



Contribution ID: 10

Type: not specified

## Machine Learning-based Spectra Classification

*Tuesday, 26 October 2021 11:15 (15 minutes)*

Spectroscopy experiment techniques are widely used and produce huge amounts of data especially in facilities with very high repetition rates. At the European XFEL, X-ray pulses can be generated with only 220ns separation in time and a maximum of 27000 pulses per second. In experiments (e.g. SCS, FXE, MID, and HED) at European XFEL, spectral changes can indicate the change of the system under investigation and so the progress of the experiment. Immediate feedback on the actual status (e.g., time-resolved status of the sample) would be essential to quickly judge how to proceed with the experiment. The major spectral changes that we aim to capture are either the change of intensity distribution (e.g., drop or appearance) of peaks at certain locations, or the shift of those on the spectrum.

Machine Learning (ML) opens up new avenues for data-driven analysis in spectroscopy by offering the possibility to quickly recognize such specific changes on-the-fly during data collection, and it usually requires lots of data that are clearly annotated. Hence, it is important that research outputs should align with the FAIR principles. For XFEL experiments, it is suggested to introduce NeXus data format standards in future experiments.

In this work, we present an example to show how Neural Network-based ML can be used for accurately classifying the system state if data is properly provided. We demonstrate a solution to automatically find the regions (or bins) with high separability where the spectra classes differ significantly. By teaching individual neural networks for each bin and combining them with a weighting technique, a robust classification of any new spectral curve can be quickly obtained.

**Presenter:** SUN, Yue