BiGmax Summer School 2021

Report of Contributions

Contribution ID: **1**　　　　　　　　　　　　　　　　Type: **not specified**

Lecture ”Four Generations of Neural Network Potentials”

Monday 13 September 2021 14:15 (1 hour)

A lot of progress has been made in recent years in the development of machine learning (ML) potentials for atomistic simulations [1]. Neural network potentials (NNPs), which have been introduced more than two decades ago [2], are an important class of ML potentials. While the first generation of NNPs has been restricted to small molecules with only a few degrees of freedom, the second generation extended the applicability of ML potentials to high-dimensional systems containing thousands of atoms by constructing the total energy as a sum of environment-dependent atomic energies [3]. Long-range electrostatic interactions can be included in third-generation NNPs employing environment-dependent charges [4], but only recently limitations of this locality approximation could be overcome by the introduction of fourth-generation ML potentials [5], which are able to describe non-local charge transfer using a global charge equilibration step. In this talk an overview about the evolution of NNPs will be given along with typical applications in large-scale atomistic simulations.

[1] J. Behler, J. Chem. Phys. 145 (2016) 170901.
[2] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, J. Chem. Phys. 103 (1995) 4129.
[3] J. Behler and M. Parrinello, Phys. Rev. Lett. 98 (2007) 146401.
[4] N. Artrith, T. Morawietz, J. Behler, Phys. Rev. B 83 (2011) 153101.
[5] T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Comm. 12 (2021) 398.

Author: BEHLER, Jörg (University of Göttingen)

Presenter: BEHLER, Jörg (University of Göttingen)

Session Classification: Machine Learning - Theory

Contribution ID: **2**   Type: **not specified**

Tutorial "Active Learning with Bayesian Optimization"

Monday 13 September 2021 17:00 (1 hour)

Gaussian process regression (GPR) is a kernel-based regression tool with intrinsic uncertainty estimation, which makes it well-suited to natural science datasets. In Bayesian optimization, GPR is coupled with acquisition functions for an active learning approach, where models are iteratively refined by addition of new data points with high information content. This tutorial will use the BOSS code to demonstrate the basic principles of Bayesian optimization and how it can be applied in N-dimensional atomistic structures search.

Author: TODOROVIĆ, Milica (Turku University)

Presenter: TODOROVIĆ, Milica (Turku University)

Session Classification: Machine Learning - Theory

Contribution ID: **3**

Type: **not specified**

Tutorial "Neural Networks"

Tuesday 14 September 2021 14:00 (1 hour)

In this tutorial, we discuss a neural network application in materials discovery. More specifically, we will showcase how to accelerate functional high entropy alloy discovery using neural network based generative model and ensemble model for the regression task. Therefore, this tutorial consists of two parts. Firstly, we discuss in detail how to construct an alloy generation scheme based on generative neural network model, attribute classifier, stochastic sampling, and density estimation model. Secondly, we demonstrate a systematic approach to combine Bayesian optimization, neural network and gradient boosting decision tree to achieve a two-stage Ensemble Regression Model (TERM). The first stage concerns composition-based regression models, aiming at fast and large-scale composition inference. The top results from the first-stage model enter a more refined model, where density functional calculations and thermodynamic calculations are included as part of the model input. Finally, the TERM outputs were evaluated based on a rank-based policy.

Author: WEI, Ye (Max-Planck-Institut für Eisenforschung GmbH)

Presenter: WEI, Ye (Max-Planck-Institut für Eisenforschung GmbH)

Session Classification: Machine Learning - Theory & Application

Contribution ID: **4**　　　　　　　　　　　　　　　　　Type: **not specified**

Lecture "Machine Learning for Electron Microscopy: from Imaging to Atomic Fabrication"

Tuesday 14 September 2021 16:45 (1 hour)

