# Yingqi Gao


### Quantitative & ML Researcher – AI Systems & Economics

*yingqi.gao.stats@gmail.com | 858-666-7501 | yingqi-gao.github.io | linkedin.com/in/yingqi-gao*


PhD-trained researcher in **quantitative** modeling and **machine learning**. Develops **interpretable**, **data-efficient** algorithms for **decision-making under uncertainty** and studies **profit-aware**, **incentive-aligned** mechanisms for **AI markets** that coordinate the pricing, exchange, and utilization of **data**, **ML models**, and **APIs** (such as **LLM** services). I aim to design **next-generation AI systems** where autonomous **agents** learn and reason through **causal and economic principles** to achieve **stable**, **adaptive**, and **intelligent** collective behavior.

## Work Experience

### PhD Researcher – Market Design & AI Economics
Feb 2023 - Present

*University of California, Los Angeles (UCLA) | Los Angeles, CA*

- Designed a **learning**-based **auction**-to-posted-price mechanism for **AI-driven digital markets** trading **data**, **ML models**, and **LLM** services, combining **SciPy** optimization with **R**-based nonparametric estimation for **data-efficient**, **profit-driven**, **no-regret** pricing that reaches optimal revenue **2× faster**.
- First-authored *Learn then Decide: A Learning Approach for Designing Data Marketplaces* (major revision at **JASA**, Theory & Methods), establishing the foundation for future **LLM-based data, privacy, and ad markets** driven by **incentive-aligned economic design**.
- Built a **PyTorch** framework for prediction-powered **uncertainty quantification** in **high-dimensional causal mediation**, benchmarked against **XGBoost** and **DoubleML**, and scaled on **Google Cloud**.

### Machine Learning Researcher – Reinforcement Learning Systems
Sep 2022 - Dec 2022

*University of California, Los Angeles (UCLA) | Los Angeles, CA*

- Engineered a **PyTorch** framework extending OP3 for object-centric **reinforcement learning**, using **variational encoding** and **attention-based decomposition** to enable **interpretable**, **modular decision-making** in **dynamic**, **multi-agent** systems.

### Graduate Researcher – Bayesian Inference & Scalable Modeling
Jun 2020 - Jun 2021

*Columbia University | New York, NY*

- Optimized a **Bayesian hierarchical sparse VAR model** for multi-subject, multi-session **fMRI** data, cutting **runtime by 80% (15 to 3 days)** via **HPC** migration and **C** integration, enabling **efficient uncertainty-aware inference** and **stable estimation of cross-system dependencies**.

## Skills

**Programming & Tools:** Python (PyTorch, TensorFlow, SciPy, scikit-learn, XGBoost, NumPy, pandas, Seaborn), R, SQL, Git, Google Cloud, HPC.

**Core Competencies**: Machine Learning, Deep Learning, Reinforcement Learning, Recommender Systems, Optimization, Mechanism Design, Causal Inference, Game Theory, Uncertainty Quantification.

## Education

### University of California, Los Angeles (UCLA)
Sep 2021 - Dec 2026 (expected)

**PhD** in Statistics & Data Science

### Columbia University
Sep 2019 - Dec 2020

**MA** in Statistics – GPA: 4.0/4.0

### University of California, San Diego (UCSD)
Sep 2015 - Jun 2019

**BS** in Probability & Statistics; **BS** in Management Science, Graduated with High Honors – GPA: 3.9/4.0