

Research Statement

I am Prasanna Mayilvahanan, a fourth-year Ph.D. student at the Max Planck Institute for Intelligent Systems in Germany under the supervision of Dr. Wieland Brendel. Previously, I completed my master's in Informatics at USI-Lugano/ETH Zürich and my undergrad at the Indian Institute of Technology, Guwahati. My broad research area is the 'foundations of deep learning,' and I have published four first-author papers along with several co-authored publications at top venues (NeurIPS, ICLR, ICML, IEEE, etc.). I have also done internships at Apple Machine Learning Research (Summer 2024) and at Cohere (Winter 2025). Specifically, my research focuses on robustness and reasoning in language and vision-language models. The core insight from my work is that, from a generalization perspective, pre-training data predominantly dictates downstream performance, while other common factors (architecture, optimizer, etc.) play relatively minor roles, though still relevant from an efficiency standpoint [1, 2, 3].

Given this emphasis on data-driven generalization, my current research interests span two closely related directions. First, I aim to develop training objectives, architectures, and curate data that enable creative generalization, allowing models to extrapolate qualitatively (e.g., combinatorially or transformationally) beyond the patterns explicitly present in their training distribution. In this direction, I am currently studying alternatives to standard next-token prediction, including multi-token prediction [5] objectives (e.g. diffusion models) and related formulations. Also, I am exploring data curation strategies that promote “Abductive Reasoning” [6] — encouraging models to generate latent hypotheses as part of achieving creative generalization. Second, I focus on reward modeling and exploration-based post-training methods, using reinforcement learning to deliberately push LLMs and MLLMs beyond their static pre-training distributions in order to discover new knowledge and strengthen reasoning capabilities [4]. In particular, my ongoing work examines reward modeling for multimodal reasoning, analyzing how the interaction between exploration bonuses and curriculum design can help models solve hard reasoning problems that lie beyond base model capabilities.

REFERENCES

- [1] Mayilvahanan et al. (2025). "LLMs on the Line: Data Determines Loss-to-Loss Scaling Laws." Proceedings of the International Conference on Machine Learning (ICML). <https://arxiv.org/abs/2502.12120>.
- [2] Mayilvahanan et al. (2024). "In Search of Forgotten Domain Generalization." Proceedings of the International Conference on Learning Representations (ICLR). <https://arxiv.org/abs/2410.08258>.
- [3] Mayilvahanan et al. (2023). "Does CLIP's Generalization Performance Mainly Stem from High Train-Test Similarity?" Proceedings of the International Conference on Learning Representations (ICLR). <https://arxiv.org/abs/2310.09562>.
- [4] Mayilvahanan et al. (2026). "MATH-Beyond: A Benchmark for RL to Expand Beyond the Base Model" Proceedings of the International Conference on Learning Representations (ICLR) <https://arxiv.org/abs/2510.11653>
- [5] Nagarajan et al. (2025). "Roll the dice & look before you leap: Going beyond the creative limits of next-token prediction" (ICML) <https://arxiv.org/abs/2504.15266>
- [6] Zahavy (2026). "LLMs can't jump" (PhilSci-Arxiv) <https://philsci-archive.pitt.edu/28024/>