I will discuss recent progress in automated experiment in electron microscopy, ranging from feature to physics discovery via active learning. The applications of classical deep learning methods in streaming image analysis are strongly affected by the out of distribution drift effects, and the approaches to minimize though are discussed. We further present invariant variational autoencoders as a method to disentangle affine distortions and rotational degrees of freedom from other latent variables in imaging and spectral data and decode physical mechanisms. Extension of encoder approach towards establishing structure-property relationships will be illustrated on the example of plasmonic structures. Finally, I illustrate transition from post-experiment data analysis to active learning process. Here, the strategies based on simple Gaussian Processes often tend to produce sub-optimal results due to the lack of prior knowledge and very simplified (via learned kernel function) representation of the system. Comparatively, deep kernel learning (DKL) methods allow to realize both the exploration of complex systems towards the discovery of structure-property relationship, and enable automated experiment workflows targeting physics (rather than simple spatial feature) discovery. The latter is illustrated via experimental discovery of the edge plasmons in STEM/EELS in MnPS3, a lesser-known 2D material.

Author:　KALININ, Sergei (Oak Ridge National Laboratory)

Presenter:　KALININ, Sergei (Oak Ridge National Laboratory)

Session Classification:　Machine Learning - Theory & Application

Contribution ID: **6**   Type: **not specified**

Tutorial "Tips, tricks, and tools for reproducible materials science"

Wednesday 15 September 2021 14:00 (1 hour)

The ability to replicate results is a key characteristic of quality science, and is growing ever more important in light of the replication crisis [1, 2].
A study can rarely be repeated using only the minimalistic descriptions provided in the "Materials and methods" section in a paper.
It is therefore important to properly document the entire knowledge generation pipeline in such a way that it could be repeated by anyone with minimal effort.
In this tutorial we will look at the tools and techniques by which we can improve the reproducibility of an experimental study through a single comprehensive example from electron microscopy.

The session will cover:

- eLabFTW, an electronic lab notebook application with a python API, which can be used to document samples and experiments.

- Various strategies of working with experimental metadata, including the HDF5 file format and JSON.

- Strategies and best practices for reproducible data analysis pipelines in jupyter notebooks, using git, conda-forge, mybinder and docker.

<div id="ioannidis">[1] Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.</div>

<dif id="schooler">[2] Schooler, J. Metascience could rescue the ,'replication crisis'. Nature 515, 9 (2014).</dif>

Author: CAUTAERTS, Niels (Max-Planck-Institut für Eisenforschung GmbH)

Presenter: CAUTAERTS, Niels (Max-Planck-Institut für Eisenforschung GmbH)

Session Classification: Learning from Complex Data

Contribution ID: **7**   Type: **not specified**

Tutorial "Neural Networks with Euclidean Symmetry for Physical Sciences"

Wednesday 15 September 2021 16:45 (1 hour)

Atomic systems (molecules, crystals, proteins, etc.) are naturally represented by a set of coordinates in 3D space labeled by atom type. This is a challenging representation to use for machine learning because the coordinates are sensitive to 3D rotations and translations and there is no canonical orientation or position for these systems. One motivation for incorporating symmetry into machine learning models on 3D data is to eliminate the need for data augmentation —the 500-fold increase in brute-force training necessary for a model to learn 3D patterns in arbitrary orientations.

Most symmetry-aware machine learning models in the physical sciences avoid augmentation through invariance, throwing away coordinate systems altogether. But this comes at a price; many of the rich consequences of Euclidean symmetry are lost: geometric tensors, point and space groups, degeneracy, atomic orbitals, etc.

We present a general neural network architecture that faithfully treats the equivariance of physical systems, naturally handles 3D geometry, and operates on the scalar, vector, and tensor fields that characterize them. We describe how the networks achieves equivariance, demonstrate the capabilities of our network using simple tasks, and provide coding examples to build these models using e3nn: a modular framework for Euclidean Neural Networks (https://e3nn.org).

Authors: GEIGER, Mario (EPFL); SMIDT, Tess (Berkeley Lab)

Presenter: SMIDT, Tess (Berkeley Lab)

Session Classification: Learning from Complex Data

Contribution ID: **8**   Type: **not specified**

Lecture "Variational methods in material sciences"

Thursday 16 September 2021 14:00 (1 hour)

Variational methods are powerful tools in image processing.
Basically we are searching for a suitable mathematical model (function) consisting of a data term and a prior
which minimizer provides a solution of the task at hand and can be computed in an efficient, reliable way. Typically this leads to non-smooth, high-dimensional
optimization problems.
This talk deals with recent results obtained by applying
variational methods for different tasks in material sciences as
- crack detection using optical flow models in image sequences,
- determination of deformation fields in electron backscatter diffraction image sequences,
- superresolution of material images by learned patch-based priors,
and
- denoising of FIB images with directional total variation priors.

