
Proposal: Transformer Models for Predicting Material Properties

Kaiwei Zhang

School of Materials Science and Engineering

Juncheng Yu

School of Aerospace Engineering

Haonan Li

Department of Power and Energy Engineering

1 Background

In recent years, artificial neural networks have emerged as a novel approach to acquire the properties of materials. Models like SchNet [1], CGCNN [2], MEGNet [3], ALIGNN [4], etc. have reached high accuracies on various properties such as formation energy, moduli, band gap, etc. They have also been used to discover new materials [5], guide material synthesis [6], uncover structure-property relationships [7], and are also proven useful in predicting many specific properties in refined areas [8, 9].

Currently, leading models are mostly GNNs and such architecture has long dominated the field. A node in the graph represents an atom, whereas an edge represents a bond. A problem that GNNs have is how to assign the contribution of different sites and atoms to the predicted value. For example, when it comes to perovskite, a material with the formula ABX_3 as shown in Figure 1(a), it is difficult to determine whether the A site is more important to the bandgap, if the B site influences energy, or whether the X element matters. This difficulty arises because global pooling ignores the diversity of nodes in the graph.

Another problem of current models is that they are not adept at capturing global features of crystal structures. For a graph with N nodes, it could take $O(N)$ layers for a node to receive the features of the farthest node if the nodes are sparsely connected, which is the case if materials are viewed as crystal graphs. Nevertheless, these models are not deep enough for information to travel across the whole graph, as shown in Figure 1(b). As a result, current models are not exceptional at properties related to the global features of a material. A typical example of a material’s global feature is its symmetry (shown in Figure 1(c)), which is strongly related to properties such as piezoelectricity and dielectricity. For a material to be piezoelectric, it must lack inversion symmetry. The dielectric constant tends to be higher when materials possess low symmetry. A good prediction of piezoelectric and dielectric constants must start with assessing the global symmetry. As current models lack a global viewpoint, they perform unsatisfactorily on such tasks. For instance, ALIGNN [4] predicts formation and total energy much better than piezoelectric and dielectric constants. Energy prediction requires less global information, as one can estimate the total energy by adding up the energies of each atom [10].

Transformers [11] may be applied to the prediction of crystal properties. We can simply view the atoms in a material’s lattice as a set of input tokens, and the coordinates can be treated by using methods similar to positional encoding. By using the attention mechanism, transformers are well-suited for tackling the aforementioned problems. The attention matrix could be utilized to assess the individual contribution and atomic interactions. On the other hand, the maximum path length of attention mechanism is $O(1)$ [11], making it ideal for grasping global features. It should also be noted that transformers are flexible to the length of the input, which is consistent with the nature of materials considering that the number of atoms inside a lattice could vary substantially.

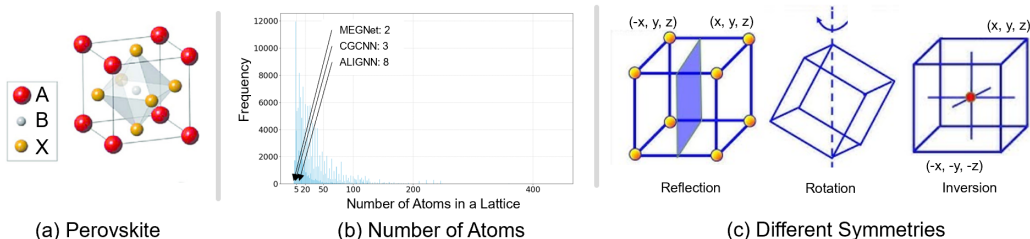


Figure 1: Illustrations

2 Related work

GNN models Various GNN models have been developed to predict the properties of materials. CGCNN [2] uses graph convolutional layers and is interpretable in a few cases by observing local chemical environments. The MEGNet [3] uses the algorithm of graph networks [12] of sequentially updating the bonds, atoms, and global attributes of a material, resulting in improved behavior. ALIGNN [4] further takes into account the impact of angles between bonds. It involves two parallel graphs with linking message pathways. Models such as BonDNet [13] and DimeNet [14] also belong to the GNN family, yet they are developed for molecules and do not work well on crystals.

Attention mechanism There have been attempts to apply attention mechanisms. GATGNN [15] uses several augmented graph attention layers (AGAT) to learn the local relationship between adjacent atoms and a global attention layer is applied at last. CTGNN [16] uses transformer aside from GNN, yet its prediction is limited to energies. It would be reluctant to recognize these models as fully effective solutions for the problems of contribution assignment and global information extraction.

3 Proposed method

Model training Our research consists of four parts. To start with, we will train a model for predicting material properties. In terms of model structure, we are hoping for a transformer model. Methods like structural encoding [17] could help transformer models adapt to graphic tasks. As for training data, Materials Project [18] is a dataset that holds calculated properties of over 150000 materials. OQMD [19], JARVIS-DFT [20] and Google’s GNoME [21] could also enrich the potential dataset. Existing benchmarks [22] will be used to evaluate our model’s capabilities in predicting formation energies, bandgaps, moduli, etc.

Attention analysis Secondly, we expect to assign contributions of each atom by analyzing the attention weights. Following the design of BERT [23], an additional classifier token [cls] could be applied to track the individual influences on the final output. If the results are ideal, they could be used to observe the contribution of different sites and elements inside a crystal. Besides, attention could also be used to study the interactions between atoms. We hope that the model could intrinsically generate higher attention weights between atoms that have smaller distances or stronger chemical bonds. To achieve a better interpretation of the learned attention, techniques including attention rollout, attention flow [24], and gradient self-attention [25] maps could be utilized.

