

## Application of machine learning and kernel density estimation for efficient data treatment on single crystal diffraction data

R. Kiyanagi<sup>1</sup>, T. Ohhara<sup>1</sup>, A. Nakao<sup>2</sup>, K. Munakata<sup>2</sup>, Y. Ishikawa<sup>2</sup>, K. Moriyama<sup>2</sup>

<sup>1</sup>J-PARC center, Japan Atomic Energy Agency, Ibaraki, Japan,

<sup>2</sup>Neutron Science and Technology Center, Comprehensive Research Organization for Science and Society, Ibaraki, Japan

ryoji.kiyanagi@j-parc.jp

“SENJU” is a TOF-Laue neutron single crystal diffractometer installed at J-PARC/MLF in Japan. This instrument is designed to study structures of inorganic materials and organic materials with relatively small cell sizes as well as magnetic structures. SENJU has 41 2D-detectors installed cylindrically surrounding the sample position. With the benefit of high intensity pulsed white neutrons generated at J-PARC, SENJU can measure a quite large 3D reciprocal space at once, which makes a quite efficient measurements possible. In order to make a good use of the efficient measurement, as a natural consequence, efficient data treatment is also demanded.

The experiments conducted at SENJU can be categorized into two types, one is for a standard structure analysis where a large number of Bragg reflections are collected, and the other is for a search for superlattice reflections including magnetic reflections where new reflections, which typically are weak, are searched within the observed 3D reciprocal space. For these two types of data, some mathematical methods are applied to efficiently treat the data.

### 1) Application of machine learning to Bragg reflection selection

In a measurement for a standard structure analysis, typically some hundreds, or often more than ten thousand, of Bragg reflections are measured owing to the 2D-detectors and the white neutrons. The issue is that the measured data have to be checked before fed into a structure analysis software, because there can be ill reflections such as ones overlapped with a powder ring coming from, e.g., radiation shields or ones very close to each other. In order to eliminate such reflections, machine learning was adapted.

Several sets of “good” and “bad” data, including simulated ones, were prepared and used as a training data. With each training data, a model was constructed based on the convolutional neural network. Most of the trainings reached models with high accuracy, namely higher than 90%, and the models, indeed, were able to distinguish “good” and “bad” data from actual data with high probability. Further improvement is envisaged with increased number of the training data, especially “bad” data.

### 2) Application of kernel density estimation to reciprocal map

In a measurement for a search for superlattice reflections, the expected reflections are typically very weak. Hence longer exposure is needed and, often, even after the long exposure, the reflections could be blurry. In order to enhance the chance to find new reflections, improvement of the visibility of the data is one option. Therefore, the kernel density estimation (KDE) method was applied to a measured reciprocal map data.

It was found that the visibility of the reflections is greatly enhanced by the application of KDE. As shown in Figure 1, the reflections that are obscure in the raw data map became much clearer in the KDE map. In the KDE map, even the splitting of the peak can be recognized. The visibility was estimated to be increased by five times by KDE, which means that the application of KDE can be equal to five times longer exposure.

In the presentation, the details of the application of the machine learning and KDE will be shown, including the information about the codes used in the calculations.

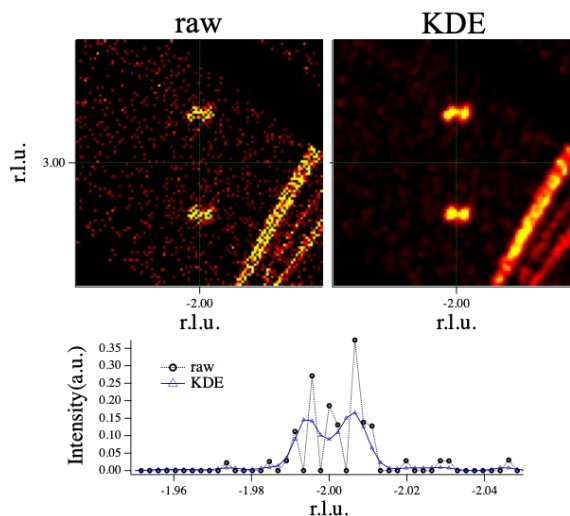


Figure 1. Comparison between raw data and KDE of a reciprocal map

**Keywords:** single crystal diffraction; neutron; machine learning; neural network; kernel density estimation;