Author: STEIDL, Gabriele (TU Berlin)

Presenter: STEIDL, Gabriele (TU Berlin)

Session Classification: Multidimensional Data Analysis

Contribution ID: **9**   Type: **not specified**

Tutorial "Introduction to EELS analysis with HyperSpy"

Thursday 16 September 2021 16:45 (1 hour)

HyperSpy (https://hyperspy.org/) is an open-source Python package for the analysis of multi-dimensional datasets. In its fourteen years of existence, its community of developers has taken it from being a simple collection of scripts for electron energy-loss data analysis to become the core of a multidisciplinary software ecosystem. In this talk, I will its evolution, ecosystem, main features andstructure—all illustrated with a bit of theory and applications. I will also briefly describe how to contribute to HyperSpy and how to create your own HyperSpy extension.

Author: DE LA PEÑA, Francisco (Université de Lille)

Presenter: DE LA PEÑA, Francisco (Université de Lille)

Session Classification: Multidimensional Data Analysis

Contribution ID: **11**   Type: **not specified**

Lecture "NFDI4Chem"

Friday 17 September 2021 14:00 (1 hour)

tba

Author: JUNG, Nicole (Karlsruhe Institute of Technology)

Presenter: JUNG, Nicole (Karlsruhe Institute of Technology)

Session Classification: Open Data

Contribution ID: **12**   Type: **not specified**

Tutorial "Sharing, publishing, and managing computational materials science data with NOMAD"

Friday 17 September 2021 16:30 (1 hour)

The main focus of this tutorial is the FAIR sharing of materials science data and how to do it with NOMAD. We will be covering the publication of new data and the exploration and download from NOMAD's existing data; both through our browser-based interface and APIs.

Author: SCHEIDGEN, Markus (Humboldt Universität zu Berlin / Fritz Haber Institut der Max Planck Gesellschaft)

Presenter: SCHEIDGEN, Markus (Humboldt Universität zu Berlin / Fritz Haber Institut der Max Planck Gesellschaft)

Session Classification: Open Data

Contribution ID: **15**　　　　　　　　　　　　　　　　　　　　Type: **not specified**

Lecture "Science in the Age of Machine Learning"

Monday 13 September 2021 15:15 (1 hour)

Traditionally the "best" observations are those with the largest signal from the most tightly controlled systems. In a wide range of phenomena –the dance of proteins in function, femtosecond breaking of molecular bonds, the gestation of fetuses –tight control is neither feasible, nor desirable. Modern machine-learning techniques extract far more information from sparse random sightings than usually obtained from set-piece experiments. I will describe on-going efforts to extract structural and dynamical information from noisy, random snapshots recorded with very poor, or non-existent timing information. Examples will include functional motions of molecular machines, and ultrafast dynamics of photo-induced reactions.

Author: OURMAZD, Abbas (University of Wisconsin Milwaukee)

Presenter: OURMAZD, Abbas (University of Wisconsin Milwaukee)

Session Classification: Machine Learning - Theory

Contribution ID: **16**   Type: **not specified**

Tutorial "Compressed sensing meets symbolic regression: learning interpretable models"

Monday 13 September 2021 18:00 (1 hour)

In this tutorial, we introduce the AI technique of symbolic regression, combined with compressed sensing for the identification of compact, interpretable models.

Specifically, we introduce the Sure-Independence Screening and Sparsifying Operator (SISSO), together with its recent variants.

The methodology starts from a set of candidate features, provided by the user, and it builds a tree of possible mathematical expression, involving linear and nonlinear operators, up to a given complexity. A compressed sensing solver finds, among billions or trillions of candidate expressions those that better explain the training data.

We will show demonstrative applications to materials science, including prediction of perovskite-materials stability and topological-insulators identification.

