EPFL-CIS & RIKEN-AIP Joint Online Workshop on Machine Learning

Abstract DAY 2: Sep. 8, 2022 (September 7 – 8, 2022)

Switzerland 10:00 / Japan 17:00 Etienne Boursier (EPFL-CIS)

Title: Gradient flow dynamics of shallow ReLU networks for square loss and orthogonal inputs

Abstract:

The training of neural networks by gradient descent methods is a cornerstone of the deep learning revolution. Yet, despite some recent progress, a complete theory explaining its success is still missing. This article presents, for orthogonal input vectors, a precise description of the gradient flow dynamics of training one-hidden layer ReLU neural networks for the mean squared error at small initialisation. In this setting, despite non-convexity, we show that the gradient flow converges to zero loss and characterise its implicit bias towards minimum variation norm. Furthermore, some interesting phenomena are highlighted: a quantitative description of the initial alignment phenomenon and a proof that the process follows a specific saddle to saddle dynamics.

Bio:

Etienne Boursier completed his PhD at ENS Paris-Saclay in September 2021, under the supervision of Vianney Perchet and entitled “Statistical Learning in a strategical environment”. During his PhD, he studied multi-agent learning, combining (online) learning with game theoretical tools. In particular, he mainly focused on the problem of Multiplayer Multi-armed bandits, but also worked on other bandits related problems, Social Learning and Utility/Privacy trade-off. Since October 2021, he is a postdoc in the Theory of Machine Learning Lab led by Nicolas Flammarion at EPFL. He is currently focusing on multitask/meta-learning and also providing theoretical insights on the empirical success of nonlinear neural networks.

Switzerland 10:30 / Japan 17:30 Chao Li (RIKEN-AIP)

Title: Structure Search for Tensor Network Learning

Abstract:

Recent works put much effort into tensor network structure search (TN-SS), aiming to select good tensor network structures involving TN-ranks, formats, and so on, for improving the performance of TNs on machine learning tasks. In this presentation, we will focus on the TN-SS problem and talk about the following three questions: 1) what is TN-SS, and what motives the studies on this issue; 2) how to resolve it, and what we need to pay for the searching; 3) how much benefit can we achieve from TN-SS for machine learning? We will first show that TN-SS can be modeled as a combinatorial optimization problem. Then, two searching algorithms, TNGA and TNLS, will be introduced to solve the problem. Last, several applications will be discussed to demonstrate the potential benefit of TN-SS for machine learning. Related works were published in (Li et al., ICML’20, ICML’22).

Bio:

Dr. Chao Li is currently an indefinite-term research scientist with the AIP center, RIKEN institute, since 2021. Before that, he was a post-doctoral researcher with RIKEN-AIP from 2018 to 2020. He obtained his bachelor's and Ph. D. degrees at Harbin Engineering University (HEU) in China in 2006 and 2017, respectively. He regularly serves as a (senior) reviewer of ICML, NeurIPS, IJCAI, AAAI, and so on. His research interests include tensor network and machine learning.

Switzerland 11:00 / Japan 18:00 Luca Viano (EPFL-CIS)

Title: Proximal Point Imitation Learning

Abstract:

This work develops new algorithms with rigorous efficiency guarantees for infinite horizon imitation learning (IL) with linear function approximation without restrictive coherence assumptions. We begin with the minimax formulation of the problem and then outline how to leverage classical tools from optimization, in particular, the proximal-point method (PPM) and dual smoothing, for online and offline IL, respectively. Thanks to PPM, we avoid nested policy evaluation and cost updates for online IL appearing in the prior literature. In particular, we do away with the conventional alternating updates by the optimization of a single convex and smooth objective over both cost and Q-functions. When solved inexactly, we relate the optimization errors to the suboptimality of the recovered policy. As an added bonus, by re-interpreting PPM as dual smoothing with the expert policy as a center point, we also obtain an offline IL algorithm enjoying theoretical guarantees in terms of required expert trajectories. Finally, we achieve convincing empirical performance for both linear and neural network function approximation.

Bio:

I am currently a ELLIS PhD student at EPFL advised by Volkan Cevher and co-advised by Gergely Neu working on optimization methods for reinforcement and imitation learning. Previously, I obtained my MSc in Computational Science from EPFL and my BSc in Physics Engineering from Politecnico di Torino.

Abstract DAY 2: Sep. 8, 2022

Switzerland 11:30 / Japan 18:30 Matthias Weissenbacher (RIKEN-AIP)
Title: Reinforcement Learning via Symmetries of Dynamics

Abstract:
Offline reinforcement learning (RL) leverages large datasets to train policies without interactions with the environment. The learned policies may then be deployed in real-world settings where interactions are costly or dangerous. Current algorithms over-fit to the training dataset and as a consequence perform poorly when deployed to out-of-distribution generalizations of the environment. We aim to address these limitations by learning a Koopman latent representation which allows us to infer symmetries of the system's underlying dynamic. The latter is then utilized to extend the otherwise static offline dataset during training; this constitutes a novel data augmentation framework which reflects the system's dynamic.

Bio:
Currently I am a Research Scientist at RIKEN AIP with the focus on generalization in Reinforcement Learning (RL). My background is however in mathematical physics in which I have completed my Master at ETH Zurich and then my PhD at the Max-Planck Institute for Physics in Munich with a focus on Sting theory. During my first Postdoc at the University of Tokyo (Kavli IPMU) my research interests shifted towards the vibrant field of AI. These days my main ambition is to evolve RL-algorithms to be able to generalize to new tasks.

