"The most widely used machine learning algorithms were invented and hardwired by humans. Can we also construct meta-learning (or meta-learning) algorithms that can learn better learning algorithms?"

Juergen Schmidhuber (link).

In this Bachelor thesis, the student will study and extend a prominent meta-learning technique termed \textit{MAML} (1). In this thesis, we will focus on applying MAML for few-shot learning and/or reinforcement learning. See the paper "Model Agnostic Meta-learning" (1) and our recently published extension "sparse-MAML" (2) which will be a starting point of this thesis.

In MAML, one tries to meta-learn a neural network initialization $\theta$ which is adjusted on a small training set $D^\tau_t$ by $K$ gradient steps. The aim of the algorithm is that the resulting weights $\phi_{\tau,K}$ parametrize a neural network function $f(x; \phi_{\tau,K})$ that generalizes i.e. has low loss on validation data $D^\tau_v$. This can be expressed by the following optimization problem

$$\min_{\theta} E_{\tau \sim \rho(\tau)} \left[ \mathcal{L}(\phi_{\tau,K}(\theta, \alpha), D^\tau_v) \right]$$

s.t. $\phi_{\tau,k+1} = \phi_{\tau,k} - \alpha \nabla_{\phi} \mathcal{L}(\phi_{\tau,k}, D^\tau_t)$ and $\phi_{\tau,0} = \theta,$ \hfill (1)

Despite the initialization, one can additionally meta-learn more meta-parameters influencing the training dynamics e.g. the learning rates $\alpha$ or the training length $K$. These extensions will be investigated by the student in this thesis. The outline of the project is:

1. Study, re-implement and reproduce the MAML algorithm and results.
2. Extend the MAML algorithm by allowing for an adaptive training length $K$ per parameter.
3. Early experiments showed that for some parameters, very large $K$ are found by MAML. This should be prevented by the student implementing accelerated gradient descent methods.

For this thesis, deep learning experience is highly recommended.

**Prerequisites:** Knowledge of Tensorflow, Pytorch or JAX

**Supervisor:** Johannes Oswald, CAB J 21.2, voswaldj@ethz.ch

**Supervising Professor:** Prof. Dr. Angelika Steger, CAB G 37.2, steger@inf.ethz.ch

**References**
