Kevin Liu

Los Angeles, CA | kliuc@ucla.edu | (925) 577-9710 | linkedin.com/in/kliuc | github.com/kliuc

Education

University of California Los Angeles (UCLA) Expected June 2025
BS Double Major in Mathematics, Statistics & Data Science

- **GPA: 3.95/4.00**
- **Math Machine Learning Research Group**: Studying deep learning theory, advised by Prof. Guido Montúfar.
- Learning Assistant Program: Supported underclassmen with study of multivariate calculus.
- Coursework: ML Theory, Deep Learning Theory, Statistical Analysis, Linear Algebra, Optimization, Probability

Experience

Student Researcher June 2024 – Present
UCLA Mathematics – Los Angeles, CA

- Performed theoretical analyses and empirical experiments on topics in deep learning theory, exploring phenomena such as *benign overfitting* in ReLU networks and *grokking* of algorithmically generated datasets.
- Developed neural networks, training methods, and data generation procedures in PyTorch, evaluating existing research and exploring novel hypotheses to deepen theoretical understanding of deep learning phenomena.

Data Science Intern Mar 2024 – May 2024
Timeplus – Remote

- Designed and implemented an end-to-end DDoS detection pipeline using Python, integrating data ingestion, streaming processing with Timeplus, GPT-3.5-based pattern analysis, and real-time alerting.
  Read: timeplus.com/post/real-time-ddos-detection
- Utilized few-shot prompting to train the LLM on DDoS traffic patterns, achieving 90% testing accuracy on the CIC-IDS2017 intrusion detection dataset, a 5% improvement over traditional neural networks.

Data Analyst Intern June 2023 – Sept 2023
Intellipro Group – Santa Clara, CA

- Designed and implemented scalable solutions for recruiting database enrichment, process automation, and API integration using Python, MongoDB, and SQL, enhancing data workflows.
- Boosted MongoDB query performance by 10x through optimized schema design and indexing strategies.

Projects

UFC Fighter Rating Engine

- Designed and implemented a *Glicko-2-based* rating system for UFC fighters using Python, optimizing algorithm parameters to minimize predictive discrepancy, achieving 68% backtesting accuracy on historical fight outcomes.
- Built an interactive Streamlit interface with real-time visualizations, including fighter rating trends, current rankings, and peak historical ratings, enabling seamless data analysis and comparison.
- Tools Used: Python (pandas, NumPy, scikit-learn, Streamlit)

Sports Betting Arbitrage Tool

- Developed an automated sports betting odds scraping and data ingestion pipeline using Python and SQLite, collecting over 15,000 odds across hundreds of events and thousands of athletes in real-time.
- Designed queries to identify arbitrage opportunities, achieving a 15% risk-free monthly return with this strategy.
- Tools Used: Python (Requests, Beautiful Soup, Selenium, SQLAlchemy, Streamlit), SQLite

Skills

Languages: Python (pandas, NumPy, PyTorch, scikit-learn, Matplotlib, SQLAlchemy), R (dplyr, ggplot2), SQL
Technologies: Jupyter Notebook, Colab, SQLite, MongoDB, Git, AWS, LaTeX