Joint E-MRS/MRS Tutorial on “Artificial Intelligence for Advancing Materials Science”

Workshop coordinators:

- E-MRS: Stefan Sandfeld (Micromechanical Materials Modelling, University of Freiberg, Germany)
- MRS: Kristofer Reyes (Materials Design and Innovation, University at Buffalo, USA)

Workshop instructors:

- Stefan Sandfeld (Micromechanical Materials Modelling, University of Freiberg, Germany)
- Ehrenfried Zschech (Fraunhofer Institute for Ceramic Technologies and Systems IKTS, Germany)
- Kristofer Reyes (Materials Design and Innovation, University at Buffalo, USA)
- Yexiang Xue (Department of Computer Science, Purdue University, USA)

General description:

The joint E-MRS/MRS tutorial on “Artificial Intelligence (AI) for Advancing Materials Science” aims to introduce scientists and engineers from the field of materials science to a novel computer based planning and execution of experiments as well as approach of data analysis in order to support the research and development of new materials and processes. It is believed that AI-based working techniques will significantly change the way how research and development will be carried out in the future. The aim of the workshop is to acquaint the participants with state-of-the-art methods applied in AI-based research and development.

Importance of AI-based working methods for young researchers and their careers:

In recent years, we have seen exciting progress in artificial intelligence (AI). AI systems are now reaching human-level and even superhuman-level performance on a range of tasks, such as speech recognition, image interpretation, machine translation (Google translate), and gameplay (DeepBlue for Chess, Watson for Jeopardy, and AlphaGo for Go). There is a general belief that AI is poised to radically transform many components of our society and economy. Just to mention one example, self-driving cars and trucks, which incorporate real-time image recognition and control, are close to becoming a reality.

Given these advances in AI and machine learning (ML), the scientific community has taken note and is exploring the use of AI for scientific discovery. Deep learning and AI reasoning methods enable scientists to uncover new types of structure in large amounts of data and design new experiments leading to the most promising areas for further experimentation. These techniques also open up new opportunities for accelerating materials research and discovery. To reap the full potential of these developments, it is critical to educate the materials science community, in particular students and young researchers: Not only can AI dramatically accelerate the pace of materials discovery; AI can also reduce the cost of scientific discovery.

Outline for the workshop:

The most promising AI approaches for scientific discovery, in general, and materials science, in particular, involves a combination of AI techniques: Machine learning, deep learning, search and optimization, reasoning, knowledge representation, and decision making. These techniques should be combined with human insights, simulations, and experimentation.

Schedule: each lecture 45 minutes: 2 lectures + coffee break + 2 lectures + lunch break + 2 lectures + coffee break + 2 lectures

1 Introduction to AI and Materials Data Science: (Stefan Sandfeld, 9.00-9.45)

The workshop begins with a general overview of the emerging field of Materials Data Science and Materials Informatics. In particular, we start with a few short examples, show some use cases and will explain the impact of AI on Materials Science. Then, basic data science concepts and methods such as data processing techniques and exploratory data analysis will be shown together with a few first steps towards concepts from machine learning.

2 Machine Learning Methods of Supervised Learning (Stefan Sandfeld, 9.45-10.30)

In the second morning session we will then dive deeper in machine learning methods of supervised learning. This includes on the one hand parametric Models (statistical Models, classification and regression with models) and on the other hand non-parametric models (k-Nearest Neighbors, Decision Trees, Support Vector Machines, etc) will be introduced.

======= coffee break (10.30-10.45) ======

3 Sequential Bayesian Experimental Design (Kris Reyes, 10.45 – 11.30)

In this module, we shall discuss sequential experimental design, and the Bayesian modeling of experimental responses. Key to this is the idea of a decision-making policy. We shall give many examples of such policies for various design tasks. Topics discussed will include Gaussian Processes, Multi-armed Bandits and Bayesian Optimization.

4 Reinforcement Learning (Kris Reyes, 11.30 – 12.15)

In this module, we shall discuss how to include problem or experiment-specific structure or constraints in sequential experimental design. This will be done through the language of Markov Decision Processes (MDP) and Reinforcement Learning (RL). We also shall briefly discuss algorithms for selecting optimal experimental actions using the MDP/RL framework.

======= lunch break (12.15-12.45) ======

5 A Gentle Introduction to Deep Learning (Stefan Sandfeld, 12.45 – 13.30)

A brief introduction to artificial neural networks (Neural Network Definition and Elements, Custom layers, Activation Functions, Loss functions) as well as to Deep Neural Networks (CNN, RNN) will be given. This is complemented by examples that focus on materials science applications, such as semantic segmentation of images.

6 Application in Materials Development and Materials Characterization (Ehrenfried Zschech, 13.30 – 14.15)

The first part of this lecture is dedicated to AI applications in materials development, where the design of experiments is demonstrated based on artificial neural networks to support the classification and predictions of parameters. The second part will consider AI application in materials characterization based on examples, such as (i) the machine learning based investigation of the morphology of hierarchical materials (3D image analysis of XCT data), and (ii) the study of the structure and chemical bonding of materials using neural networks and machine learning in spectroscopic and diffraction data analysis.

======= coffee break (14.15-14.30) ======

7 AI Problem Solving, Reasoning, Randomization, and Human Computation (Yexiang Xue, 14.30 – 15.15)

In this part, I will talk about how AI can be used as a problem-solving tool for scientific discovery. AI reasoning and learning technologies provide efficient ways to navigate in a high dimensional space, which can benefit materials science since many applications require searching for candidate solutions that satisfy given properties from a large compositional space. After introducing core AI concepts, such as knowledge representation, constraint satisfaction, search, and reasoning, etc. the focus will be on a recent approach based on hashing and randomization, which makes a significant speedup possible. This will be demonstrated based on applications including the phase-map identification problem in high-throughput materials discovery.

8 Constrained Unsupervised Learning and Applications in Materials Science (Yexiang Xue, 15.15 – 16.00)

The core computational models in unsupervised learning will be introduced, such as clustering, singular value decomposition, principal component analysis, and non-negative matrix decomposition. I will also introduce neural network-based generative models based on embeddings and their applications in natural language processing and in materials science. The second part of the talk will be focused on embedding constraint reasoning into unsupervised and generative models. The necessity of constraint embedding in generating physically meaningful solutions will be shown as well as a number of approaches based on constraint programming and decision diagrams. Finally, examples from natural language processing, robotics, operational research, and materials science will be shown, in which constraint embedded generative models lead to better results.

9 Closing Remarks (Stefan Sandfeld, 16.00-16.05)