

Modern machine learning tools for crystallography

Paolo Bosetti¹, Matteo Ragni¹, Matteo Leoni²

¹Department Of Industrial Engineering, University Of Trento, Trento, Italy, ²Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy
E-mail: paolo.bosetti@unitn.it

The development of diffraction (from X-rays, neutrons, electrons) has been the key to understand the atomic arrangement in materials and complex systems. The structure of simple compounds can nowadays be solved in a rather straightforward way on small computers.

Some crystallography problems, however, are still open. Two of them are of particular relevance: a fast and accurate identification of all constituents in a mixture (for the identification of possible unknown and of minor phases) and the general structural solution of large structures. The available tools, based on a deterministic approach, seem to fail in the solution of the general case.

The fast development GPU-based computing has revived the interest in general approximator based upon machine learning approach [1]. Fast training algorithms, developed for massively parallel consumer GPUs, have brought the complexity of deep multi-layer perceptron to desktop computing and embedded systems. As a consequence, the applications of neural networks to solve the most diverse tasks has exploded in the last few years [2].

In the recent past, the applications of machine learning in crystallography have been limited to the formulation of clustering and principal component analysis approaches to datasets, but did not massively capture the interest of the community. We will show here that an approximator, trained on a dataset that allow to sufficiently generalize a problem, can help in solving some of the toughest challenges.

In particular, the problems of classification of diffraction data, of identification of the phases in a mixture or of structure solution (phasing), can be expressed with a machine learning approach and a neural network can be trained to not only answer to a specific query, but also to express the confidence on such classification. Albeit still in its infancy, the approach has the potential to have a deep impact in the field and largely extend the current state of the art. Theory and applications will be provided.

[1] Goodfellow, I. et al. (2016). Deep learning. MIT Press.

[2] Liu, W. et al. (2016). Neurocomputing. 234, 11-26.



Keywords: [Deep Learning](#), [Diffraction](#), [Machine Learning](#)