Hello, I'm

Yingqi Gao

yingqi.gao.stats@gmail.com
linkedin.com/in/yingqi-gao
yingqi-gao.github.io

PhD candidate in statistics researching machine learning and economics, building real-world, data-driven solutions that help corporations maximize profits by transforming unstructured data and complex insights into strategic decisions. I develop market-driven mechanisms for digital economies, spanning data, privacy, and AI derivatives; and design AI systems that adapt to economic incentives, including LLM-powered ad auctions and uncertainty-aware chatbots.

Education

University of California, Los Angeles - *PhD in Statistics & Data Science*   SEP 2021 - JUN 2026

Columbia University in the City of New York - *MA in Statistics in Statistics*   SEP 2019 - DEC 2020

University of California, San Diego - *BS in Probability & Statistics and BS in Management Science*   SEP 2015 - JUN 2019

Experience

PhD Researcher - *UC Los Angeles*   FEB 2023 - PRESENT

- Designed a pioneering two-phase pricing mechanism that doubles the speed of approaching oracle revenue, maximizing customer retention and dataset sales with nonparametric value estimation and optimization in SciPy.
- Expanded a first-of-its-kind online adaptive pricing algorithm, converting Qiu et al. (2022)'s repeated density estimation method from R to Python via rpy2, achieving fast reget convergence competitive with state-of-the-art online learning mechanisms under comparable but varied settings.
- Architecting a decentralized privacy marketplace where an LLM empowers users with granular controls, turning forced surrender into incentive-aligned, sustainable, and profitable exchanges.
- Enhancing AI chatbots, especially in healthcare AI, with uncertainty quantification, enabling statistical confidence assessment to deliver more reliable and trustworthy recommendations.

Machine Learning Researcher - *UC Los Angeles*   SEP 2022 - DEC 2022

- Engineered a more flexible, accurate, efficient, and scalable zero-shot generalization framework for object-centric reinforcement learning, building upon Veerapaneni et al. (2020)'s OP3 model in PyTorch and achieving high success rates in block stacking and rearrangement tasks.
- Programmed a unified action interface with variational encoding, enabling aggregate action modeling and attention-based decomposition for more interpretable and adaptive multi-object control.
- Converted high-level research concepts into functional code, collaborated with researchers to extend the model to a more complex billboard ball (BiB) game environment in Gym, and delivered a report, presentation, and live demo.

Graduate Researcher - *Columbia University*   JUN 2020 - JUN 2021

- Optimized a Bayesian hierarchical sparse VAR model for multi-subject, multi-session fMRI analysis, cutting days of Gibbs sampling runtime in R with HPC, while assessing and strengthening model stability via sensitivity analysis.

Research Assistant - *Columbia Business School*   DEC 2019 - MAR 2020

- Developed a patent identification algorithm, outperforming NBER's method with higher precision and refined detection of academic affiliations, via data analysis, refined regular expressions, and optimized extraction in SAS.

Lead Data Analyst - *UC San Diego*   MAR 2018 - JUN 2018

- Led a team of six in a fast-paced statistical project course, driving data analysis across five real-world topics in R, guiding model development, decision-making, collaboration, and report writing, earning a top-tier evaluation.

Skills

Programming & Tools: Python (PyTorch, TensorFlow, SciPy, rpy2), R, SAS, SQL, HPC, distributed processing

Economics: game theory, causal inference, A/B testing, mechanism design, auctions, econometrics