Speaker, Title and Abstract :

International Conference / School on Machine Learning Physics

International Lectures

Lei Wang (Chinese Academy of Sciences)
"Generative models for physicists"
These lectures will cover:
1. Motivation and a dictionary between generative models and statistical physics,
2. Boltzmann machines, autoregressive models, variational autoencoders, normalizing flows, diffusion models, and more, from a physicist's perspective,
3. Applications of generative models to many-body problems.
It ends with a remark that the Universe is a generative model. A model that generates everything from THE action.

Sebastian Goldt (SISSA)
"Analytical approaches to the learning dynamics of two-layer neural networks"
This course will discuss the dynamics of stochastic gradient descent with neural networks. Starting from classical work by D. Saad and S. Solla [PRL, PRE '95], we will use tools from statistical physics to derive an analytical description of the learning dynamics of two-layer neural networks trained on a simple data model. This framework will allow us to discuss a couple of recent developments: (1) the connection to mean-field approaches for describing the dynamics of wide networks; (2) the extension to more complex models of data, via the Gaussian equivalence principle and beyond; and (3) more fine-grained approaches to early learning dynamics (one step of SGD, "escaping mediocrity"). Throughout the lectures, we will illustrate our results on popular examples such as the teacher-student setup or Gaussian mixture classification, and try to unveil common motifs across the different approaches.

Sven Krippendorf (LMU Munich)
"Machine learning as a method in mathematical physics"
In these lectures I will discuss how machine learning methods can be adapted to answer questions in mathematical physics. Concretely, I will discuss how machine learning methods can be used to detect symmetries. I plan to introduce the relevant ML techniques alongside it

such that there is no pre-requisite of knowing ML techniques. Finally, I also plan to give an overview on where we have successfully applied this framework as of now in more specialised physics questions, highlighting which technological adaptations were crucial in the process.

International Conference

Gary Shiu (U. Wisconsin)

"Learning from Topology: Cosmological Parameter Inference from the Large-scale Structure"

A challenge common to different scientific areas is to effectively infer from big, complex, higher-dimensional datasets the underlying theory. Persistent homology is a tool in computational topology developed for recognizing the ``shape" of data. Such topological measures have the advantages that 1) they are stable against experimental noise, 2) they probe multiscale, non-local characteristics of a dataset, 3) they provide interpretable statistics that encode information of all higher-point correlations. In this talk, I will focus on the applications of persistent homology to cosmological inference. I'll show how the constraints on primordial non-Gaussianities and cosmological parameters derived from persistent homologies are generally tighter than those from the redshift-space power spectrum and bispectrum combined. I'll also present a recent work in which we proposed a neural network model to map persistence images to cosmological parameters. Through a parameter recovery test, we demonstrate that our model makes accurate and precise estimates, considerably outperforming conventional Bayesian inference approaches. This provides a proof-of-concept that topological data analysis (via persistent homology) and machine learning can be combined for cosmological inference.

Jesse Thaler (MIT / IAIFI)

"The Hidden Geometry of Particle Collisions"

Since the 1960s, particle physicists have developed a variety of data analysis techniques for the goal of comparing experimental measurements to theoretical predictions. Despite their numerous successes, these techniques can seem esoteric and ad hoc, even to practitioners in the field. In this talk, I explain how many particle physics analysis tools have a natural geometric interpretation in an emergent "space" of collider events induced by the Wasserstein metric. This in turn suggests new machine learning strategies to interpret point cloud data sets from collider physics and beyond.

James Halverson (Northeastern U / IAIFI)

"Rigorous Applications of Machine Learning: from Particles to the Poincare Conjecture"

Though powerful, machine learning techniques are often stochastic, error-prone, and blackbox. How, then, should we apply them in fields like theoretical physics and pure mathematics that place a high value on rigor and understanding? In this talk I will demonstrate how applied ML may be made rigorous by conjecture generation and solution verification, through applications in knot theory related to the Poincare conjecture. I'll also show that rigor may arise by using ML theory, with new approaches to Riemannian metric flows and quantum field theory.

