

DEEP FRIDAYS

Seminars and meetings with D.L. researchers from academia and industries

Spring Session 2023: Young Researchers around the World

Organizer: Prof. **Andrea Asperti**

When: all Fridays at 17 p.m.

Where: on line, via Teams: [Click here to join the meeting](#)

Calendar

April 14 2022

speaker: **Hans Kersting**

affiliation: **Yahoo Research**

title: The beneficial role of stochastic noise in SGD

abstract: The data sets used to train modern machine-learning models are often huge, e.g. millions of images. This makes it too expensive to compute the true gradient over all data sets. In each gradient descent (GD) step, a stochastic gradient is thus computed over a subset (“mini-batch”) of data. The resulting stochastic gradient descent (SGD) algorithm, and its variants, is the main workhorse of modern machine learning. Until recently, most machine-learning researchers would have preferred to use GD, if they could, and considered SGD only as a fast approximation to GD. But new research suggests that the stochasticity in SGD is part of the reason why SGD works so well. In this talk, we investigate multiple theories on the advantages of the noise in SGD, including better generalization in flatter minima (“implicit bias”) and faster escapes from difficult parts of the landscapes (such as saddle points and local minima). We highlight how correlating noise can help optimization and zoom in on the question which noise structure would be optimal for SGD.

<https://proceedings.mlr.press/v162/orvieto22a.html>,

<https://arxiv.org/pdf/2206.04613.pdf>,

https://openreview.net/forum?id=cNrglG_OAeu

bio: Hans Kersting is a research scientist at Yahoo research in the scalable ML team. Previously he worked on stochastic optimization for deep neural networks at INRIA/ENS Paris and received a PhD in Bayesian machine learning from Max-Planck-Institute Tübingen. He co-authored a text book on probabilistic numerics, a discipline that treats computation as Bayesian inference. His current work is

focused on scalable modeling and training in supervised learning with large data sets.

April 21 2022

speaker: **Sam Foreman**

affiliation: **Argonne National Laboratory**

title: Generative Modeling and Efficient Sampling Techniques for Lattice Gauge Theories

abstract: In this talk we describe how ideas from lattice gauge theory have contributed to recent advances in generative modeling, and look at some ongoing work in this direction. In particular, we consider the problem of efficiently sampling configurations from probability distributions in high-dimensions, and highlight some of the major challenges. We will look at some of the common problems experienced when scaling models to modern supercomputers with thousands of GPUs, and what's next for large-scale distributed training.

bio: Sam Foreman is a computational scientist with a background in high energy physics at the Argonne Leadership Computing Facility. He is generally interested in the application of machine learning to computational problems in physics, particularly within the context of high-performance computing. Sam's current research focuses on using deep generative modeling to help build better sampling algorithms for simulations in lattice gauge theory. Personal website: <https://samforeman.me>

April 28 2022

speaker: **Dieter Büchler**

affiliation: **Max Planck Institute for Intelligent Systems**

title: Dynamic Real-world Reinforcement Learning through muscular robots

abstract: Despite decades of robotics research, current robots still struggle to acquire general and flexible dynamic skills on human level. Tasks, such as table tennis, represent this set of dynamic problems that appear easy to learn for humans but pose a steep challenge for anthropomorphic robots. In this talk, I will argue that hardware plays a crucial role for the generation of such skills. In particular, muscular actuation (i) enables robust long-term training such as required with reinforcement learning and (ii) fail-safe execution of explosive motions that allow robots to safely explore in dynamic regimes. Stay tuned for table tennis playing, ball smashing and precisely controlled soft muscular robots.

bio: Dieter Büchler is a research group leader at the Max Planck Institute for Intelligent Systems within the Empirical Inference department in Tübingen (Germany) led by Bernhard Schölkopf. Previously, Dieter received his Ph.D. from the Technische Universität Darmstadt in 2019 under the supervision of Jan Peters, an M.Sc. in Biomedical Engineering from the Imperial College London, and a B.Eng. in Information and Electrical Engineering from the HAW Hamburg (Germany). His research interests are in the areas of control, machine learning and robotics for soft robots.

May 5 2022

speaker: Konstantin Vonzhenin

affiliation: LCQB - Sorbonne University

title: Sequence-based Deep Learning methods to predict protein-protein interaction networks

abstract: The recent development of a wide spectrum of Deep Learning methods has allowed rapid progress in the field of Bioinformatics. In particular, the Natural Language Processing models allow extracting important information directly from amino acid sequences of proteins. This feature can be used in order to predict interactions between various proteins without using any structural information at the level of input. This is essential in cases where protein structures are either unknown or can not be formed due to the instability of a protein. The main goal of this project is to be able to use only sequential information in order to make generalized predictions of interactions between different species in a single community.

bio: Konstantin is a PhD student at Sorbonne University under the supervision of Alessandra Carbone. His research interests are centered around applications of Deep Learning, particularly in Bioinformatics. Having a General Physics background from his undergraduate studies, he previously obtained a Master's Degree in Computational Neuroscience and published a work titled "Multilevel development of cognitive abilities in an artificial neural network" which shows what mechanisms are essential for our brain to perform complex conscious processing of information.

May 12 2022

speaker: Mariana Clare

affiliation: European Centre for Medium-Range Weather Forecasts (ECMWF)

title: Could deep learning methods replace numerical weather models?

abstract: Over the last year, there has been a rise in deep learning methods being used to produce highly accurate weather predictions. These methods include Graph Neural Networks and Recurrent Neural Networks, amongst others, and have been applied by some of the world's leading tech companies. Some works have made claims that these deep learning methods are more accurate than the existing state-of-the-art numerical weather models. But is this claim justified? And if it's not, will there be deep learning methods in the future that can provide a better forecast than numerical models? In this talk, I will discuss the methods used and their advantage and limitations over existing numerical weather models.

bio: Mariana is a researcher at ECMWF working on applying machine learning techniques to assess uncertainty in weather forecasts. In particular she focuses on building trustworthy and interpretable techniques that can thus be used operationally in day-to-day forecasts. She recently received a PhD from Imperial, focussing on developing and using advanced numerical and statistical techniques to assess erosion and flood risk in coastal zones. By training she is a mathematician, having done her undergraduate degree in Mathematics at the University of Oxford.

May 19 2022

speaker: **Giorgia Ramponi**

affiliation: **ETH AI Center**

title: Multi-Agent Reinforcement Learning

abstract:

bio: Giorgia is a Postdoctoral Researcher at ETH AI Center, advised by prof. Niao He and prof. Andreas Krause and supported by Google Brain. Her research interests lie in machine learning and mathematical modeling, focusing on Reinforcement Learning and Multi-Agent Learning. Previously she worked on Social Network Analysis with Marco Brambilla and Stefano Ceri, and on Networking with Gaia Maselli. In June 2021, she completed her PhD in Information Technology at Politecnico di Milano (with honors), advised by Marcello Restelli defending the dissertation "Challenges and Opportunities in Multi-Agent Reinforcement Learning".