



Vorträge
Tenure Track „Machine Learning“

Dienstag, 2. April 2019, Hörsaal 10

Levent Sagun
(École Polytechnique Fédérale de Lausanne)

9:00 Uhr – 9:20 Uhr: Didaktischer Vortrag

„Backprop Algorithm and Automatic Differentiation“

One of the main driving forces behind modern deep neural networks is the backprop algorithm. In this brief lecture, we will start with an elementary review of deep learning from the point of view of differentiable programming. Then we will introduce automatic differentiation and in particular, we will focus on the backprop algorithm. We will provide fundamental examples, discuss their computational complexity, and review their modern applications.

9:50 Uhr – 10:35 Uhr: Wissenschaftlicher Vortrag

„Over-parametrization in neural networks: an overview and a definition“

An excursion around the ideas for why the stochastic gradient descent algorithm works well on training deep neural networks leads to considerations about the underlying geometry of the related loss function. Recently, we gained a lot of insight into how tuning SGD leads to better or worse generalization properties on a given model and task. Furthermore, we have a reasonably large set of observations that lead to the conclusion that more parameters typically lead to better accuracies as long as the training process is not hampered. In this talk, I will speculatively argue that as long as the model is over-parameterized (OP), all solutions are equivalent up to finite size fluctuations.

We will start by reviewing some of the recent literature on the geometry of the loss function, and how SGD navigates the landscape in the OP regime. Then we will see how to define OP by finding a sharp transition described by the models fitting abilities to its training set. Finally, we will discuss how this critical threshold is connected to the generalization properties of the model, and argue that life beyond this threshold is (more or less) as good as it gets.