Author: GHIRINGHELLI, Luca M. (Fritz Haber Institute of the Max Planck Society)

Co-author: SBAILÒ, Luigi (NOMAD Lab at FHI)

Presenter: GHIRINGHELLI, Luca M. (Fritz Haber Institute of the Max Planck Society)

Session Classification: Machine Learning - Theory

Contribution ID: **17**

Type: **not specified**

Lecture "Machine-learning aided atom probe tomography: status and (possible) directions"

Tuesday 14 September 2021 15:00 (1 hour)

Atom probe tomography (APT) is a materials analysis technique that provides sub-nanometer resolution compositional mapping. The data is in the form of a point cloud containing often millions of atoms, and to each of these points is assocaited an elemental nature. By interrogating the point cloud, the local composition of a material or a phase of a specific microstructural feature can be reported. APT is often referred to as "data-intensive" technique, and has long made use of many clustering-type techniques (DBSCAN, NN etc.) to facilitate data extraction, which are all now often classified as belonging to machine-learning.

In this presentation, I will review some of the recent developments from MPIE in the application of machine-learning techniques to atom probe analysis workflows –i.e. beyond just extraction of data from the point cloud –targeting faster and more efficient, reliable and reproducible data analysis.

Authors: SAXENA, Alaukik (Max-Planck-Institut für Eisenforschung GmbH); GAULT, Baptiste (Max-Planck-Institut für Eisenforschung GmbH); FREYSOLDT, Christoph (Max-Planck-Institut für Eisenforschung GmbH); WEI, Ye (Max-Planck-Institut für Eisenforschung GmbH); LI, Yue (Max-Planck--Institut für Eisenforschung GmbH)

Presenter: GAULT, Baptiste (Max-Planck-Institut für Eisenforschung GmbH)

Session Classification: Machine Learning - Theory & Application

Contribution ID: **18**   Type: **not specified**

Lecture "Neural Networks with Euclidean Symmetry for Physical Sciences"

Wednesday 15 September 2021 15:00 (1 hour)

Atomic systems (molecules, crystals, proteins, etc.) are naturally represented by a set of coordinates in 3D space labeled by atom type. This is a challenging representation to use for machine learning because the coordinates are sensitive to 3D rotations and translations and there is no canonical orientation or position for these systems. One motivation for incorporating symmetry into machine learning models on 3D data is to eliminate the need for data augmentation —the 500-fold increase in brute-force training necessary for a model to learn 3D patterns in arbitrary orientations.

Most symmetry-aware machine learning models in the physical sciences avoid augmentation through invariance, throwing away coordinate systems altogether. But this comes at a price; many of the rich consequences of Euclidean symmetry are lost: geometric tensors, point and space groups, degeneracy, atomic orbitals, etc.

We present a general neural network architecture that faithfully treats the equivariance of physical systems, naturally handles 3D geometry, and operates on the scalar, vector, and tensor fields that characterize them. We describe how the networks achieves equivariance, demonstrate the capabilities of our network using simple tasks, and provide coding examples to build these models using e3nn: a modular framework for Euclidean Neural Networks (https://e3nn.org).

Authors: SMIDT, Tess (Berkeley Lab); GEIGER, Mario (EPFL)

Presenter: SMIDT, Tess (Berkeley Lab)

Session Classification: Learning from Complex Data

Contribution ID: **19**   Type: **not specified**

Lecture "Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery"

Wednesday 15 September 2021 17:45 (1 hour)

The increasing availability of data from the first three paradigms of science (experiments, theory, and simulations), along with advances in artificial intelligence and machine learning (AI/ML) techniques has offered unprecedented opportunities for data-driven science and discovery, which is the fourth paradigm of science. Within the arena of AI/ML, deep learning (DL) has emerged as a game-changing technique in recent years with its ability to effectively work on raw big data, bypassing the (otherwise crucial) manual feature engineering step traditionally required for building accurate ML models, thus enabling numerous real-world applications, such as autonomous driving. In this talk, I will present our ongoing research in AI and high performance data mining, along with illustrative real-world scientific applications. In particular, we will discuss approaches to gainfully apply DL on big data (by accelerating DL and enabling deeper learning) as well as on small data (deep transfer learning) in the context of materials science. I will also demonstrate some of the software tools developed in our group.

