Machine learning techniques for data analysis in materials science

Claudio RONCHETTI - *ENEA, Centro Ricerche Casaccia Roma*

Science has been characterized by four paradigms: empirical models, formulation of physical / chemical laws, computational / simulator models, and data-driven techniques. Our project is based on the exploitation of the fourth paradigm to speed up scientific discovery. In recent years, there has been rapid progress in the development of deep learning neural networks for materials science. Such networks have shown that they can predict accurately and quickly the materials' properties: those networks represent the molecule or crystalline material as a graph where the atoms and their atomistic bonds correspond respectively to the nodes and the edges of the graph. Graph Neural Networks (GNNs) are a class of the deep learning methods designed to perform inference on data described by graphs, providing an easy way to perform node-level and edge-level prediction and classification tasks. In the material science, there have been developed GNN architectures-based neural networks such as Crystal Graph Convolutional Neural Networks (CGCNN, and more recent Geometric-Information-Enhanced Crystal Graph Network, GeoCGNN), MatErials Graph Network (MEGNet) and Atomistic Line Graph Neural Network (ALIGNN). We used and compared different networks to predict a property of materials (formation energy), and then we applied the transfer learning method to adapt the best neural network to our domain and to improve the results.