

Title: Fingerprinting local environments with applications to machine learning interatomic potentials

Abstract: A class of supervised machine learning approaches aim at predicting a label (i.e., the value of some quantity) from an input data vector. For example, it is common practice to recognize a person's face from the set of data points (pixels) in a digital image frame. The same techniques are effectively useful in computational condensed matter physics problems for prediction of atomic contributions to a given physical quantity from the arrangement of the neighboring atoms of the individual atoms. Then one needs a "descriptor" that quantifies the environment of an atom such that it can be fed as input to a supervised machine learning tool. We review the basic ideas, techniques and challenges of fingerprinting the local environment as an essential ingredient of training interatomic potentials from e.g., ab initio reference samples. We exemplify the local fingerprinting by applying it to predict the excited and ground-state energy surfaces of large atomic systems.