universität wien

Fakultät für Mathematik

## Vorträge
## Tenure Track „Machine Learning“

**Mittwoch, 27. März 2019, Hörsaal 13**

**Rima Al-Aifari**
**(ETH Zürich)**

### 9:00 Uhr – 9:20 Uhr: Didaktischer Vortrag

**„Lineare Algebra und Geometrie, Analysis, Numerische Mathematik, Programmieren, Wahrscheinlichkeitstheorie und Statistik“**

In this first course on machine learning, we introduce and discuss the three major concepts of machine learning: data, model and training. We then go on to motivating the study of machine learning by considering one of its four pillars that will be discussed in this course: dimensionality reduction. To do so we present a specific problem of dimensionality reduction and discuss how it can be solved.

### 9:50 Uhr – 10:35 Uhr: Wissenschaftlicher Vortrag

**„Three unstable problems in imaging“**

We introduce two problems of reconstructing objects from measured data and one conceptually different problem of machine learning: image classification. While these problems are very different from a theoretical point of view, they have in common that they are unstable in nature: small perturbations of the input may result in unpredictable changes in the output. For the first two problems, we provide a mathematical analysis of the inversion problem and possible remedies for regularization. For the third problem, we discuss the different sources of instabilities and suggest alternative learning procedures to achieve better robustness results.

---

**Donnerstag, 28. März 2019, Seminarraum 14**

**Philipp Christian Petersen**
**(University of Oxford)**

### 15:00 Uhr – 15:20 Uhr: Didaktischer Vortrag

**„Empirical Risk Minimisation with Finite Model Classes“**

Based on the example of weather forecasting, we will introduce the basic notions of statistical learning theory including empirical risk minimisation. Next, we will identify overfitting as one of the main issues of empirical risk minimisations. Finally, we will prove and discuss a result demonstrating that with high probability, overfitting is avoided if the underlying hypothesis classes are finite and sufficiently many data points were taken.

### 15:50 Uhr – 16:35 Uhr: Wissenschaftlicher Vortrag

**„Deep Neural Networks and Partial Differential Equations: Approximation Theory and Structural Properties“**

Novel machine learning techniques based on deep learning have achieved remarkable results in many areas such as image classification, game intelligence, or speech recognition. Driven by these successes, many scholars have started using them in areas which are not traditionally associated with machine learning. For instance, more and more researchers are employing deep neural networks to develop tools for the discretisation and solution of partial differential equations. Two reasons can be identified to be the driving forces behind the increased interest in neural networks in the area of the numerical analysis of PDEs. On the one hand, powerful approximation theoretical results have been established which demonstrate that neural networks can represent functions from the most relevant function classes with a minimal number of parameters. On the other hand, highly efficient machine learning techniques for the training of these networks are now available and can be used as a black box. In this talk, we will give an overview of some approaches towards the numerical treatment of PDEs with neural networks and study the two aspects above. We will recall some classical and some novel approximation theoretical results and tie these results to PDE discretisation. Afterwards, providing a counterpoint, we analyse the structure of network spaces and deduce considerable problems for the black box solver. In particular, we will identify a number of structural properties of the set of neural networks that render optimisation over this set especially challenging and sometimes impossible.

---

**Freitag, 29. März 2019, Hörsaal 15**

**Quentin Berthet**
**(University of Cambridge)**

**9:00 Uhr – 9:20 Uhr: Didaktischer Vortrag**

**„Gradient descent for convex optimization“**

We will introduce the algorithm of gradient descent for unconstrained minimization of convex functions. We will prove some mathematical guarantees for this algoritm, give some illustrations on simple functions, and possible applications and extensions, to the field of machine learning.

**9:50 Uhr – 10:35 Uhr: Wissenschaftlicher Vortrag**

**„Theory and applications in statistical learning problems“**

Modern techniques in data storage and acquisition have brought forth a new scientific age, driven by information. While this profound change opens the door to scientific advances in many disciplines, it also poses to mathematicians the challenge of extracting valuable information from this inundation of data. As datasets are becoming larger and more complex, new statistical techniques are required. Designing these new methods, and giving theoretical guarantees about their reliability and efficiency, also requires new mathematical ideas.