Global view demonstration Thirdly, we wish to demonstrate the transformer’s advantage of capturing global information. The model’s ability to handle symmetry information could be illustrated by comparing the distribution of encoded vectors with point group categories. We also expect our model to reach better accuracy at predicting properties related to global information, such as piezoelectric dielectric constant.

Further discussion Lastly, if the previous sections are smoothly accomplished, we hope to further investigate the potential of our model. We could compare it with fine-tuned LLMs [26, 27] in sorting tasks, and investigate the element embeddings [10] for chemical instincts. We could also use our model to discover novel materials within a ternary or quaternary system [5].

References

- [1] Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet—a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24), 2018.
- [2] Tian Xie and Jeffrey C Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, 120(14):145301, 2018.
- [3] Chi Chen, Weike Ye, Yunxing Zuo, Chen Zheng, and Shyue Ping Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9):3564–3572, 2019.
- [4] Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, 2021.
- [5] Weiye Xia, Ling Tang, Huaijun Sun, Chao Zhang, Kai-Ming Ho, Gayatri Viswanathan, Kirill Kovnir, and Cai-Zhuang Wang. Accelerating materials discovery using integrated deep machine learning approaches. *Journal of Materials Chemistry A*, 11(47):25973–25982, 2023.
- [6] Xingjun Duan, Zhi Fang, Tao Yang, Chunyu Guo, Zhongkang Han, Debalaya Sarker, Xinmei Hou, and Enhui Wang. Maximizing the mechanical performance of ti3alc2 -based max phases with aid of machine learning. *Journal of Advanced Ceramics*, 11(8):1307–1318, 2022.
- [7] Xiang Zhang, Zichun Zhou, Chen Ming, and Yi-Yang Sun. Gpt-assisted learning of structure–property relationships by graph neural networks: Application to rare-earth-doped phosphors. *The Journal of Physical Chemistry Letters*, 14(50):11342–11349, 2023.
- [8] Tim Hsu, Tuan Anh Pham, Nathan Keilbart, Stephen Weitzner, James Chapman, Penghao Xiao, S Roger Qiu, Xiao Chen, and Brandon C Wood. Efficient and interpretable graph network representation for angle-dependent properties applied to optical spectroscopy. *npj Computational Materials*, 8(1):151, 2022.
- [9] Koushik Pal, Cheol Woo Park, Yi Xia, Jiahong Shen, and Chris Wolverton. Scale-invariant machine-learning model accelerates the discovery of quaternary chalcogenides with ultralow lattice thermal conductivity. *npj Computational Materials*, 8(1):48, 2022.
- [10] Tian Xie and Jeffrey C Grossman. Hierarchical visualization of materials space with graph convolutional neural networks. *The Journal of chemical physics*, 149(17), 2018.
- [11] A Vaswani. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017.
- [12] Peter W Battaglia, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.
- [13] Mingjian Wen, Samuel M Blau, Evan Walter Clark Spotte-Smith, Shyam Dwaraknath, and Kristin A Persson. Bondnet: a graph neural network for the prediction of bond dissociation energies for charged molecules. *Chemical science*, 12(5):1858–1868, 2021.
- [14] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*, 2020.
- [15] Steph-Yves Louis, Yong Zhao, Alireza Nasiri, Xiran Wang, Yuqi Song, Fei Liu, and Jianjun Hu. Graph convolutional neural networks with global attention for improved materials property prediction. *Physical Chemistry Chemical Physics*, 22(32):18141–18148, 2020.
- [16] Zijian Du, Luozhijie Jin, Le Shu, Yan Cen, Yuanfeng Xu, Yongfeng Mei, and Hao Zhang. Ctgnn: Crystal transformer graph neural network for crystal material property prediction. *arXiv preprint arXiv:2405.11502*, 2024.

- [17] Chengxuan Ying, Tianle Cai, Shengjie Luo, Shuxin Zheng, Guolin Ke, Di He, Yanming Shen, and Tie-Yan Liu. Do transformers really perform badly for graph representation? *Advances in neural information processing systems*, 34:28877–28888, 2021.
- [18] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.
- [19] Scott Kirklin, James E Saal, Bryce Meredig, Alex Thompson, Jeff W Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. The open quantum materials database (oqmd): assessing the accuracy of dft formation energies. *npj Computational Materials*, 1(1):1–15, 2015.
- [20] Kamal Choudhary, Kevin F Garrity, Andrew CE Reid, Brian DeCost, Adam J Biacchi, Angela R Hight Walker, Zachary Trautt, Jason Hattrick-Simpers, A Gilad Kusne, Andrea Centrone, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj computational materials*, 6(1):173, 2020.
- [21] Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.
- [22] Alexander Dunn, Qi Wang, Alex Ganose, Daniel Dopp, and Anubhav Jain. Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm. *npj Computational Materials*, 6(1):138, 2020.
- [23] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, volume 1, page 2. Minneapolis, Minnesota, 2019.
- [24] Samira Abnar and Willem Zuidema. Quantifying attention flow in transformers. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4190–4197, Online, July 2020. Association for Computational Linguistics.
- [25] Oren Barkan, Edan Hauon, Avi Caciularu, Ori Katz, Itzik Malkiel, Omri Armstrong, and Noam Koenigstein. Grad-sam: Explaining transformers via gradient self-attention maps. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 2882–2887, 2021.
- [26] Kevin Maik Jablonka, Philippe Schwaller, Andres Ortega-Guerrero, and Berend Smit. Leveraging large language models for predictive chemistry. *Nature Machine Intelligence*, 6(2):161–169, 2024.
- [27] Seongmin Kim, Yousung Jung, and Joshua Schrier. Large language models for inorganic synthesis predictions. *Journal of the American Chemical Society*, 146(29):19654–19659, 2024.