

Richard Zhuang

EDUCATION

University of California, Berkeley

Expected Graduation: May. 2025

Double Major in Computer Science (Honor) and Applied Mathematics

GPA: **4.0/4.0**

Relevant Coursework: Machine Learning, Deep Learning, Foundations of Large Language Models (LLMs), Statistical Learning Theory, Probability Theory and Stochastic Processes, Techniques of Data Science, Computer Algorithms.

Honors and Awards: Dean's List - Fall 2022, Spring 2023, Fall 2023.

ACADEMIC PROJECTS

EmbedLLM: Learning Compact Representations of LLMs

Jan. 2024 - Nov. 2024

Advised by Prof. Jiantao Jiao (**First-Authored, ICLR 2025 Spotlight**)

- Designed and implemented a **encoder-decoder based reconstruction system** to extract model embeddings that simultaneously facilitates model correctness forecast, model routing, and benchmark accuracy predictions.
- Conducted probing experiments to analyze information contained in the learned embeddings.

PokerBench: Training LLMs to become Professional Poker Players

Jan. 2024 - Sep. 2024

Advised by Prof. Gopala Anumanchipalli (**First-Authored, AAAI 2025**)

- Established a **benchmarking dataset** with over 20K test and 560K training examples of game scenarios utilizing optimal decision labels verified by poker strategy solvers.
- Performed **supervised fine-tuning** on Llama-3-8B, achieving a 52% improvement in test set accuracy.

Evolving AI Collectives to Enhance Human Diversity and Enable Self-Regulation

Sep. 2023 - Feb. 2024

Advised by Prof. Dawn Song (**Co-Authored, ICML 2024**)

- Conducted prompt engineering and performed simulations for the Prisoner's Dilemma and story relay games between multiple LLM agents.
- Analyzed game results through data visualization to discern behavior disparities between multi-agent and single-agent systems.

EXPERIENCES

Bespoke Labs (bespokelabs.ai)

Jan. 2025 - May. 2025

Research Intern

- Conducting research in developing function calling agents and reasoning models through reinforcement learning and synthetic data curation.

Delphi Group, Carnegie Mellon University

May. 2024 - Sep. 2024

Software Development Engineer Intern

- Contributed to real-time epidemiology forecasting tools by developing and alpha-testing R packages.
- Enhanced package documentations by establishing tutorial notebooks on time-series modeling using **ARIMA**, **ETS**, and **Fourier Decomposition methods**.

Mindscape Lab, University of California, San Francisco

Sep. 2023 - Sep. 2024

Machine Learning Researcher Intern

- Implemented a **Bayesian hierarchical generative model** of patient movement patterns to aid the development of a disease transmission simulation system.
- Adapted image reconstruction algorithms like **Masked Autoencoder** and **Vision Transformer** to recover omniscient transmission pattern from partial observations.

SKILLS

- Languages: Python, Java, R, SQL, MATLAB, JavaScript, Lean
- Tools & Packages: PyTorch, Transformers, HuggingFace, vLLM, LitGPT, Ray, NumPy, Pandas, Matplotlib, AWS