AKPIK 2: Machine Learning Prediction and Optimization Tasks

Time: Tuesday 9:30–10:45 Location: BEY/0127

AKPIK 2.1 Tue 9:30 BEY/0127

Bayesian Optimization for Mixed-Variable Problems in the Natural Sciences — •Yuhao Zhang[1], Ti John[2], Matthias Stosiek[1], and Patrick Rinke[1,3] — [1]School of Natural Sciences Physics Department, Technical University of Munich, Germany — [2]Department of Computer Science Aalto University, FInland — [3]Munich Center for Machine Learning, Germany

Optimizing expensive black-box objectives over mixed search spaces is a common challenge across the natural sciences. Bayesian optimization (BO) offers sample-efficient strategies through probabilistic surrogate models and acquisition functions. However, its effectiveness diminishes in mixed or high-cardinality discrete spaces, where gradients are unavailable and optimizing the acquisition function becomes computationally demanding. In this work, we generalize the probabilistic reparameterization (PR) approach of Daulton et al. to handle non-equidistant discrete variables, enabling gradient-based optimization in fully mixed-variable settings with Gaussian process surrogates. With real-world scientific optimization tasks in mind, we conduct systematic benchmarks on synthetic and experimental objectives to obtain an optimized kernel formulations and demonstrate the robustness of our generalized PR implementation. We additionally show that, when combined with a modified BO workflow, our approach can efficiently optimize highly discontinuous and discretized objective landscapes. This work establishes a practical BO framework for addressing fully mixed optimization problems encountered in the natural sciences.

AKPIK 2.2 Tue 9:45 BEY/0127

Overparametrization bends the landscape: BBP transitions at initialization in simple Neural Networks — •Brandon Livio Annesi, Chiara Cammarota, and Dario Bocchi — Sapienza Univeristy, Rome Italy

High-dimensional non-convex loss landscapes play a central role in the theory of Machine Learning. Gaining insight into how these landscapes interact with gradient-based optimization methods, even in relatively simple models, can shed light on this enigmatic feature of neural networks. In this talk, I will focus on a prototypical simple learning problem, which generalizes the Phase Retrieval inference problem by allowing the exploration of overparametrized settings. Using techniques from field theory, we analyze the spectrum of the Hessian at initialization and identify a Baik*Ben Arous*Péché (BBP) transition in the amount of data that separates regimes where the initialization is informative or uninformative about a planted signal. Crucially, we demonstrate how overparameterization can bend the loss landscape, shifting the transition point, even reaching the information-theoretic weak-recovery threshold in the large overparameterization limit, while also altering its qualitative nature. We distinguish between continuous and discontinuous BBP transitions and support our analytical predictions with simulations. In the case of discontinuous BBP transitions strong finite-N corrections allow the retrieval of information at a signal-to-noise ratio (SNR) smaller than the predicted BBP transition. In these cases we provide estimates for a new lower SNR threshold that marks the point at which initialization becomes entirely uninformative.

AKPIK 2.3 Tue 10:00 BEY/0127

Training convolutional neural networks with the forward - forward algorithm — •Matthias Schröter[1,2], Frauke Alves[3], and Riccardo Scodellaro[3] — [1]Institute for Diagnostic and Interventional Radiology, University Medical Center Göttingen, Robert Koch-Straße 40, 37075 Göttingen, Germany — [2]Max Planck Institute for Dynamics and Self-Organization, 37075 Göttingen, Germany — [3]Translational Molecular Imaging, Max Planck Institute for Multidisciplinary Sciences, 37075 Göttingen, Germany.

Recent successes in image analysis with deep neural networks are achieved almost exclusively with Convolutional Neural Networks (CNNs) trained using the backpropagation (BP) algorithm. In a 2022 preprint, Geoffrey Hinton proposed the Forward - Forward (FF) algorithm as a biologically inspired alternative, where positive and negative examples are jointly presented to the network and training is guided by a locally defined goodness function. Here, we extend the FF paradigm to CNNs. This talk compares FF and BP training across different datasets (MNIST, CIFAR 10, CIFAR 100) discusses different optimization strategies, and provides insights into the inner workings of FF trained networks using Class Activation Maps.

AKPIK 2.4 Tue 10:15 BEY/0127

Modeling resonant soliton interactions in the Kadomtsev-Petviashvili equation using PINNs — •Gerald Kämmerer — Universität Duisburg

Resonant two-soliton interactions in the Kadomtsev-Petviashvili (KP) equation are modeled using Physics-Informed Neural Networks (PINNs). This framework directly solves the KP equation by incorporating the governing partial differential equation residuals into the loss function, specifically focusing on Y-shaped resonances and web-like patterns that occur under specific resonance conditions. Comparisons with known algebraic solutions show a good agreement in capturing characteristic interaction patterns. To accelerate the learning of complex dynamics, progressive training strategies and symmetry-informed network architectures are implemented, embedding the equation's inherent coordinate symmetries. The results demonstrate that PINNs can capture the rich dynamics of resonant soliton interactions, offering a framework for exploring parameter regimes beyond traditional numerical methods.

AKPIK 2.5 Tue 10:30 BEY/0127

Phase Transitions reveal Accuracy Hierarchies in Deep Learning — •Ibrahim Talha Ersoy[1], Andrés Fernando Cardozo Licha[2], and Karoline Wiesner[1] — [1]Universität Potsdam, Institut für Astronomie und Physik, Potsdam, Deutschland — [2]Universidade Federal Fluminense, Instituto de Física, Niterói, Brazil

Training Deep Neural Networks relies on the model converging on a high-dimensional, non-convex loss landscape toward a good minimum. However, much of the phenomenology of training remains ill understood. We focus on three seemingly disparate phenomena: the observation of phase transitions akin to statistical physics, the ubiquity of saddle points, and mode connectivity which is key for the active research area of model merging. We bring these into a single explanatory framework, that of the geometry of the loss and error landscapes. We show analytically that phase transitions in DNN learning are governed by saddle points in the loss landscape. Furthermore, we present a simple, easy to implement and fast algorithm, using the L2 regularizer as a tool, to explore the geometry of error landscapes. We demonstrate its use for efficiently finding paths connecting global minima by confirming the mode connectivity for DNNs trained on the MNIST data set to then use it to show numerically that saddle points in DNN loss landscapes mark transitions between distinct models that encode distinct digits of the MNIST data. Our work establishes the geometric origin of key DNN training phenomena and reveals hierarchically ordered accuracy basins analogous to phases in statistical physics.