Long-Gang Pang (Frankfurt U)

"Deep Learning for High Energy Nuclear Physics"

High energy nuclear physics has long been captivated by the mysteries of the universe, particularly the formation and evolution of hot and dense quark-gluon plasma (QGP) in the early universe. While QGP can be recreated in the laboratory through relativistic heavy ion collisions, capturing its behavior directly has proven challenging due to its short lifetime. Instead, scientists have turned to detecting the four-momentum of final state hadrons, which are produced after the QGP freezes out and form a point cloud in the momentum space that contains valuable information about QGP properties and initial nuclear and partonic structure. In recent years, deep learning techniques have emerged as a powerful tool for solving inverse problems and variational problems in high energy nuclear physics. By leveraging optimization, auto-differentiation, and advanced algorithms, these models can learn complex relationships between input data and desired outputs, paving new avenues to studying QGP properties and initial conditions. This breakthrough has significant implications for advancing our understanding of fundamental physics and developing new technologies for exploring the universe.

Haiping Huang (Sun Yat-sen U)

"Geometric computation in deep neural networks"

We propose a mode decomposition learning that can interpret the weight matrices as a hierarchy of latent modes. These modes are akin to patterns in physics studies of memory networks, but the least number of modes increases only logarithmically with the network width and even becomes a constant when the width grows further. The mode decomposition learning not only saves a significant large amount of training costs but also explains the network performance with the leading modes, displaying a striking piecewise power-law behavior. The modes specify a progressively compact latent space across the network hierarchy, making a more disentangled subspace compared to standard training. Our mode decomposition learning is also studied in an analytic online learning setting, which reveals multiple stages of learning dynamics with a continuous specialization of hidden nodes. Extension of this framework to geometry aware computation in various architectures of different nature is also discussed.

REFERENCES

1. C. Li and H. Huang, Phys. Rev. Research 5, L022011 (2023)
2. Z. Lin and H. Huang, unpublished (2023)
3. Y. Wang, M. Xie, W. Huang, and H. Huang, unpublished (2023)

Masashi Sugiyama (RIKEN AIP)

"Towards Trustworthy Machine Learning from Weakly Supervised, Noisy, and Biased Data"

When training a machine learning system, the training data suffers from a variety of uncertainties such as insufficient information, label noise, and bias. In this talk, I will give an overview of our recent research on reliable machine learning, including weakly supervised classification, noisy label classification, and transfer learning. Finally, I would like to discuss with physicists how basic machine learning technology can be further developed.

Phiala Shanahan (MIT)

" Generative models for first-principles calculations of the structure of matter "

Novel approaches to machine learning are transforming the landscape of theoretical physics. In this context, I will discuss opportunities for generative models to accelerate first-principles theory calculations in particle and nuclear physics. Particular challenges to this paradigm include incorporating complex symmetries into model architectures, scaling models to the large number of degrees of freedom of state-of-the-art numerical studies, and designing machine-learning-accelerated algorithms that are provably exact. I will illustrate the potential of this approach by describing first studies that demonstrate that ML-accelerated sampling can be orders of magnitude more efficient than traditional algorithms, such as Hamiltonian/hybrid Monte Carlo, in the context of lattice quantum field theory calculations for nuclear physics.

Eun-Ah Kim (Cornell U)

"Data-centric learning of Quantum Many-body States with Classical Machines"

Decades of efforts by the quantum matter research community drove a "data revolution." Modern experimental modalities produce high-dimensional data in large volumes. Unprecedented control and new facilities imply new dimension and new knobs, such as timeresolved probing or scanning probing. Moreover, through recent advances in quantum simulators, quantum many-body dynamics can be simulated in various quantum computing platforms. Such many-body states are probed through projective measurements resulting in bit-strings that reside in exponentially large dimensional space. I will discuss how to learn the nature of quantum many-body states encoded in the data of the new era through data-centric approaches using machine learning. Ultimately, such "learning" should aim to accelerate discoveries and gain new insights. A synergy between data science and quantum matter physics is essential for this. I will present cases of fruitful collaborations that led to new insights and started to shape an approach to data sets of the new era.

Sven Krippendorf (LMU Munich)

"Physics to understand neural network dynamics"

Large neural networks (e.g. large language models) require a large amount of energy to train while our brain is much more energy efficient. To optimise our design of neural networks, I will discuss how physics models can play a significant role in quantifying the dynamical behaviour of neural networks. This includes a duality between the physical system of a dynamical scalar field in an expanding Universe and neural networks trained via gradient descent. In the second part I will introduce collective variables to describe neural network dynamics and discuss their behaviour throughout training. I will comment on applications of this framework to optimise our design of neural network training.

