Mohit Kulkarni

kulkarnimohit01@gmail.com | m2kulkarni | m2kulkarni | m2kulkarni.github.io

Education

ETH Zurich and University of Zurich *2023 – 2025*
M.Sc, Neural Systems and Computation

Indian Institute of Technology, Kanpur *2019 – 2023*
B.S, Mathematics and Scientific Computing. Minor in Machine Learning

Publications

Kulkarni, M.*, Chaudhry, H.*, Pehlevan, C. "From Recall to Reasoning: Understanding the Role of Associative Memory in Hybrid Architectures" `Submitted to ICML 2026.`

Chaudhry, H.*, **Kulkarni, M.***, Pehlevan, C. "Test-time scaling meets associative memory: Challenges in subquadratic models." `ICLR workshop on New Foundations in Associative Memory, 2025.`

Daie, K., . . . , **Kulkarni, M.**, Botvinick, M.M., Svoboda, K. "Optical brain computer interface for measuring local circuit plasticity during learning.", `In preparation.`

Experience

Research Intern, Cohere *Dec 2025 – May 2026*
- Investigating reinforcement learning methods for long-context reasoning in language models; resolved infrastructure bottlenecks for training at extended context lengths.

Research Assistant, Emerge Lab, New York University *Aug 2025 – Dec 2025*
Prof. Eugene Vinitsky
- Investigating meta-RL methods for k-shot adaptation to enable agents to rapidly align with cooperative partners.
- Developed a high-performance `welfare-diplomacy` simulator in `pufferlib` to study emergent cooperation and welfare in multi-agent reinforcement learning.

Research Fellow, SEAS, Harvard University *Sep 2024 – Aug 2025*
Prof. Cengiz Pehlevan
- Explored efficiency versus reasoning trade-offs in linear attention and hybrid architectures. Pre-trained and scaled inference for hybrid models of size 150M-1.3B to understand mathematical reasoning capabilities.
- Designed and implemented an experimental framework for pretraining and fine-tuning hybrid architectures combining `flash-linear-attention` and `OLMo`.

Research Assistant, Allen Institute for Neural Dynamics *Sep 2020 – May 2023*
Dr. Karel Svoboda, Dr. Kayvon Daie
- Built a scalable and efficient data analysis pipeline on `GCP` using `DataJoint`. Developed `RNN` models to understand learning in biological and artificial networks, comparing model predictions with experimental `2-Photon` recordings.

Selected Projects

GPU-accelerated Terminal Emulator | *Personal Project* *GITHUB*
- Developed a C++/OpenGL based terminal emulator with support for colors, advanced text rendering (`ANSI X3.64`) and ligatures; now integrating LLM function calling for seamless in-terminal AI capabilities.

Spiking Neural Network Benchmark | *SNUFA Workshop* *WEBSITE*
- Created `SNUFA100` and `SNUFA100_sentence` datasets for evaluating Spiking Neural Networks; implemented baseline with surrogate gradient descent.

Honors and Awards

Heyning-Roelli Foundation Scholarship | For Masters Thesis at Harvard *2024*
Brain Computation and Learning Workshop Travel Grant | IISc Bangalore *2023*
Cosyne 2022, Undergraduate Travel Grant | Lisbon, Portugal *2022*
INSPIRE Scholarship | Awarded throughout Bachelors *2019 – 2023*
All India Rank 637 | JEE Advanced *2019*