TT 1: Tutorial: Physics Meets Machine Learning (joint session DY/TUT/TT)

Machine learning has revolutionized many application fields such as computer vision and natural language processing. In physics there is a growing interest in using machine learning to enhance the analysis of experimental data and to devise and optimize experiments or numerical simulations. On the other hand physicists use their intuition and methods from statistical physics and complex systems theory to better understand the working principles of modern machine learning methods. This tutorial session introduces some subfields within this area and the basic methods involved.

Organized by Sabine Andergassen (Tübingen), Martin Gärttner (Heidelberg), Moritz Helias (Jülich), and Markus Schmitt (Cologne)

Time: Sunday 16:00–18:15 Location: HSZ 01

Tutorial TT 1.1 Sun 16:00 HSZ 01

Machine Learning for Quantum Technologies — •Florian Marquardt — Max Planck Institute for the Science of Light and Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

Machine learning is revolutionizing science and technology. In the past few years, it has become clear that it promises significant benefits as well for the development of quantum technologies. In this tutorial I will first give a brief introduction to neural networks. I will then discuss a number of areas and examples in which machine learning is being successfully applied in this context. These include measurement data analysis and quantum state representation, approximate quantum dynamics, parameter estimation, discovering strategies for hardware-level quantum control, the optimization of quantum circuits, and the discovery of quantum experiments, discrete quantum feedback strategies, and quantum error correction protocols.

Reference: "Artificial intelligence and machine learning for quantum technologies", M. Krenn, J. Landgraf, T. Foesel, and F. Marquardt, Phys. Rev. A 107, 010101 (2023).

Tutorial TT 1.2 Sun 16:45 HSZ 01

The Unreasonable Effectiveness of Gaussians in the Theory of Deep Neural Networks — •Zohar Ringel — Racah Institute of Physics, Hebrew University in Jerusalem

Physical Sciences are in many ways the success story of explaining fundamental phenomena using simple math [1]. The fact that physical phenomena could be arranged in that manner is remarkable. Yet this simplicity does not necessarily carry over to life sciences or data sciences. Indeed prominent authors have argued against our desire to rely on neat mathematical structures when analyzing big data [2].

In the past half-decade several results have emerged which balance mathematical simplicity with data-induced complexity. These could be seen as a middle ground between the above juxtaposing views. The common divider here is the use of Gaussian distributions as approximants of various different quantities in deep neural networks (DNNs). Specifically these Gaussians emerge when describing outputs of DNNs with random weights, outputs of trained DNNs at random times, outputs of fixed DNNs over random input data, and fluctuations of hidden DNN pre-activations. In this tutorial I will present these quantities, provide arguments supporting their Gaussianity, and outline several theoretical implications.

[1] The Unreasonable Effectiveness of Mathematics in the Natural Sciences. Wigner (1960)

[2] The Unreasonable Effectiveness of Data. Halevy, Norvig, Pereira (2009)

Tutorial TT 1.3 Sun 17:30 HSZ 01

Computing learning curves for large machine learning models using the replica approach — •Manfred Opper — Inst. für Softwaretechnik und Theor. Informatik, TU Berlin — Centre for Systems Modelling and Quantitative Biomedicine, University of Birmingham, UK

Methods of statistical physics have been used for a long time to mathematically analyse the typical performance of machine learning models in the limit where both the number of data and the number of parameters (such as network weights) is large. By defining Boltzmann-Gibbs probability distributions over parameters where the cost function of the machine learning problem plays the role of a hamiltoninan, one can derive analytical expressions for training errors and generalisation errors using the corresponding partition functions and free energies in terms of a usually small number of order parameters.

Since the models depend on a set of random data to be learnt, additional appropriate statistical (so-called quenched) averages of free energies over this 'disorder' have to be performed. The replica approach is a prominent analytical tool from the statistical physics of disordered systems to solve this nontrivial technical challenge.

In this tutorial I will give an introduction to this approach. Starting with an explicit calculation for simple single layer perceptrons, I will then argue how the method can be applied to more complex problems such as kernel machines (support vector machines and Gaussian processes) and multilayer networks.