DY 11: Focus Session: Physics Meets ML II – Understanding Machine Learning as Complex Interacting Systems (joint session DY/TT)

Machine-learning has recently entered and is now transforming many fields of science, enabling discoveries in a data-driven manner. As a scientific method, however, ML often lacks one defining feature: Explainability. We here seek discussions with pioneers in understanding, explaining, and improving machine learning methods from the point of view as a physical system of interacting elements. In fact, the history of approaching neuronal networks and problems of inference and learning as a problem of statistical physics has a long history, with a number of important discoveries early on. The close relation between spin glasses and neuronal networks are being currently exploited to address pressing questions, such as the remarkable generalization properties of neuronal networks despite their massive overparameterization and their behavior reminiscent of renormalization group transformations.

Organized by Sabine Andergassen (Tübingen) and Moritz Helias (Jülich)

Time: Monday 15:00–18:30 Location: ZEU 250

Invited Talk DY 11.1 Mon 15:00 ZEU 250
The challenge of structured disorder in statistical physics — •Marc Mezard — Bocconi University, Milano

Statistical physics offers many interesting tools to study machine learning. In most cases it needs to use a statistical ensemble of data. Most of the theoretical work has relied on unstructured data. Yet, the highly structured character of data used in training deep networks is a crucial ingredient of their performance. Modelling structured data, analyzing the learning and the generalization of deep networks trained on this data, are major challenges. This talk will describe several recent developments in this direction.

Invited Talk DY 11.2 Mon 15:30 ZEU 250
The emergence of concepts in shallow neural-networks — •Elena Agliari — Piazzale A. Moro 5, 00185 Roma

In the first part of the seminar I will introduce shallow neural-networks from a statistical-mechanics perspective, focusing on simple cases and on a naive scenario where information to be learnt is structureless. Then, inspired by biological information-processing, I will enrich the framework and make the network able to successfully and cheaply handle structured datasets. Results presented are both analytical and numerical.

Invited Talk DY 11.3 Mon 16:00 ZEU 250
Adaptive Kernel Approaches to Feature Learning in Deep Neural Networks — •Zohar Ringel — Racah Institute of Physics, Hebrew University in Jerusalem

Following the ever-increasing role of deep neural networks (DNNs) in our world, a better theoretical understanding of these complex artificial objects is desirable. Some progress in this direction has been seen lately in the realm of infinitely overparameterized DNNs. The outputs of such trained DNNs behave essentially as multivariate Gaussians governed by a certain covariance matrix called the kernel. While such infinite DNNs share many similarities with the finite ones used in practice, various important discrepancies exist. Most notably the fixed kernels of such DNNs stand in contrast to feature learning effects observed in finite DNNs. Such effects are crucial as they are the key to understanding how DNNs process data. To accommodate such effects within the Gaussian/kernel viewpoint, various ideas have been put forward. Here I will provide a short overview of those efforts and then discuss in some detail a general set of equations we developed for feature learning in fully trained/equilibrated DNNs. Interestingly, our approach shows that DNNs accommodate strong feature learning via mean-field effects while having decoupled layers and decoupled neurons within a layer. Furthermore, learning is achieved not by compression of information but rather by increasing neuron variance along label-relevant directions in function space.

DY 11.4 Mon 16:30 ZEU 250
Interpreting black-box ML with the help of physics — •Miriam Klopotek — University of Stuttgart, SimTech Cluster of Excellence EXC 2075, Stuttgart, Germany

Complexity is an unavoidable part of systems with emergent or even so-called intelligent capabilities. Ultimately, it stems from the many microscopic constituents with multiple possible states, which introduces a vast space of degrees of freedom. This is true both for many-body systems as well as modern machine learning (ML) systems. Today, the latter suffer notoriously from the 'black-box problem', i.e. they are inherently opaque. We argue that an engagement with physics can offer deep insights ultimately for a theory of operation and thus an interpretation, as well as powerful ways to assess their reliability and shortcomings. We show some results for a case study with beta-variational autoencoders (β-VAEs) [1], which we trained on data from a well-characterized model system of hard rods confined to 2D lattices [2].

[1] D. P. Kingma and M. Welling, ICLR 2014. D. J. Rezende, S. Mohamed, and D. Wierstra, ICML 2014, p. 1278-1286.
[2] P. Quiring, M. Klopotek and M. Oettel, Phys. Rev. E 100, 012707 (2019).

