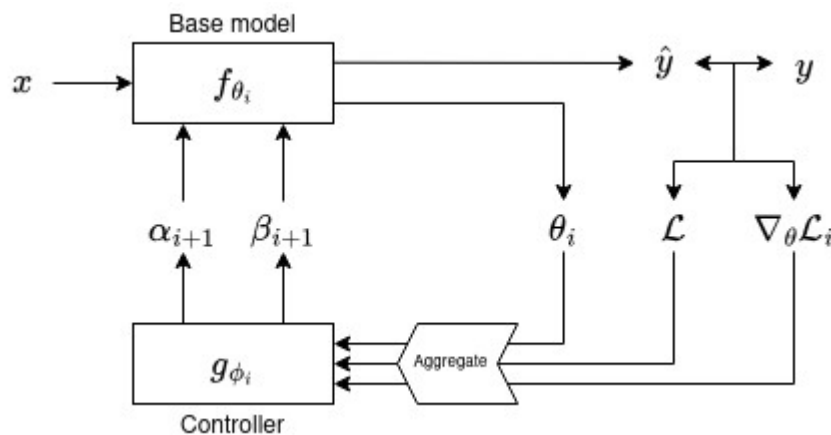


Master Thesis: Online Hyperparameter Adaptation via Learned Controllers

Thesis Idea:

The process of selecting optimal hyperparameters is a critical challenge in training machine learning models. Despite significant progress in hyperparameter search strategies such as grid search or random search, most methods rely on static hyperparameter schedules or manual tuning, which are often suboptimal and time-consuming. Instead, the goal of this Master Thesis is to pair a machine learning model with a light-weight controller that dynamically adjusts the used training hyperparameters throughout the course of the training. The controller learns to optimize hyperparameters online, based on the model's evolving training dynamics, including loss trends, gradient statistics, and performance metrics. To maintain a small controller size, these training dynamics must be captured in a selection of aggregated features. A first concept for this algorithm is shown below. The machine learning model, to which the controller is to be attached, will be provided. The student's task will be to create the aggregated features and to design and evaluate the controller model. The evaluation could investigate whether the controller accelerates convergence, enhances generalization, improves training stability, and how its performance compares to established static scheduling methods. Layer-wise adaptation approaches could also be explored.



1. Train base model
2. Train controller
3. Repeat steps 1 and 2 until base model converges. Update hyperparameters every N steps.

What we offer:

- GPU infrastructure for model training and evaluation
- Exchange with experts in the field of machine learning
- Working on a relevant research topic with a wild range of application
- Option to write a scientific research paper about the Master Thesis

What we are looking for:

- **background in machine learning (ML)**
- **experience with Python and PyTorch**
- motivation to not just apply machine learning but conduct foundational research
- Preferred: strong understanding of hyperparameters and optimizers

Contact:

Anthony Richardson

Email: antmen@uni-bremen.de