Author: AGRAWAL, Ankit (Northwestern University)

Presenter: AGRAWAL, Ankit (Northwestern University)

Session Classification: Learning from Complex Data

Contribution ID: **20**   Type: **not specified**

Lecture "HyperSpy: theory and applications"

Thursday 16 September 2021 15:00 (1 hour)

HyperSpy (https://hyperspy.org/) is an open-source Python package for the analysis of multidimensional datasets. In its fourteen years of existence, its community of developers has taken it from being a simple collection of scripts for electron energy-loss data analysis to become the core of a multidisciplinary software ecosystem. In this talk, I will its evolution, ecosystem, main features andstructure—all illustrated with a bit of theory and applications. I will also briefly describe how to contribute to HyperSpy and how to create your own HyperSpy extension.

Author: DE LA PEÑA, Francisco (Université de Lille)

Presenter: DE LA PEÑA, Francisco (Université de Lille)

Session Classification: Multidimensional Data Analysis

Contribution ID: **21**   Type: **not specified**

Lecture "Structuring, analyzing, and harvesting big data in materials science electron microscopy"

Friday 17 September 2021 15:00 (1 hour)

Every day, experimental materials science data is being collected in thousands of laboratories around the world. However, the diversity of instruments, vendor software packages and (proprietary) data formats, lab cultures, and the focus mostly on new discoveries causes most of this data to end up in a black hole in terms of accessibility to the scientific community (including in many cases the lab in which the data was acquired). Using the example of transmission electron microscopy we will present our attempts to encourage scientists to annotate and contribute more of their data to the scientific community and benefiting themselves in the process. We are striving to achieve this by being able to offer relevant online data processing and analyzing capabilities and tools to annotate data in more and more automated ways. As more and more data is being accumulated, big data techniques can be applied to benefit from the added value that sharing of experimental data sets produces.

Authors: KOCH, Christoph (Humboldt-Universität zu Berlin); KÜHBACH, Markus (Max-Planck-Institut für Eisenforschung GmbH); SHABIH, Sherjeel (Humboldt-Universität zu Berlin)

Presenter: KOCH, Christoph (Humboldt-Universität zu Berlin)

Session Classification: Open Data

Contribution ID: **22**

Type: **Poster**

Robust recognition and exploratory analysis of crystal structures via Bayesian deep learning

Tuesday 14 September 2021 19:00 (2 hours)

Due to their ability to recognize complex patterns, neural networks can drive a paradigm shift in the analysis of materials science data. Here, we introduce ARISE, a crystal-structure identification method based on Bayesian deep learning. As a major step forward, ARISE is robust to structural noise and can treat more than 100 crystal structures, a number that can be extended on demand. While being trained on ideal structures only, ARISE correctly characterizes strongly perturbed single- and polycrystalline systems, from both synthetic and experimental resources. The probabilistic nature of the Bayesian-deep-learning model allows to obtain principled uncertainty estimates, which are found to be correlated with crystalline order of metallic nanoparticles in electron-tomography experiments. Applying unsupervised learning to the internal neural-network representations reveals grain boundaries and (unapparent) structural regions sharing easily interpretable geometrical properties. This work enables the hitherto hindered analysis of noisy atomic structural data from computations or experiments.

Author: LEITHERER, Andreas (Fritz Haber Institute of the Max Planck Society)

Co-authors: Dr ZILETTI, Angelo (Fritz Haber Institute of the Max Planck Society); Dr GHIR-INGHELLI, Luca M. (Fritz Haber Institute of the Max Planck Society)

Presenter: LEITHERER, Andreas (Fritz Haber Institute of the Max Planck Society)

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **31**   Type: **Poster**

Teaching solid mechanics to artificial intelligence

Tuesday 14 September 2021 19:00 (2 hours)

I will present our latest progress in using machine learning for solving non-linear solid mechanics. The presentation will be based on our published work here: https://www.nature.com/articles/s41524-021-00571-z

Authors: MIANROODI, Jaber (Max-Planck-Institut für Eisenforschung GmbH); H. SIBONI, Nima (Max-Planck-Institut für Eisenforschung GmbH); RAABE, Dierk (Max-Planck-Institut für Eisenforschung GmbH)

Presenter: MIANROODI, Jaber (Max-Planck-Institut für Eisenforschung GmbH)

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **32**　　　　　　　　　　　　　　　　　　　Type: **Poster**

Multidimensional Photoemission Spectroscopy: proposal & demonstration of data (infra)structure

Tuesday 14 September 2021 19:00 (2 hours)

The complexity of photoemission data is rapidly increasing,
as new technological breakthroughs have enabled multidimensional parallel acquisition of multiple observables. Most of the community is currently using heterogeneous data formats and workflows.

