An interpretable deep learning model to extract material structure-property relationship

Tien-Sinh Vu[1], Minh-Quyet Ha[1], Duong-Nguyen Nguyen[1], Viet-Cuong Nguyen[2] and Hieu-Chi Dam[1]

[1]Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
[2]HPC SYSTEMS Inc., 3-9-15 Kaigan, Minato, Tokyo 108-0022, Japan

Abstract

Nowadays, the recent development of automated experiments and high-performance computers has made it possible to obtain vast amounts of experimental and computational material data, which increasingly requires developing explainable and interpretable material informatics approaches to accelerate the understanding of physical and chemical phenomena [1]. However, although machine learning and deep learning methods using existing data show very optimistic high prediction accuracy for materials properties, it is difficult to estimate how predictive these approaches will be for new materials. Moreover, it is also quite challenging to expect these approaches to lead to a deeper understanding of physicochemical phenomena. In this study, we developed a deep learning model for learning the representations of atomic sites from the data of material structures. We applied the attention mechanism (AM) [2] to describe the interactions between atomic sites. Incorporating AM into deep learning models provides deeper insight into interpreting the relationship between the model inputs and outputs and improves the performance of the models. In addition to enabling the accurate prediction of specific target properties, the training process allows the model with AM to learn the representation of all the atomic sites in a material structure and their impact on each other. We evaluate the proposed method's predictive capability and interpretability by using three datasets, including the QM9 benchmark dataset [3] and our two homemade datasets: Fullerene molecules and Pt-Graphene dataset, which are obtained from first-principles molecular dynamics simulations. The experimental comparison confirmed that the prediction performance of the proposed model is comparable with the state-of-the-art material representations from other deep learning methods. Furthermore, the interpretability of the proposed method was discussed by focusing on specific physical properties and evaluating the attention contribution of the constituent atom local structures to the final representation of the structures compared with the results obtained from first-principles calculations. The results of these evaluation experiments confirm that the structure attention scores obtained by the proposed method provided important information for understanding the structure-property relationship of materials.

References

[1] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. Von Lilienfeld, Physical review letters 108, 058301 (2012)

[2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, in Advances in Neural Information Processing Systems, Vol. 30

[3] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. Von Lilienfeld, Scientific data 1, 1 (2014).