XRD+ prediction model using SVM and ANN

Hao Zhang[1], Prathyush P Menon[1], Gino Hrkac[1]

[1] *Department of Engineering, University of Exeter, Exeter, UK.*

X-ray diffraction (XRD) is a commonly used material analysis technique that can be used to determine information such as the structure, composition, and crystal structure of substances. However, interpreting and analysing XRD data requires experienced experts and usually takes time and effort. Subjective factors influence traditional data analysis methods, and analysts may draw different conclusions. In recent years, machine learning technology has made XRD data analysis faster, more efficient, and more accurate, so more researchers have begun applying machine learning technology to XRD data analysis.

XRD data is usually extensive and complex, and traditional analysis methods require tedious processing and analysis by experienced experts. Machine learning technology can use algorithms and models to process these data, automatically extract features and patterns, and quickly and accurately classify and predict.

In our model we start from our self-coded spacegroup database that uses crystallographic properties / symmetry and point groups to generate a data set of simplified XRD images. The model is designed in such a way that it can generate mixed systems based on surface cuts and interface matching and producing corresponding XRD+ images for such mixed systems. This XRD/XRD+ database together with materials project entries are used for data-driven symmetry identification and property prediction, and compared to predictions from inorganic crystal structure database (ICSD) and materials project [1].
We start with an experimental full-profile XRD pattern as an input. Then we use support vector machine (SVM) and artificial neural network (ANN) algorithms for symmetry classification. The symmetry recognition is performed through classification models based on our theoretical XRD data base and phase diagrams derived from the composition vector. The system is trained with known experimental XRD patterns using SVM and ANN to test its accuracy (training:testing;80:20). Once optimised for a composition vector, the model is then used to analyse the full XRD pattern. In addition we are using a fully convolutional neural network (FCN) concatenated with multilayer perceptron (MLP) using a composition vector for property regression, see figure 1 schematic.

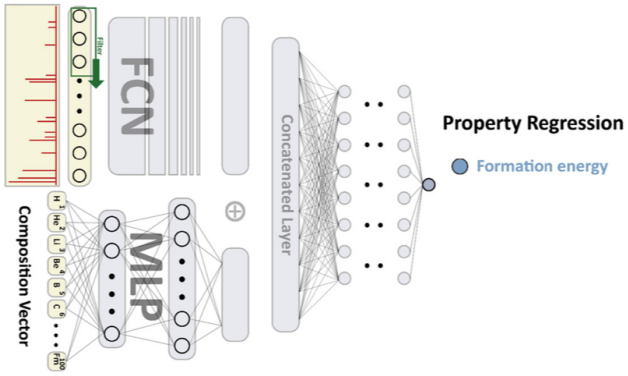

Figure 1: Schematic of XRD/XRD+ model predictor [1].

The results from the FCN-MLP can be compared to DFT and MD simulations, e.g. formation energy. The results are compared to those obtained from a well-established crystal graph convolutional neural network (CGCNN).

References

[1] B.D. Lee, J. Lee, W.B. Park, et al., Powder X-Ray Diffraction Pattern Is All You Need for Machine-Learning-Based Symmetry Identification and Property Prediction. Adv. Intell. Syst., 4: 2200042 (2022).