Machine learning for material property prediction and applications in high entropy alloys

Jun Ni

State Key Laboratory of Low-Dimensional Quantum Physics, Department of Physics, Tsinghua University, Beijing 100084, China;
Email address: junni@mail.tsinghua.edu.cn

In this talk, I will introduce our recent work on how to predict properties of high entropy alloys using machine learning models accurately with small data set. For high entropy alloys, the large data set is challenging to obtain and the structures of high entropy alloysis also difficult to determine. We build a deep learning framework that combines elemental convolution and Generative Adversarial Networks(GAN). The framework improves the prediction performance of graph neural networks in a small data set while significantly reducing the computational consumption when using accurate structures as input for prediction. To improve the accuracy of machine learning predictions, it is usually to use material structures as descriptors. However, the complexity of computing material structures limits the practical use of these models. To address this challenge and improve prediction accuracy in small data sets, we develop a generative network framework: Elemental Features enhanced and Transferring corrected data augmentation in Generative Adversarial Networks (EFTGAN). Combining the elemental convolution technique with Generative Adversarial Networks (GAN), EFTGAN provides a robust and efficient approach for generating data containing elemental and structural information that can be used not only for data augmentation to improve model accuracy, but also for prediction when the structures are unknown. Applying this framework to the FeNiCoCrMn/Pd high-entropy alloys, we successfully improve the prediction accuracy in a small data set and predict the concentration-dependent formation energies, lattices, and magnetic moments in quinary systems accurately. The model is also applied to the inorganic crystal data obtained from the Materials Project. We choose a data set containing 5830 samples to predict the bulk modulus and shear modulus. A wide range of inorganic crystals, from simple metals to complex minerals, are included in the data set covering 87 elements and 7 lattice systems. The model shows a significant improvement in performance compared to other graph-based neural networks. This shows that the EFTGAN model is capable of processing larger samples and more complex structures. This study provides a new algorithm to improve the performance and usability of deep learning with structures and without structures as inputs, which is effective and accurate for the prediction and development of materials for small data sets.