Xin Ma



xinma@mail.ustc.edu.cn | Homepage | GitHub | Google Scholar

Education

University of Science and Technology of China (USTC) **Sep 2024 – Present**
M.S. in Data Science (Statistics) *Anhui, China*

South China University of Technology (SCUT) **Sep 2020 – Jun 2024**
B.S. in Mathematics and Applied Mathematics (Statistics) | Rank: 1/69 *Guangdong, China*

- Coursework: Mathematical Analysis, Advanced Algebra, Real Analysis, Complex Analysis, Abstract Algebra, Probability Theory, Stochastic Processes, Regression Analysis, Multivariate Statistical Analysis, and more.

Selected Awards

(2024, 2025, 2026) Graduate Academic Scholarship, USTC (First Class 2024, Second Class 2025, First Class 2026).
(2024) Outstanding Graduate, SCUT.
(2022) First Prize (Guangdong Province), China Undergraduate Mathematical Contest in Modeling.
(2021–2024) National Encouragement Scholarship (top 3% national-wide).
(2020) National Scholarship (top 0.2% national-wide).

Publications

1. (2026) "*More Edits, More Stable: Understanding the Lifelong Normalization in Sequential Model Editing.*"
Xin Ma, Wei Chen, Qi Liu, Derong Xu, Zhi Zheng, Tong Xu, Enhong Chen.
The 43rd International Conference on Machine Learning (ICML 2026).

Research Experience

Understanding Lifelong Normalization in Sequential Model Editing **Jun 2025 – Jan 2026**
First author, ICML 2026

- **Problem.** Sequential model editing aims to continually update factual knowledge in LLMs without retraining from scratch. However, Most editing methods degrade after thousands of sequential edits. Editors that stay stable at the million-edit scale often rely on Lifelong Normalization (LN), an unexplained component whose removal causes collapse.
- **Theory.** Showed LN performs recursive Bayesian estimation of a drifting gradient distribution, yielding updates with **bounded norm** and **asymptotic orthogonality**, which explains why "more edits" means "more stable".
- **Method & results.** The resulting editor, **StableEdit**, stays stable to **2M edits** across 4 LLMs and 4 benchmarks, improving generalization by ∼**7.4 points** over UltraEdit.

Entropy-KL Selective Fine-Tuning for the SFT-then-RL Pipeline **Dec 2025 – Jan 2026**
Collaborative project

- **Problem & method.** Low-data SFT before RL sharpens the policy and drifts from the pretrained distribution, curbing RL exploration. This project masks high-entropy and high-KL tokens from imitation and regularizes them with entropy and KL terms instead.
- **Contribution (theory).** Explained the mechanism through a token-level gradient analysis: cross-entropy gradients are dominated by high-entropy tokens, whose one-hot forcing term drives entropy collapse and parameter drift, while the regularized update on masked tokens is **label-free**; also showed why selective regularization outperforms global regularization.
- **Evidence & results.** On Qwen3-4B/8B, the method beats standard SFT by up to **+5.1 pass@32** with less parameter drift and remains best after RL (DAPO); ablations confirm that selective beats global regularization.

Irreversible Points in On-Policy Distillation **Aug 2026 – Present**

- **Problem.** In on-policy distillation the teacher scores every token of the student's rollout, yet some student errors are corrected by the teacher while others become **irreversible**: past them, neither student nor teacher recovers a correct answer.
- **Ongoing.** Studying how these points form, what changes at them, and whether interventions, such as having the student reject its earlier reasoning and restart, can recover them.

Skills

Programming	Familiar with Python and has written C++, MATLAB, R
Mathematics	Familiar with real analysis, linear and abstract algebra, probability theory, stochastic processes, Bayesian statistics, regression and multivariate statistical analysis
ML/Research	Familiar with PyTorch, HuggingFace Transformers, PEFT/LoRA, Weights & Biases; Git, Linux, and Slurm job submission on GPU servers