pdf]

1. Yixiao Huang, Xiaoli WU, Rosa H. M. Chan and Pooled Resource Open-Access ALS Clinical Trials Consortium, “**Stratification and Survival Prediction for Amyotrophic Lateral Sclerosis Patients**”. IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), 2022, pp. 1-5, doi: 10.1109/BHI56158.2022.9926946. [pdf]

ACADEMIC EXPERIENCE

Reviewer: AISTATS 2024,2025, ICLR 2025; TMLR

HONORS AND AWARDS

2024 ICML Travel Award

Jun, 2024

The College of Engineering Dean’s Scholarships

Maximum of 5 recipients in the College of Engineering every year

Aug, 2022

The 46th International Collegiate Programming Contest - Asia Kunming Regional

Silver Medal (Rank 40th/697)

April, 2022

Chi-Li Pao Foundation Scholarships for Academic Outstanding Students

Maximum of 2 recipients in the College of Engineering every year

Aug, 2021

Teaching

EE127/227AT - Optimization Models in Engineering (UC Berkeley)

Graduate Student Instructor

Fall, 2025

EE2331 - Data Structures and Algorithms (CityU)

In-class tutoring

Fall, 2022

MA1201 - Calculus and Basic Linear Algebra II (CityU)

After-class tutoring

Fall, 2020