We propose a new data format based on NeXus, a hierarchically organized hdf5 structure. The aim is to immediately enable preprocessed data and metadata shareability according to FAIR principles, employing the existing storage and archiving infrastructure such as Zenodo, OpenAIRE and Nomad/FAIRmat. Ultimately, the multidimensional photoemission spectroscopy (MPES) format is designed to allow high-performance automated access, providing experimental databases for high-throughput material search.

The MPES format is based on the creation of a standardized set of classes that univocally identify experimental observables and metadata. Such dictionary is complemented by "application definitions", i.e. ontologies that constrain the existence of specific elements in a file.

Our approach involves reaching out to the community using a website with a wiki structure. By this medium, we wish to favour acceptance, to avoid conflict with different low-level preprocessing workflows, and to create continuously updated documentation.

As a demonstrator of the potential of our approach, we present the workflow we developed for our data pipeline, originating from time-resolved angularly-resolved photoemission spectroscopy.

Authors:　PINCELLI, Tommaso (Fritz Haber Institute of the Max Planck Society);　AGUSTSSON, Steinn Ymir (JGU Mainz);　XIAN, Rui Patrick (Fritz Haber Institute of the Max Planck Society);　Dr KUTNYAKHOV, Dmytro (DESY Photon Science, 22607, Hamburg, Germany);　Prof. ACREMANN, Yves (Laboratory for Solid State Physics, ETH Zurich, 8093, Zurich, Switzerland);　HARTELT, Michael (Technical University Kaiserslautern);　Dr DENDZIK, Maciej (KTH Royal Institute of Technology, Stockholm);　Dr DONG, Shuo (Fritz Haber Institute of the Max Planck Society);　Dr BEAULIEU, Samuel (CELIA, CNRS-Universitè de Bordeaux);　Mr MAKLAR, Julian (Fritz Haber Institute of the Max Planck Society);　Dr SCHEIDGEN, Markus (Humboldt-Universität zu Berlin);　RETTIG, Laurenz (Fritz Haber Institute of the Max Planck Society);　ERNSTORFER, Ralph (Fritz Haber Institute of the Max Planck Society)

Presenters:　PINCELLI, Tommaso (Fritz Haber Institute of the Max Planck Society);　AGUSTSSON, Steinn Ymir (JGU Mainz)

Session Classification:　Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **33** Type: **Poster**

A materials informatics framework to discover patterns in atom probe tomography data.

Tuesday 14 September 2021 19:00 (2 hours)

To quantify chemical segregation at multiple length scales in APT in a semi-automatic way, we propose a multi-stage strategy. First, we collect composition statistics from APT datasets for 2x2x2 nm voxels. These voxel compositions are then clustered in compositional space using Gaussian mixture models to automatically identify key phases. Next, based on this compositional classification we employ DBSCAN in physical space at voxel resolution to detect individual precipitates at a small fraction of the effort needed for single-atom-based algorithms. This framework was used to identify and disentangle plate-like Zr-rich and topologically complex Cu-rich precipitates in two APT datasets from Fe-doped Sm-Co magnets, each containing approximately 500 million ions. Upon segmentation of each precipitate using DBSCAN a new approach based on principle component analysis (PCA) is applied to 2D slices of the complex Cu-rich precipitates to further decompose them into approximate planar regions along with their junctions. For each precipitate, the actual distribution of atomic fractions is compared to the expected distribution of a random alloy. This step helps to quantitatively assess clustering within a given precipitate for each atomic species. Finally, for each quasi-planar precipitate, a triangular grid is superimposed to investigate in-plane compositional fluctuations, thickness, and 1D composition profiles.

