

## 5th Summer School

# MATH@NTUA

MATHEMATICS OF MACHINE AND STATISTICAL LEARNING

ATHENS, 19-23 JUNE 2023

### ABSTRACTS

1. **Constantine Caramanis** (University of Texas, Austin & Archimedes Research Unit)

**Title:** Online Learning

**Abstract:** In this session, we turn to dynamic problems in machine learning, where we must simultaneously learn and make decisions over time. In the so-called batch ML setup, we see all our data at once, and must find a single decision that fits the data as well as possible. Yet in many applications, data are revealed sequentially over time, and we must do our best at each point in time to take the best action. Though we will not focus on applications, there are in fact many and diverse applications of online learning. Important applications include dynamic pricing, network routing, portfolio management, and real-time bidding in advertising.

This mini-course will cover three key ideas of online learning. We will first discuss online gradient descent and variants. Next, we will cover the basic ideas from follow-the-leader type algorithms. Finally, we will move to multi-armed bandits, and will cover the basic results for adversarial and stochastic bandits.

The ideas from convex and non-convex optimization you learned from Panayiotis Mertikopoulos will be directly relevant for much of this mini-course.

2. **Daniele Durante** (Università Bocconi, Milano)

**Title:** Statistical Learning of Networks

**Abstract:** Network data encoding complex relationship structures among a set of entities are ubiquitous in several disciplines covering, among others, social sciences, neurosciences, economics, ecology and genetics. Although the field of network science provides nowadays a complete set of models and methods for studying complex connectivity structures, the relevance of such data and the challenges associated with the analysis of modern networks, motivate still active and ongoing research in this field. The scope of this short course is to provide an overview of both classical and more recent algorithmic strategies and generative models which enable statistical learning of complex network structures. These will include, among others, community detection algorithms, force-directed placement solutions for graph drawing, exponential random graph models, stochastic block models and latent space models. The computational methods associated with these formulations will be also discussed and the practical performance will be illustrated via a number of real-world applications from political sciences, neuroscience and criminology.

3. **Panayotis Mertikopoulos** (National and Kapodistrian University of Athens, Archimedes Research Unit)

**Title:** Introduction to optimization for machine learning

**Abstract:** The quality of a machine learning model depends to a large extent on the optimization algorithms that have been used for its training. In this series of introductory lectures, we will examine a range of optimization algorithms – both stochastic and deterministic, for both convex and non-convex problems – and study their basic theoretical guarantees. Specifically, we will start with an overview of the classical theory of gradient descent in convex programming problems, the accelerated gradient descent algorithm of Nesterov, as well as the applications of the theory in stochastic and batch learning problems. Next, we will introduce non-convex programming problems, where we will study the convergence properties of the above algorithms (guarantees of avoiding saddle points, convergence rate, etc.), both in their deterministic and stochastic versions. Finally, if time permits, we will discuss more specific topics such as adaptive algorithms (like AdaGrad), second-order methods (Newton’s algorithm, etc.), and/or applications to min-max type problems

4. **Konstantinos Spiliopoulos** (Boston University)

**Title:** Normalization effects on deep neural networks and deep learning for scientific problems

**Abstract:** We study the effect of normalization on the layers of deep neural networks. A given layer  $i$  with

$N_i$  hidden units is allowed to be normalized by  $1/N_i^{\gamma_i}$  with  $\gamma_i \in [1/2, 1]$  and we study the effect of the choice of the  $\gamma_i$  on the statistical behavior of the neural network's output (such as variance) as well as on the test accuracy on the MNIST and CIFAR10 data sets. We find that in terms of variance of the neural network's output and test accuracy the best choice is to choose the  $\gamma_i$ 's to be equal to one, which is the mean-field scaling. We also find that this is particularly true for the outer layer, in that the neural network's behavior is more sensitive in the scaling of the outer layer as opposed to the scaling of the inner layers. The mechanism for the mathematical analysis is an asymptotic expansion for the neural network's output. An important practical consequence of the analysis is that it provides a systematic and mathematically informed way to choose the learning rate hyperparameters. Such a choice guarantees that the neural network behaves in a statistically robust way as the  $N_i$  grow to infinity.

Time permitting, I will discuss applications of these ideas to design of deep learning algorithms for scientific problems including solving high dimensional partial differential equations (PDEs), closure of PDE models and reinforcement learning.

**Relevant papers:**

- 1) K. Spiliopoulos and J. Yu, "Normalization effects on deep neural networks", 2023, AIMS Journal on Foundations of Data Science, 2023, Volume 5, Issue 3: pp. 389-465, arXiv: <https://arxiv.org/abs/2209.01018>, github: [https://github.com/kspiliopoulos/NENN\\_Deep](https://github.com/kspiliopoulos/NENN_Deep)
- 2) K. Spiliopoulos and J. Yu, "Normalization effects on shallow neural networks and related asymptotic expansions", 2021, AIMS J. on Foundations of Data Science , June 2021, Vol. 3, Issue 2, pp. 151-200, arXiv: <https://arxiv.org/abs/2011.10487>, github: [https://github.com/kspiliopoulos/NENN\\_Shallow](https://github.com/kspiliopoulos/NENN_Shallow)
- 3) Justin Sirignano, K. Spiliopoulos and Jonathan MacArt, "PDE-constrained Models with Neural Network Terms: Optimization and Global Convergence", 2023, Journal of Computational Physics, Vol. 481, 15 May 2023, 112016, arXiv: <https://arxiv.org/abs/2105.08633>
- 4) Justin Sirignano, K. Spiliopoulos, "Asymptotics of Reinforcement Learning with Neural Networks", Stochastic Systems, Vol. 12, No. 1, March 2022, pp. 2-29, arXiv: <https://arxiv.org/abs/1911.07304>
- 5) Justin Sirignano, K. Spiliopoulos, "DGM: A deep learning algorithm for solving partial differential equations", 2018, J. of Comp. Physics, Vol. 375, pp. 1339-1364, arXiv: <https://arxiv.org/abs/1708.07469>
- 6) Justin Sirignano, K. Spiliopoulos, "Stochastic gradient descent in continuous time", 2017, SIAM J. on Financial Mathematics, Vol. 8, Issue 1, pp. 933-961, arXiv: <https://arxiv.org/abs/1611.05545>

5. **Costas Smaragdakis** (National Technical University of Athens and University of Crete)

**Hands-On Training:** Learning the Black-Scholes price formula

**Abstract:** This hands-on tutorial offers a step-by-step approach to replicating the Black-Scholes pricing formula through a neural network model. Beginning with a brief overview of option pricing and the Black-Scholes model, we will implement a programming code to train the model using a dataset of Black-Scholes formula inputs and corresponding option prices. The tutorial focuses on practical implementation, enabling participants to understand basic concepts of the process behind the adaptation of the trainable parameters of simple models to estimate complex formulas.

*Participants can bring their laptops along and follow the tutorial live.*

	Monday	Tuesday	Wednesday	Thursday	Friday
09:30-11:00	Durante	Mertikopoulos	Durante	Durante	Spiliopoulos
11:00-11:30	Break	Break	Break	Break	Break
11:30-13:00	Mertikopoulos	Caramanis	Caramanis	Smaradgakis	Spiliopoulos
13:00-14:30	Lunch	Lunch	Lunch	Lunch	Lunch
14:30-16:00	Mertikopoulos	Mertikopoulos	Caramanis	Q&A session	Smaradgakis
16:00-			Caramanis		



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ-ΣΕΜΦΕ



H.F.R.I.  
Hellenic Foundation for  
Research & Innovation