

Exploring gradient-based optimization techniques for deep learning

Deep neural networks (DNNs) are usually trained with gradient-based optimization techniques. While stochastic gradient descent works usually quite well, more recent techniques like ADAM are known to be more robust. In addition, natural gradient methods have been proposed but are less often used in practice. Here we will investigate when one optimization technique is preferable over others in deep learning. We will consider standard benchmark datasets like MNIST and investigate the influence of re-parametrization of networks on the optimization.

Goals & Tasks

- Review of the state-of-the-art on natural gradient methods.
- Implementation of DNNs for standard benchmark tasks and very simple toy tasks for conceptual experiments.
- Exploring the influence of re-parametrization of networks on the optimization.

Contact

Robert Legenstein
robert.legenstein@igi.tugraz.at

Qualifications

- Interest in deep learning.
- Experience with Python based deep learning frameworks such as TensorFlow or PyTorch are beneficial.
- Registered to one of the following:
 - ✓ **Bachelor Thesis**
 - ✓ **Project, Seminar/Project**
 - Master Thesis**

Extension for a Master Thesis is possible on interesting results.