Yikuan Li

ykuanli@outlook.com Google Scholar
github.com/Kuaaannn linkedin.com/in/yikuan-li x.com/Yi_Kuan_Li

Education

University of Warwick Sep. 2024 – Jun. 2027 (expected)
BSc Mathematics and Statistics Coventry, UK

- **Year 1:** 88.8/100 (first class; top 10 of 269). **Year 2:** 88.5/100 (first class; top 10 of 257).
- **GPA equivalent:** 4.0/4.0. **Academic Excellence Prize** in Years 1 and 2.
- **Selected coursework:** Optimization (MSc-level), Algorithms, Java Programming, Analysis I–III, Probability & Statistics, Stochastic Processes, Multilinear Algebra, Multivariable Analysis, ODEs & PDEs.

Research Interests

Pre-training and post-training of large language models, with a focus on optimizers, parameter-efficient fine-tuning (PEFT), and efficient training.

Publications and Preprints

Are Parameter-Efficient Fine-tuning Methods Really Different?
Yikuan Li, Pinyan Lu, Fanghui Liu. [Preprint]

Research Experience

Parameter-efficient fine-tuning · FLINT Lab Jun. 2025 – Present
Research with Prof. Fanghui Liu DCS, Warwick; SJTU

- Investigated orthogonal matrix parameterizations, evaluating initialization strategies, parameter efficiency, runtime overhead, and generalization.
- Compared downstream performance and retention of pretrained capabilities using a consistent evaluation protocol; found limited practical benefits from preserving singular values in orthogonal fine-tuning.
- Co-authored the preprint *Are Parameter-Efficient Fine-tuning Methods Really Different?*

Post-training of large language models · RAISE Lab Jun. 2026 – Present
Research with Dr. Paris Giampouras DCS, Warwick

- Investigate efficient evolution strategies (ES) for post-training large language models.

Understanding Gradient Orthogonalization in the Muon Optimizer Oct. 2026 – Present
Third-year essay · Supervisor: Dr. Clarice Poon WMI, Warwick

- Conduct theoretical analysis to understand why gradient orthogonalization may improve optimization.

Independent Online Coursework

Independently studied course materials and completed assignments for **Stanford** CS229 (Machine Learning), CS231N (Computer Vision), and CS224N (NLP with Deep Learning); **MIT** 18.S096 (Matrix Calculus for ML and Beyond) and 6.S184 (Flow Matching and Diffusion Models).

Technical Skills

Programming: Python, SQL, MATLAB, R, Java
Machine learning: NumPy, PyTorch, Hugging Face Transformers, PEFT
Tools: LaTeX, Git, Linux, Bash, Slurm, Claude Code, Codex, Cursor
Languages: English (fluent), Mandarin (native), Cantonese (fluent)

Additional Experience

Flywheel May 2024 – Aug. 2024
Data Analyst Intern Hangzhou, China

- Developed time-series forecasting pipelines using Prophet, SARIMA, and LSTM models to predict gross merchandise value (GMV) around major promotional events.
- Automated data consolidation and reporting in Python, reducing manual processing time by approximately 50%.