Machine Learning meets Environmental Science

Friday, the 25th of September 2020: 9:00-14:00

- **Meeting Venue:**
  Neue Aula, Audimax, Geschwister-Scholl-Platz
- **Organizers**
  Prof. Martin Butz (ML Cluster, Dep. Of Computer Science)
  Prof. Christiane Zarfl (Center for Applied Geoscience)
- **Registration is required by Email (cm-sekretariat@inf.uni-tuebingen.de) until Sept 22, 2020**

Program

9:00 Welcome and Introduction: Cluster of Excellence – Machine Learning for Science
Philipp Berens

9:10 ML Transfer Center and simulation based inference
Álvaro Tejero-Cantero & Jakob Macke

9:40 Day-ahead optimization of production schedules for saving electrical energy costs
Thomas Stüber & Michael Menth

10:00 Uncovering hidden structure in climate data
Bedartha Goswami

10:20 Coffee Break

10:40 Learning spatiotemporal distributed generative graph neural networks
Martin Butz

11:00 Modeling environmental processes in rivers
Christiane Zarfl

11:20 Can plants learn? Coupling models and data in eco-evolutionary research
Sara Tomiolo (with Katja Tielbörger)

11:40 Improving the spatial prediction of soil organic carbon content in two contrasting climatic regions by stacking machine learning models and rescanning covariate space
Ruhollah Taghizadeh-Mehrjardi (with Thomas Scholten)

12:00 Snack Break

12:30 ML and mobile robots in environmental science
Andreas Zell

12:50 Turbulent transport of energy, momentum and matter by large data sets obtained from airborne probing of the lower atmosphere
Jens Bange

13:20 Status of the CRC 1253 CAMPOS - Catchments as Reactors: Metabolism of Pollutants on the Landscape Scale
Christiane Zarfl

13:40 Funding options in the ML Cluster of Excellence and beyond
Tilman Gocht

14:00 End

Aim of the day:

Environmental science studies spatiotemporal dynamics of various processes and on different topics, including climate and weather, geology, hydrology, vegetation and agriculture, various forms of pollution (e.g. of organic pollutants), to name just a few. In all these cases, multiple, often interrelated data sources are available at varying degrees of spatial and temporal granularity. Moreover, human activities, such as river dam building, CO2 release, plantations, industry etc., strongly influence the unfolding dynamics. Critical principles – such as basic laws in physics – apply universally in such systems. Environmental science has strong expertise in modeling the underlying processes – typically by systems of partial or ordinary differential equations.

As a result, besides the expertise about the underlying processes, environmental science offers two types of data – real-world data as well as data from the respectively available models of the considered environmental system. This offers essentially the perfect basis for a meaningful, science-driven application of ML algorithms. On the one hand, the parameters of the differential equations may be optimized more effectively by means of state-of-the-art gradient-based approximation approaches from ML. On the other hand, the available models may be augmented or fully substituted by distributed spatiotemporal, generative neural network approaches, such as convolutional networks, graph networks, autoencoders, recurrent neural networks, and combinations thereof.

Seeing that models are available to pre-train and analyze potentially applicable ML architectures, expertise is available to tune these models to the actual underlying processes, and that real world data is available to further train and test the generalizability of these ML architectures, it is time that ML meets Environmental Science! The aim is to foster collaboration with a focus on two main potential strands. First, available models of differential equations and involved prior situation assumptions may be optimized by means of state-of-the-art ML techniques. Second, ML techniques and particularly distributed, generative artificial neural networks may be designed to infer the processes and structures that generate particular data patterns, thus enabling (i) the fast, efficient, and accurate simulation of environmental processes and (ii) the consideration of impacts of human actions, including the potential to derive optimal actions for steering the environmental system towards a desired (stable / homeostatic) direction.