Ryan Zheng

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

Education

University of California, Los Angeles — Los Angeles, CA
B.S. in Physics — *Sep. 2024 – June 2028*

Experience

Generative AI Intern — January 2025 – Present
Scale AI — *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, to next-gen SWE agents solving real world GitHub issues.
- Execute rigorous quality assurance reviews for critical datasets powering state-of-the-art model training, evaluating contributor and intern deliverables against strict customer specifications to ensure data integrity and compliance standards.
- Design robust problems, construct futureproof 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*

- Collaborated on a research study of the conterminous United States' travel patterns during the COVID-19 pandemic based on origin-destination data.
- Created lightweight and robust data aggregation tool using Java to process and aggregate large datasets of national travel data.
- Coauthored research paper accepted to *COTA International Conference of Transportation Professionals 2024 Conference Proceeding.*

Projects

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

Assocation for Computing Machinery AI — Project Member
Sigma Pi Sigma Honor Society — Member
Society of Physics Students — Member

Awards

ICPC Break the Binary Third Place — 2024
USA Physics Olympiad Qualifier — 2023

Technical Skills

Languages: Java, Python, Swift, C++, SQL, HTML/CSS, R
Frameworks: PyTorch, Hugging Face, React, Jekyll
Developer Tools: Git, VS Code, XCode, Cursor, Jupyter, Godot