

Ryan Zheng

925-201-9539 | ryanzheng@nmokey.com | linkedin.com/in/nmokey | github.com/nmokey

EDUCATION

University of California, Los Angeles

B.S. in Physics, Data Science Engineering Minor

Los Angeles, CA

Sep. 2024 – June 2027

- **Relevant Coursework:** Data Structures & Algorithms, Computer Architecture, Multivariable Calculus, Linear Algebra, Differential Equations, Modern Physics
- **GPA:** 3.879
- **Languages:** C++, Java, Python, Swift, SQL, HTML/CSS, R
- **Frameworks:** PyTorch, Hugging Face, React, Jekyll
- **Developer Tools:** Git, Cursor, VS Code, XCode, Jupyter

EXPERIENCE

Generative AI Intern

Scale AI

January 2025 – Present

San Francisco, CA (Remote)

- Contribute to training and evals for SOTA reasoning and agentic models across 5+ clients and 100+ tasks, ranging from abstract visual reasoning, to deep research studies, to next-gen SWE agents solving real world GitHub issues.
- Execute rigorous quality assurance reviews for critical datasets powering SWE agents in Java, C++, Python, Go, and Rust, evaluating intern deliverables against customer specs to ensure data integrity. Improve data quality rating by organizing and leading in-person project workshops for contributing interns.
- Construct robust Docker testing environments and write comprehensive rubrics for industry standard AI benchmarks such as the Aider LLM Leaderboards.

Research Assistant

June 2022 – July 2022

University of Texas at Austin

Austin, TX

- Developed data aggregation tool using Java to process national travel datasets for research study of the conterminous United States' travel patterns during the COVID-19 pandemic. Optimized aggregation algorithm to reduce runtime to less than 50%.
- Coauthored research paper accepted to *2024 COTA International Conference of Transportation Professionals*.

PROJECTS

Kaggle S&P500 Prediction | ACM AI

October 2025 - Present

- Design and iterate neural networks, such as decision trees, feed-forward networks, and Long Short-Term Memory (LSTMs) models to predict forward returns of S&P 500 for Kaggle competition.
- Improve data preprocessing and model evaluation by implementing KNN imputation for dataset NaN values and k-fold cross validation. Competition scoring in progress.

R1 Reasoning | ACM AI

March 2025 - June 2025

- Implemented reinforcement learning from human feedback (RLHF) system using Group Relative Policy Optimization (GRPO) to fine-tune Qwen2.5-7B-Instruct model for mathematical reasoning tasks. Improved out-of-the-box model accuracy on test data by 17 percentage points.
- Identified and debugged issues with repetition rewards, correctness metric, and dataset parameters. Implemented custom correctness checking via regex pattern matching and repetition detection using n-gram analysis.
- Optimized memory usage through gradient checkpointing, 8-bit optimizers, and automatic GPU memory management for multi-GPU training.

PUBLICATIONS

Chen, Y., Jiao, J., & Zheng, R. (2024). Exploring changes in trip generation and impacts of built environment between regular and essential trips: A study based on the contiguous United States. *Proceedings of the CICTP 2024* (pp. 3317–3326). Presented at the CICTP 2024. <https://doi.org/10.1061/9780784485484.314>

CLUBS AND ORGANIZATIONS

Association for Computing Machinery AI (Projects Officer) · Sigma Pi Sigma Honor Society

AWARDS AND CERTIFICATIONS

ICPC Break the Binary Third Place (2024) · USA Physics Olympiad Qualifier (2023)