



Self-Distillation for Language Models

PROJECT PROPOSALS FOR MASTER'S THESES & SEMESTER PROJECTS

Motivation

Language models can become their own teachers through *self-distillation* (SD). The student generates responses, while the same model, conditioned on privileged context such as verifier feedback, expert demonstrations, or user preferences, acts as its teacher (Hübotter et al., 2026; Shenfeld et al., 2026; Zhao et al., 2026; Buening et al., 2026). Minimizing the KL divergence between their next-token distributions provides a dense learning signal, enabling the student to internalize this guidance without requiring the additional context at inference time. The *on-policy* nature of the algorithm enables the model to improve task performance while limiting forgetting on out-of-distribution tasks. This makes it a promising foundation for continual learning.

Possible Project Directions

We offer theoretical and empirical projects in three broad areas:

- **Methods for continual learning through self-distillation.** Develop and improve SD algorithms that enable language models to acquire new knowledge and skills while preserving existing capabilities.
- **Theoretical understanding of self-distillation.** Study the principles and learning dynamics behind SD, including when it improves model capabilities and when it fails.
- **Curricula for self-distillation.** Investigate how the selection and ordering of tasks and feedback can support effective learning as a model's capabilities evolve.

Ideal Candidate

An ideal candidate has strong academic performance or prior research experience, a solid foundation in machine learning and probability theory, and good programming skills. Experience with reinforcement learning or LLM fine-tuning is helpful. Above all, we value a strong drive to learn and explore new concepts.

Contact

If you are interested, please contact us to discuss these projects in more detail:

- Frederike Lübeck (✉ frederike.luebeck@inf.ethz.ch)
- Anton Baumann (✉ antbaumann@ethz.ch)
- Patrik Wolf (✉ patrik.wolf@inf.ethz.ch)

Please attach your CV, transcripts, and briefly explain why you are interested in this project.

References

- Buening, T. K., Hübotter, J., Pásztor, B., Shenfeld, I., Ramponi, G., and Krause, A. (2026). Aligning language models from user interactions. In *Third Conference on Language Modeling*.
- Hübotter, J., Lübeck, F., Behric, L. D., Baumann, A., Bagatella, M., Marta, D., Hakimi, I., Shenfeld, I., Buening, T. K., Guestrin, C., and Krause, A. (2026). Reinforcement learning via self-distillation. In *Forty-third International Conference on Machine Learning*.
- Shenfeld, I., Damani, M., Hübotter, J., and Agrawal, P. (2026). Self-distillation enables continual learning. In *Forty-third International Conference on Machine Learning*.
- Zhao, S., Xie, Z., Liu, M., Huang, J., Pang, G., Chen, F., and Grover, A. (2026). Self-distilled reasoner: On-policy self-distillation for large language models.