Switzerland 12:00 / Japan 19:00 Sebastian Neumayer (EPFL-CIS)
Title: Lipschitz Function Approximation using DeepSpline Neural Networks

Abstract:
In this talk, we investigate NNs with prescribed bounds on the Lipschitz constant. One possibility to obtain Lipschitz-constrained NNs is to impose constraints on the architecture. Here, it turns out that this significantly limits the expressivity if we use the popular ReLU activation function. In particular, we are unable to represent even simple continuous piece-wise linear functions. On the contrary, using learnable linear splines instead fixes this problem and leads to maximal expressivity among all component-wise activation functions. From the many possible applications of Lipschitz-constrained NNs, we discuss one in more detail to see that the theoretical observations also transition into improved performance.

Bio:
Sebastian Neumayer studied mathematics at TU Kaiserslautern and received his PhD in 2020 at TU Berlin under the supervision of Gabriele Steidl. Currently, he is a postdoctoral researcher in the Biomedical Imaging Group at EPFL. His main research interests are centered around convex analysis, inverse problems, and theoretical aspects of neural networks. In the past months, he has focused on studying stability properties of neural networks and designing new network architectures.

Switzerland 12:30 / Japan 19:30 Koichi Tojo (RIKEN-AIP)
Title: A method to construct exponential family by representation theory

Abstract:
Exponential families play an important role in the field of information geometry, statistics and machine learning. By definition, there are infinitely many exponential families. However, only a small part of them is widely used. We want to give a framework to deal with these "good" families systematically. In light of the observation that the sample spaces of most of them are homogeneous spaces of certain Lie groups, we proposed a method to construct exponential families on homogeneous spaces by taking advantage of representation theory in [1]. This method generates widely used exponential families such as normal, gamma, Bernoulli, categorical, Wishart, von Mises-Fisher, and hyperboloid distributions. In this talk, we will explain the method and its properties. [1] K. Tojo, T. Yoshino, A method to construct exponential families by representation theory, arXiv:1811.01394

Bio:
Koichi Tojo is a postdoctoral researcher in RIKEN AIP at Mathematical Science Team. He received Ph.D. (Mathematical Science) under Prof. Taro Yoshino from the University of Tokyo in 2018. His research interests include representation theory, Lie group theory, information geometry and machine learning.

Abstract DAY 2: Sep. 8, 2022

Switzerland 13:00 / Japan 20:00 Raphael Reinauer (EPFL-CIS)
Title: The Topological BERT: Transforming Attention into Topology for Natural Language Processing

Abstract:
In recent years, the introduction of the Transformer models sparked a revolution in natural language processing (NLP). BERT was one of the first text encoders using only the attention mechanism without any recurrent parts to achieve state-of-the-art results on many NLP tasks. This talk introduces a text classifier using topological data analysis. We use BERT's attention maps transformed into attention graphs as the only input to that classifier. The model can solve tasks such as distinguishing spam from ham messages, recognizing whether a sentence is grammatically correct, or evaluating a movie review as negative or positive. It performs comparably to the BERT baseline and outperforms it on some tasks. Additionally, we propose a new method to reduce the number of BERT's attention heads considered by the topological classifier, which allows us to prune the number of heads from 144 down to as few as ten with no reduction in performance. Our work also shows that the topological model displays higher robustness against adversarial attacks than the original BERT model, which is maintained during the pruning process. To the best of our knowledge, this work is the first to confront topological-based models with adversarial attacks in the context of NLP.

Bio:
Raphael Reinauer is currently a Postdoctoral Fellow in the Laboratory for Topology and Neuroscience at the École Polytechnique Fédérale de Lausanne in Switzerland. He completed his Ph.D. in Mathematics at the University of Münster in 2020 under the supervision of Michael Joachim. His research focuses on applied topological data analysis, machine learning, and natural language processing.

Switzerland 13:30 / Japan 20:30 Tomasz M. Rutkowski (RIKEN-AIP)
Title: Machine Learning Approaches for EEG-derived Early-onset Dementia Neuro-biomarker Development

Abstract:
Brain-computer interface (BCI) and efficient machine learning (ML) algorithms belonging to the so-called `AI for social good' domain contribute to the well-being improvement of patients with limited mobility or communication skills. We will review our recent results from a project focusing on developing a dementia digital neuro-biomarker for early-onset prognosis of a possible cognitive decline utilizing a passive BCI approach. We will report findings from elderly volunteer pilot study groups analyzing EEG responses in a classical short-term memory evaluating oddball paradigms, reminiscent images, and emotional evaluation implicit learning tasks. Results using feature engineering approaches using signal complexity/criticality and information geometry employing Riemannian geometry tools, as well as end-to-end training machine models, will be discussed. The reported pilot studies showcase the vital application of artificial intelligence (AI) for an early-onset mild cognitive impairment (MCI) prediction in the elderly.

Bio:
Tomasz (Tomek) M. RUTKOWSKI is an applied AI/ML research scientist at the RIKEN Center for Advanced Intelligence Project (AIP), a research fellow at The University of Tokyo, Japan and Nicolaus Copernicus University in Torun, Poland. Tomasz's research interests include computational neuroscience, primarily passive brain-computer interfacing (BCI) dementia biomarkers elucidation, computational modeling of brain processes, and AI for social good applications. More information and publications are available at http://tomek.bci-lab.info/