In this presentation, I will present some recent results on modern techniques for statistical learning problems, including applications to natural language processing.

---

**Freitag, 29. März 2019, Hörsaal 11**

**Mathias Staudigl**
**(Maastricht University)**

**15:00 Uhr – 15:20 Uhr: Didaktischer Vortrag**

**„20 minutes on statistical learning“**

In this course I will introduce students into the general statistical machine learning framework commonly known as the Probably Approximately Correct (PAC) learning model. My aim is to introduce students to the basic concepts of statistical machine learning, and to highlight its connections with optimization, game theory and information theory. To do so, I focus mostly on the most basic, but important, binary classification learning problem. The course will be self-contained, and uses techniques on probabilistic concentration inequalities, which, however, will be explained during the lecture. No specific background knowledge is required.

**15:50 Uhr – 16:35 Uhr: Wissenschaftlicher Vortrag**

**„Mirror Descent Learning“**

”Learning” can be very loosely defined as the ”ability to improve performance after observing data”. Over the past two decades, there has been an explosion of both applied and theoretical work on machine learning. Applications of learning methods are ubiquitous: they include systems for face detection and face recognition, prediction of stock markets and weather patterns, speech recognition, learning user’s search preferences, placement of relevant ads, and much more. The success of these applications has been paralleled by a well-developed theory. I shall call this latter branch of machine learning ”learning theory”. The main challenge in learning theory is to device forecasting rules that are as good on average as a forecaster designed under the fictitious assumption that the data are known a-priori. The performance measure involved here is distribution-free, as it acts at the level of individual sequences, and is known as the regret of the learner. The effective construction of such low-regret forecasters involves learning algorithms, and uses ideas from optimization and game theory. In this talk I give an overview on recent advances on learning algorithms based on mirror descent. In the first part of the talk, I give a short description of the learning setting, and its relation to functional aggregation techniques in non-parametric statistics, and dual averaging in optimization. I will then present recent results on online regret minimization based on mirror descent in the adversarial and multi-player online learning setting of statistical machine learning. I will spend some time explaining this fundamental concept in statistical learning theory, as well as its close connection to typical merit functions employed in (online) convex optimization. Departing from here, I will present recent results how to transform learning rules guaranteeing small regret into learning rules guaranteeing small dynamic regret, a generalization recently proposed in a multi-player non-stationary environment [1]. The second half of the lecture will establish connections between regret learning and dynamical system approaches in convex

optimization, culminating in the class of stochastic mirror descent dynamics (SMD) [2, 3]. SMD is a random projection-differential dynamical system, which has received recently tremendous interest in the machine learning community. We will give an overview on the long-run behavior of this dynamical systems, by establishing close shadowing properties with the corresponding deterministic system using the theory of asymptotic pseudo trajectories, as well as derive some localization estimates for occupation measures. We also give an overview on recent generalizations of SMD in the machine learning community. The last part of the talk establishes the connection between mirror descent dynamics and the recently developed class of algorithms, called Hessian-barrier techniques [4].

References

[1] Duvocelle, B., Mertikopoulos, P., Staudigl, M., & Vermeulen, D. (2018). Learning in time-varying games. arXiv preprint arXiv:1809.03066.
[2] Mertikopoulos, P., & Staudigl, M. (2018). On the convergence of gradient-like flows with noisy gradient input. SIAM Journal on Optimization, 28(1), 163-197.
[3] Mertikopoulos, P., & Staudigl, M. (2018). Stochastic mirror descent dynamics and their convergence in monotone variational inequalities. Journal of optimization theory and applications, 179(3), 838-867.
[4] Bomze, I. M., Mertikopoulos, P., Schachinger, W., & Staudigl, M. (2018). Hessian barrier algorithms for linearly constrained optimization problems. arXiv preprint arXiv:1809.09449.