Sebastian Goldt (SISSA)

"The Gaussian world is not enough -- how training data shapes neural representations"

What do neural networks learn from their data ? We discuss this question in two learning paradigms: supervised classification with feed-forward networks, and masked language modelling with transformers. First, we give analytical and experimental evidence for a “distributional simplicity bias”, whereby neural networks learn increasingly complex distributions of their inputs. We then show that neural networks learn from the higher-order cumulants (HOCs) more efficiently than lazy methods, and show how HOCs shape the learnt features. We finally characterise the distributions that are learnt by single- and multi-layer transformers, and discuss implications for designing efficient transformers.

Matthias Troyer (Microsoft)

"Accelerating scientific discovery with Azure Quantum Elements"

Chemical and materials science impacts 96% of manufactured goods and 100% of humanity. Advancements in this space will enable scientists to help solve many of society’s most pressing problems and unlock unprecedented growth. However, to date, technology has not been able to deliver the scale, speed and accuracy required to rapidly accelerate progress. That is all about to change. I will explore how advancements in cloud technologies, artificial intelligence, high performance computing, and quantum computing are accelerating progress for scientists around the world. As part of this, I will also share breakthroughs in molecular simulation in the cloud that are enabling new applications for computational chemists and materials scientists, advances in quantum computing, and how industrial scientists are getting started using these methods today.

Gert Aarts (U Swansea)
"From lattice field theory to machine learning and back"
Recently, machine learning has become a popular tool to use in fundamental science, including lattice field theory. Here I will report on some recent ideas, including the Inverse Renormalisation Group and quantum-field theoretical machine learning. The latter combines insights of lattice field theory and machine learning to provide a fresh perspective on (hopefully) both.

David Shih (Rutgers U)
"Machine Learning for Fundamental Physics from the Smallest to the Largest Scales"
What new particles and interactions exist beyond the Standard Model? What is the nature of dark matter? What is the origin of the universe? Essential questions of fundamental physics such as these are being confronted with an unprecedented amount of high quality data from the LHC and astronomical surveys. Powerful and cross-cutting machine learning techniques such as generative modeling, density estimation and anomaly detection are increasingly being applied to these datasets, vastly enhancing their discovery potential. In my talk, I will showcase some highlights from this ongoing machine learning revolution that span the range from the smallest scales (LHC data) to the largest scales (astronomical data). I will describe how new techniques developed for model-independent new physics searches and fast simulation at the LHC can also be applied to data from the Gaia space telescope to map out the Milky Way dark matter density, discover new stellar streams, and upsample galaxy simulations.

Lei Wang (Chinese Academy of Sciences)
"A deep variational free energy approach to dense hydrogen"
Dense hydrogen, the most abundant matter in the visible universe, exhibits a range of fascinating physical phenomena such as metallization and high-temperature superconductivity, with significant implications for planetary physics and nuclear fusion

research. Accurate prediction of the equations of state and phase diagram of dense hydrogen has long been a challenge for computational methods. In this talk, we present a deep generative model-based variational free energy approach to tackle the problem of dense hydrogen, overcoming the limitations of traditional computational methods. Our approach employs a normalizing flow network to model the proton Boltzmann distribution and a fermionic neural network to model the electron wavefunction at given proton positions. The joint optimization of these two neural networks leads to a comparable variational free energy to previous coupled electron-ion Monte Carlo calculations. Our results suggest that hydrogen in planetary conditions is even denser than previously estimated using Monte Carlo and ab initio molecular dynamics methods. Having reliable computation of the equation of state for dense hydrogen, and in particular, direct access to its entropy and free energy, opens new opportunities in planetary modeling and high-pressure physics research.

Machine Learning Physics Project talks

Project A01

Akio Tomiya

"Machine learning for lattice field theory"

The A01 team has innovatively combined computational physics and machine learning to address challenges in lattice QCD. We've introduced a Self-learning Monte Carlo with an equivariant Transformer, enhancing simulations of complex physical systems like the double exchange model. Additionally, our gauge covariant neural networks, tailored for lattice QCD, reinterpret traditional smearing as extended residual neural networks. Lastly, we've optimized the path for gauge theories, using gauge-covariant networks to tackle the sign problem in lattice field theories. Collectively, our work showcases the transformative potential of machine learning in advancing lattice QCD simulations.

Project A02

Ahmed Hammad

"Advanced machine learning to enhance the particle collider search"

The LHC search for new physics suffers from large background contamination and the small cross section for the new physics. To alleviate these problems we are obligated to utilize deep and advanced machine learning models to single out the new physics signatures. Although there are many proposed ML for the LHC analysis we focus on the contrastive learning models.