Author: SAXENA, Alaukik (Max-Planck-Institut für Eisenforschung GmbH (Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE)))

Co-authors: Mr POLIN, Nikita (Max-Planck-Institut für Eisenforschung GmbH); Prof. RAABE, Dierk (Max-Planck-Institut für Eisenforschung GmbH); Prof. BERKELS, Benjamin (RWTH Aachen University); GAULT, Baptiste (Max-Planck-Institut für Eisenforschung GmbH); FREYSOLDT, Christoph (Max-Planck-Institut für Eisenforschung GmbH)

Presenter: SAXENA, Alaukik (Max-Planck-Institut für Eisenforschung GmbH (Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE)))

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **34**   Type: **Poster**

Deep generative models for the design of Dual phase steel microstructures

Tuesday 14 September 2021 19:00 (2 hours)

Dual Phase (DP) steels are an important family of steel grades used widely in the automotive industry because of their beneficial properties such as high ultimate tensile strength (GPa range), low initial yield stress, and high early-stage strain hardening. The DP steel microstructure consists of soft ferritic grains, which are mainly responsible for ductility, and hard martensitic zones, which give these steels their strength. However, due to the huge parameter space, establishing the microstructure-property relationship in DP steels is a combinatorial challenge. To tackle this challenge we use data from crystal plasticity simulations and train machine learning models to automatically extract the patterns from DP steel microstructures without any need of hand-designed features. In particular, I will present my work on using Variational Auto Encoders (VAEs, a generative machine learning model) to learn the low dimensional latent representations and to generate synthetic EBSD data for DP steels. Also, I will present, that the learned latent space can serve as a design space for microstructure design.

Author: KUSAMPUDI, Navyanth (Max-Planck-Institut für Eisenforschung GmbH)

Presenter: KUSAMPUDI, Navyanth (Max-Planck-Institut für Eisenforschung GmbH)

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **37**   Type: **Poster**

Learning Dynamics of STEM by Enforcing Physical Consistency with Phase-Field Models

Tuesday 14 September 2021 19:00 (2 hours)

In this poster, we present our research goals of a recently BiGmax funded project towards learning dynamics of scanning transmission electron microscopy (STEM) by incorporating physical consistency with phase-field models. The primary idea of this project is to develop machine learning (ML)-based modeling of an interpretable coarse-grained dynamic model utilizing in situ STEM video sequences fulfilling a suitable dynamical phase-field equation. The modeling approach aims to discover governing equations by utilizing the video sequence data and prior physics knowledge that is directly compatible with analytic theories or subsequent ML-based analysis.

Authors: LIEBSCHER, Christian (Max-Planck-Institut für Eisenforschung GmbH); FREYSOLDT, Christoph (Max-Planck-Institut für Eisenforschung GmbH); MIANROODI, Jaber (Max-Planck-Institut für Eisenforschung GmbH); SREEKALA, Lekshmi (Max-Planck-Institut für Eisenforschung GmbH); WANG, Ning (Max-Planck-Institut für Eisenforschung GmbH); GOYAL, Pawan (Max Planck Institute for Dynamics of Complex Technical Systems); BENNER, Peter (Max Planck Institute for Dynamics of Complex Technical Systems); BUKKA, Sandeep Reddy (Max Planck Institute for Dynamics of Complex Technical Systems)

Presenters: SREEKALA, Lekshmi (Max-Planck-Institut für Eisenforschung GmbH); GOYAL, Pawan (Max Planck Institute for Dynamics of Complex Technical Systems)

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data

Contribution ID: **38**   Type: **Poster**

Dynamic structure investigation of biomolecules with pattern recognition algorithms and X-ray experiments

Tuesday 14 September 2021 19:00 (2 hours)

Author: KOTOBI, Amir (Deutsches Elektronen-Synchrotron DESY)

Co-authors: SCHWOB, L. (Deutsches Elektronen-Synchrotron DESY); MEISSNER, R. (nstitute of Polymers and Composites, Hamburg University of Technology); BARI, S. (Deutsches Elektronen-Synchrotron DESY)

Presenter: KOTOBI, Amir (Deutsches Elektronen-Synchrotron DESY)

Session Classification: Poster Session: Machine Learning -Theory & Applications / Open Data