S. Pozdnyakov[1,2]
E. Mazhnik[1]
I. Kruglov[1, 2]
A. Oganov[1, 2, 3, 4, 5]
A. Yanilkin[1, 2]

MACHINE LEARNING POTENTIAL

[1]Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region 141700, Russian Federation
[2] Dukhov Research Institute of Automatics (VNIIA), Moscow 127055, Russian Federation
[3]Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Moscow 143026, Russian Federation
[4]Department of Geosciences, State University of New York, Stony Brook, NY 11794-2100, USA
[5]International Center for Materials Design, Northwestern Polytechnical University, Xian 710072, China

sergei.pozdnyakov@gmail.com

Various problems in computational chemistry demand predicting energies and forces for a given atomic structures. There are two main approaches to get this predictions. The first one is to use ab initio calculations, like DFT methods, the second one is classic potentials. The first approach is very computationally expensive, while classic potentials usually provide insufficient accuracy. Machine learning algorithms provide both sufficient accuracy and low computational costs.
In this work we developed such machine learning potential.

The general scheme of machine learning potentials is that they are trained on some amount of atomic structures for which energies or/and forces are calculated via ab initio calculations and after using them energies or/and forces for a new unseen atomic structures can be predicted very fast.

First we fitted pair potentials. At this step we assumed that there were only two-particles interactions, which are described by pair potentials. We approximated this function with B-splines, and found the best pair potential with the lowest error on energies and forces. The model with only pair interaction showed relatively high accuracy, thereby we concluded that this part gives the largest contribution into the total potential.

After that we included three-particles interactions, which are described by three-particles potential. In contrast to two-particles potential this is a function of three variables. It was also approximated by three-dimensional B-splines. Following this idea other many-particles interactions can be described, but it would require functions of higher number of variables, so extending this approach in this way is very computationally costly.

In order to improve the accuracy of machine learning potentials and include other many-particles interactions we decided to follow existing approaches in this field. We chose the linear regression of Bispectrum components, which is called SNAP.