15 min. break

Invited Talk DY 11.5 Mon 17:00 ZEU 250
Analysing the dynamics of message passing algorithms — •Manfred Opper[1,2] and Burak Cakmak[1] — [1]Institut für Softwaretechnik und Theoretische Informatik, Technische Universität Berlin, 10587, Germany — [2]Centre for Systems Modelling and Quantitative Biomedicine, University of Birmingham, B15 2TT, United Kingdom

Message passing algorithms are deterministic methods which are designed for efficiently computing marginal statistics for probabilistic, Bayesian data models used in machine learning and statistics. Such algorithms have been developed in parallel within the machine learning and the statistical physics communities. They often provide highly accurate approximations at a much higher speed compared to exact Monte Carlo sampling. The fixed points of such algorithms can be analysed for high dimensional models (under the assumption of specific data distributions) using the replica method of statistical physics. In this talk we will focus on the dynamical properties of the algorithms. Applying dynamical functional techniques to the nonlinear dynamics, the degrees of freedom which interact via a random matrix can be decoupled in the limit of large systems resulting in exact stochastic single node dynamics. For general dynamical models, it is hard to further analyse this effective dynamics due to the occurence of memory terms. Surprisingly, for message passing algorithms memory terms are absent and exact results for convergence rates and stability can be derived for specific data distributions.

Invited Talk DY 11.6 Mon 17:30 ZEU 250
Deep Learning Theory Beyond the Kernel Limit — •Cengiz Pehlevan — Harvard University, USA

Deep learning has emerged as a successful paradigm for solving challenging machine learning and computational problems across a variety of domains. However, theoretical understanding of the training and generalization of modern deep learning methods lags behind current practice. I will give an overview of our recent results in this domain, including a new theory that we derived by applying dynamical field theory to deep learning dynamics. This theory gives insight into internal representations learned by the network under different learning rules.

DY 11.7 Mon 18:00 ZEU 250
Solving the Bethe–Salpeter equation with exponential con-

vergence — •Markus Wallerberger[1], Hiroshi Shinaoka[2], and Anna Kauch[1] — [1]TU Wien, Vienna, Austria — [2]Saitama University, Japan

The Bethe–Salpeter equation plays a crucial role in understanding the physics of correlated fermions, relating to optical excitations in solids as well as resonances in high-energy physics. Yet, it is notoriously difficult to control numerically, typically requiring an effort that scales polynomially with energy scales and accuracy. This puts many interesting systems out of computational reach.

Using the intermediate representation and sparse modeling for two-particle objects on the Matsubara axis, we develop an algorithm that solves the Bethe–Salpeter equation in $O(L^8)$ time with $O(L^4)$ memory, where L grows only logarithmically with inverse temperature, bandwidth, and desired accuracy. This opens the door for computations in hitherto inaccessible regimes. We benchmark the method on the Hubbard atom and on the multiorbital weak-coupling limit, where we observe the expected exponential convergence to the analytical results. We then showcase the method for a realistic impurity problem.

[1] M. Wallerberger et al., Phys. Rev. Research 3, 033168 (2021)

DY 11.8 Mon 18:15 ZEU 250

Making machines untangle the parquet equations — •Samuel Badr[1], Anna Kauch[1], Hiroshi Shinaoka[2], Karsten Held[1], and Markus Wallerberger[1] — [1]TU Wien, Vienna, Austria — [2]Saitama University, Saitama, Japan

Diagrammatic theories at the two-particle level are increasingly important in understanding the subtle interplay of phenomena occurring in strongly correlated electron systems. The parquet equations are a centerpiece of many such theories, since they are the simplest unbiased topological classification of two-particle diagrams. However, due to their eponymous interlocking structure, the parquet equations are vexingly difficult to solve, requiring prohibitive amounts of memory.

We tackle this problem using the recently developed, machine learning inspired, techniques: firstly, the overcomplete intermediate representation, a highly compressed model for two-particle objects which is guaranteed to converge exponentially; secondly, a sparse set of Matsubara frequencies tailored to the structure of the parquet equations. This allows us to perform convolutions and frequency shifts at no loss of accuracy.

We benchmark our solver for the Hubbard atom, where we reproduce analytic results, and then showcase the solver for more extended systems.