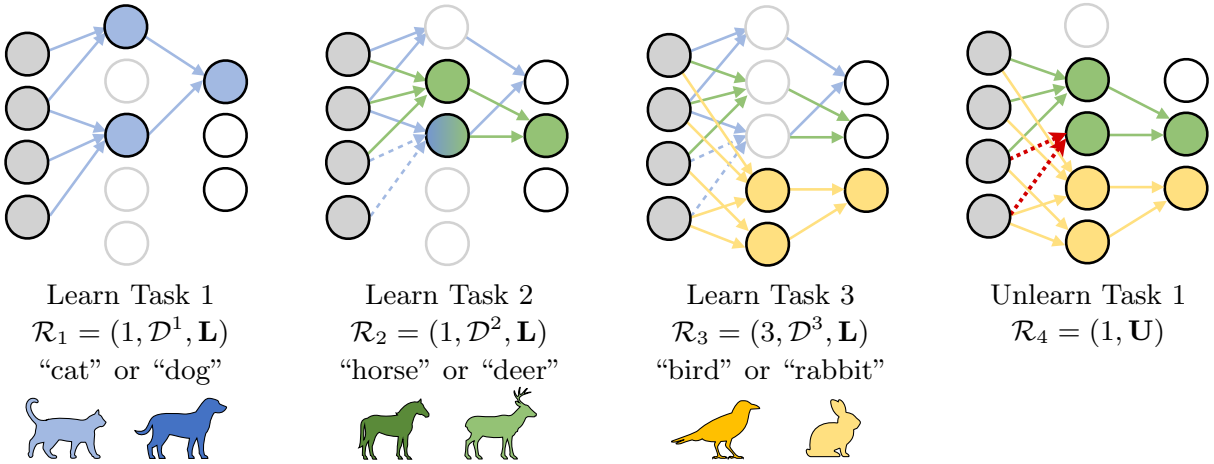
Enhanced Privacy-Awareness in Lifelong Learning



Lifelong learning enables models to incrementally acquire new knowledge without forgetting previously learned information. Contrarily, machine unlearning focuses on deliberately forgetting certain previous knowledge from pretrained models when requested, in order to comply with data privacy regulations. Enabling efficient lifelong learning with the capability to selectively unlearn sensitive information from models presents a challenging problem with contradicting objectives. Recent work [1] formalized this problem, and proposed a solution from the perspective of simultaneously preventing catastrophic forgetting and allowing forward knowledge transfer during task-incremental learning, while ensuring exact task unlearning and minimizing memory requirements, based on a single neural network model to be adapted. In this project, we will build on this framework to satisfy per-sample privacy guarantees under strict memory limitations, and study this problem in different applications.

[1] O. Özdenizci et al., "Privacy-aware lifelong learning", ICLR 2025.

Goals & Tasks

- Review of state-of-the-art on lifelong deep learning and machine unlearning methods.
- Simulating and benchmarking with the existing PALL framework [1].
- Extending this framework with exact class-unlearning and selective sample-unlearning techniques.

Qualifications

- Interest in privacy-preserving ML.
- Experience with the Python based deep learning framework PyTorch.
- Registered to one of the following:
  - ☐ **Bachelor Thesis**
  - ✓ **Seminar Project**
  - ✓ **Master Thesis**

Contact

Ozan Özdenizci: oezdenizci@tugraz.at