Contrastive learning is mainly based on learning the similarity between pairs of input from the same class and dissimilarity between the pairs from different classes by mapping them in different regions in the latent space of the model. The contrastive learning, specifically Siamese network, shows a large classification performance over the other state of art ML models. Moreover, we present a multi-scale transformer model with cross attention that is able to analyze events with different momentum scales, features, etc. Finally, we construe our results by using interpretability methods like central kernel alignment (CKA) and Grad-Cam.

Project A03

Eiji Saitoh

“Deciphering quantum fingerprints in electric conductance”

When the electric conductance of a nano-sized metal is measured at low temperatures, it often exhibits complex but reproducible patterns as a function of external magnetic fields called quantum fingerprints in electric conductance. Such complex patterns are due to quantum–mechanical interference of conduction electrons; when thermal disturbance is feeble and coherence of the electrons extends all over the sample, the quantum interference pattern reflects microscopic structures, such as crystalline defects and the shape of the sample, giving rise to complicated interference. Although the interference pattern carries such microscopic information, it looks so random that it has not been analyzed. Here we show that machine learning allows us to decipher quantum fingerprints [1]; fingerprint patterns in magneto-conductance are shown to be transcribed into spatial images of electron wave function intensities (WIs) in a sample by using generative machine learning. The output WIs reveal quantum interference states of conduction electrons, as well as sample shapes. The present result augments the human ability to identify quantum states, and it should allow microscopy of quantum nanostructures in materials by making use of quantum fingerprints.

[1] S. Daimon, K. Tsunekawa, S. Kawakami, T. Kikkawa, R. Ramos, K. Oyanagi, T. Ohtsuki & E. Saitoh, Nature Com. 13, 3160, (2022).

Project A04

Koji Hashimoto

“Quantum, spacetime and machine learning”

The group A04 is searching for possible relations among quantum concepts, spacetimes and machine learning. Specifically we elaborate on two situations, (1) neural networks allows a spacetime interpretation (2) neural network output provides geometric information. We review some of the research conducted under these two directions.

Project B01

Ryo Karakida

"Understanding deep learning algorithms through learning regimes"

There are many global minima in overparameterized neural networks, and the selection of a minimum depends on the algorithm's configuration. Understanding systematically which minimum (state or "regime") can be achieved after learning would provide valuable insights for algorithm development. Here, we introduce two attempts. One is an analysis of the gradient regularization algorithm in a solvable diagonal linear network. It reveals a preferable implicit bias caused by the finite step size towards a sparse solution referred to as the rich regime. The other is the second-order optimization in the infinite width limit of deep nets. We identify the maximum update parameterization (muP) in widely-used second-order optimization methods, revealing desirable hyperparameters for the feature regime that realizes feature learning beyond the lazy regime in infinite width. It enables us to choose the hyperparameters that work across different widths.

Project B02

Takashi Takahashi

"Mean-field analysis of self-training with pseudo-labels"

Self-training (ST) is a widely used approach in semi-supervised learning. ST assigns pseudo-labels to unlabeled data points based on the model's own predictions, and then updates the model by fitting it to the pseudo-labels. This updating step is iterated several times. Unfortunately, the theoretical foundations for designing an effective update scheme in ST remain unclear, although the generalization performance of the final model depends on the details of the update scheme used at each stage, such as the size of the unlabeled data and the method for constructing the pseudo labels.

To deepen our understanding of ST, we provide a mean-field theory of ST in a solvable setup. Specifically, we consider the ST of the linear model that minimizes the ridge-regularized cross-entropy loss when the data are generated from a two-component Gaussian mixture. Leveraging this mean-field theory, we investigate how the generalization performance of the model obtained through ST is influenced by the details of the update methods at each step. Our findings demonstrate that ST can achieve performance equivalent to supervised learning using the true labels in the limit of a large number of updates if the update size per iteration could be infinitesimal under vanishing regularization. Conversely, in scenarios where the update size per iteration might be large under vanishing regularization, our results reveal that ST may lead to a classification boundary incapable of correct classification.

Project B03

Syo Kamata

"Reliability test for uncertainty quantification in the machine-learning inference from the neutron-star data to the equation of state"

Understanding QCD in the regime of finite density is a vital challenge in the modern nuclear physics, and neutron stars by far provide the most reliable and robust constraints on the dense matter EoS. In this talk, we introduce our progress based on a neural network to estimate the EoS from experimentally observed MR probability distributions.