## Ian Joffe

# Data Scientist and Machine Learning Engineer

ianjoffe.github.io ianjoffe@uchicago.edu

EDUCATION (323) 434-3245

## University of Chicago | M.S. in Statistics | September 2023 – May 2025 (Present)

Leads M.S. class in total coursework completed. Also planning to complete course requirements for Applied Math M.S.

Completed Ph.D.-level sequences in theoretical statistics, applied statistics, and probability

## University of California, Berkeley | B.S. in Electrical Engineering and Computer Science | Graduated 2023

Minors in Mechanical Engineering and Data Science

## **EXPERIENCE**

### Data Science & ML Researcher at UChicago (planning to submit projects for publication this year)

Spring 2024 - Present

- Quantifying the role of sharpness on generalization in width-varying two-layer neural networks using theoretical bounds and experiments | Repo
- Solving for the dynamic interaction between statistical predictive models and markets betting on the same outcome
- Strategies for improving the calibration of LLMs on forecasting current events, including debiasing of latent representations and incorporation of proper scoring, supervised by Aaron Schein | Repo

### Data Science Intern at Visimo.ai

Summer 2023

- Deployed YOLO model for video object segmentation in internal R&D project, results were demoed to potential clients
- Developed and implemented novel application of self-supervised training for contracted deep learning project
- Performed comprehensive publications review on SOTA methods for LLM fine-tuning in jargon-heavy domains

## Data Science & Analytics Intern at TransUnion – Model Governance Team

Summer 2022

- Built data analytics dashboard to monitor millions of model outputs and automatically detect distribution shifts
- Researched historic distribution shifts to contextualize them in macroeconomic trends and assess significance to the company
- Used Spark to write an API to help 100+ coworkers access and alter databases

## Data Science & Cognitive Science Teaching Assistant at UChicago

Fall 2024 & Winter 2025

- Data 27100: Capstone project course for data science majors. Managed teams working on a physics emulator for Fermilab and a RAG chatbot for an environmental nonprofit
- CogSci 20002: Taught coding and mathematical modeling in the context of neuroscience, psychology, and linguistics

### **LEADERSHIP & ORGANIZATIONS**

## President, Project Manager at Political Computer Science

2019-2023

- As president oversaw 11 projects, new member recruitment, and social events. Increased member retention by over 50%
- PM for university-wide scientific poll to gather and dissect data on student sentiment. Also investigated use of social media ads by the 2020 presidential candidates, created NLP model to detect polarized speech

## **Project Manager at Sports Analytics Group at Berkeley**

2019-2023

- PM for team using ML to develop new metrics and write reports offering strategy advice for Berkeley's baseball team
- Worked with the Denver Nuggets to quantify "brand health" through social media activity, other market data

#### **Engineers Without Borders**

2019-2023

Compiled sanitation and health data for reports to national office for a team that implemented water pipelines in Panama

### PROJECTS, SKILLS, & MISC.

- Languages: Advanced in Python, also strong in R, Java, C, SQL | Tools: Tableau, Solidworks, Git, Slurm
  Libraries: PyTorch, Jax, Pandas, Numpy, SciKit-Learn, Matplotlib, Plotly & Dash, Shiny, Weights & Biases, PyMC3, Penzai, SciPy, LightGBM & xgboost, beautifulsoup
- Open Source: Submitted PR to X/Twitter Community Notes repo lowering model loss by 10% and catching bug
- **Graduate Course Projects**: Literature review and extension on info theory for predictions markets; Review and implementation on the loss landscape complexity of phase retrieval; A model to estimate scaling laws from compute market data; Review on dynamical systems for stable RNN training
- **Personal Projects:** Programming, simulating, and solving the card-based word game Quiddler; Experimenting with algorithmic stock trading; 3D Printing objects for everyday use at home
- Private math tutor for elementary, middle, and high school students
- Attended NeurIPS, ACM Economics & Computation, and SIAM Mathematics of Data Science conferences
- Counselor at two-month overnight summer camp Hess Kramer (Summers 2019 and 2021)