

Machine Learning Prediction of Valence and Coordination from EELS Spectra of Iron Containing Compounds

Samuel P. Gleason^{1,2}, Deyu Lu³ and Jim Ciston¹

¹ National Center for Electron Microscopy Facility, Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, United States.

² Department of Chemistry, University of California, Berkeley, CA, United States.

³ Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY, United States.

Electron energy loss spectroscopy (EELS) allows identification of atoms present in a sample as well as providing detailed information about bonding, distributions and locations of atoms, and their coordination numbers and oxidation states [1]. However, the quantitative analysis of EELS data typically requires manual labeling of a spectrum using known standards, a process which can be challenging when signal to noise ratios are poor or when reliable experimental or theoretical standards do not exist. Furthermore, accurate simulation of EEL spectra is often required to infer details of valence and local bonding from experimental spectra. This is especially challenging for transition metal L-edges, often requiring beyond-DFT methods tailored to treat correlated electronic systems. Machine learning models, trained on a broad set of simulated data and experimental standards, capable of predicting structural information from EELS spectra are a promising route to mitigate this problem due to the increased availability of computational datasets where spectra and structural information have been generated [2]. Previous work on this topic has generated models that leveraged these computational and experimental datasets and have been successful in extracting structural information from a fundamentally similar analysis method, X-ray absorption spectroscopy (XAS), and have also significantly accelerated analysis of these spectra [3, 4].

In this work, we train a machine learning model capable of predicting the Fe valence and Fe-O coordination environment of a compound based on its EELS spectrum [2]. To build the model, we benchmark simulated Fe L-edge EELS spectra against direct experimental measurements of classes of iron containing compounds, using either available FEFF database from the Materials Project [5] or by generating our own dataset using the Quancy code base. We augment this simulated dataset with common experimental sources of uncertainty, such as detector noise, background subtraction artifacts, and energy offsets. This model can be used to significantly accelerate the analysis of EELS spectra of mixtures of unknown iron containing compounds, including ones for which there are few theoretical or experimental standards available, such as glassy oxides. Additionally, this methodology can, in principle, be extended to the analysis of other compounds by expanding the training data [7].

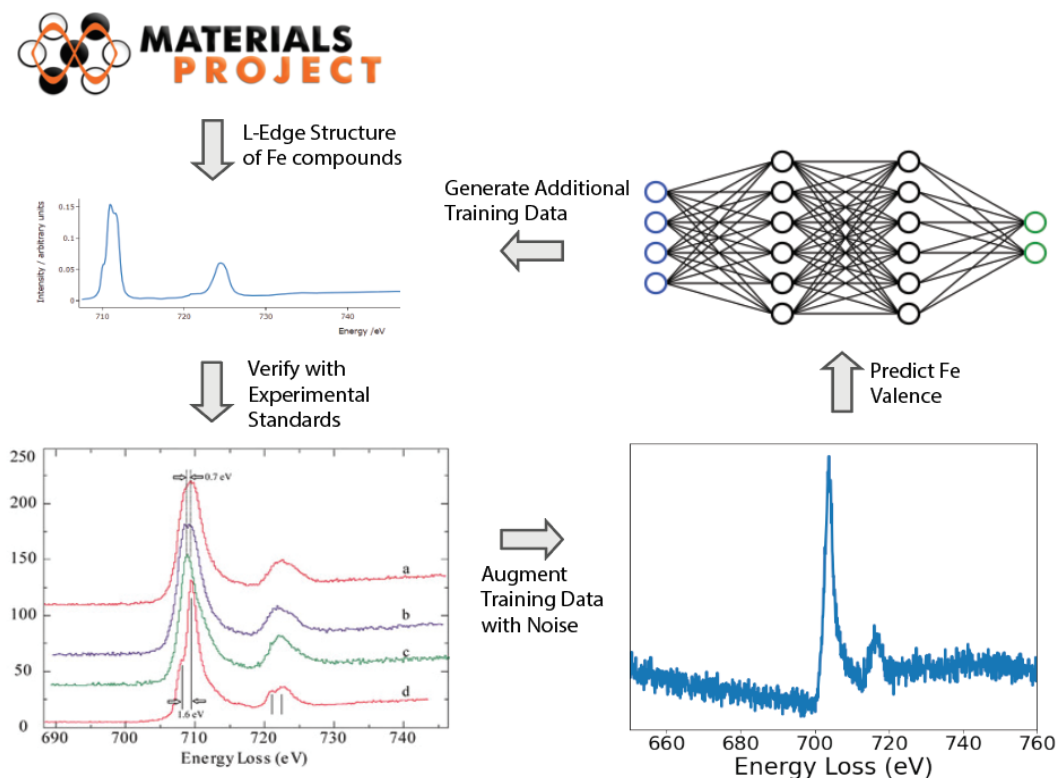


Figure 1. Schematic illustrating the data acquisition and model training pipeline. We first extract compounds from the Materials Project with simulated L-edge spectra (top left). We then verify the accuracy of these spectra using reference standards from experiments (bottom left) [6]. After verifying the accuracy of the simulated data, we augment our training data to better match the spectra that will be seen in experiments by adding features such as detector noise and artifacts generated by background subtraction (bottom right). After assembling our training set, we build a model to extract features such as Fe valence and coordination environment from the L-edge EEL spectrum (top right). To address issues such as sparsity of data for particular Fe compounds, such as Fe hydroxides, we increase our training data by using the Quancy code base to calculate additional spectra (middle arrow). We then continue this cycle until the predictions reach an acceptable level of accuracy.

References:

- [1] E Yasuda et al., in “Carbon Alloys” (Elsevier Science, Oxford) p. 239.
- [2] A Jain et al., *APL Materials* **1** (2013), p. 011002.
- [3] SE Reyes-Lillo et al., *Phys. Rev. Materials* **3** (2019), p. 033604.
- [4] J Timoshenko et al., *J. Phys. Chem. Lett.* **8** (2017) p. 5091.
- [5] Y Chen et al., *Scientific Data* **8** (2021), p. 153.
- [6] Y Jin, H Xu and AK Datye, *Microscopy and Microanalysis* **12** (2006) p. 124.
- [7] This work was primarily funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences in the program “4D Camera Distillery: From Massive Electron Microscopy Scattering Data to Useful Information with AI/ML.” Electron Microscopy use at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the Center for Functional Nanomaterials (CFN), which is a U.S. Department of Energy Office of Science User Facility, at Brookhaven National Laboratory under Contract No. DE-SC0012704.