Self-Supervised Graph Neural Networks for Accurate Prediction of Néel Temperature

Jian-Gang Kong(孔建刚)[1], Qing-Xu Li(李清旭)[1,2], Jian Li(李健)[1,2,3],
Yu Liu(刘羽)[4], and Jia-Ji Zhu(朱家骥)[1,2,3*]

[1]*School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[2]*Institute for Advanced Sciences, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[3]*Southwest Center for Theoretical Physics, Chongqing University, Chongqing 401331, China*
[4]*Inspur Electronic Information Industry Co., Ltd, Beijing 100085, China*

Antiferromagnetic materials are exciting quantum materials with rich physics and great potential for applications. On the other hand, an accurate and efficient theoretical method is highly demanded for determining critical transition temperatures, Néel temperatures, of antiferromagnetic materials. The powerful graph neural networks (GNNs) that succeed in predicting material properties lose their advantage in predicting magnetic properties due to the small dataset of magnetic materials, while conventional machine learning models heavily depend on the quality of material descriptors. We propose a new strategy to extract high-level material representations by utilizing self-supervised training of GNNs on large-scale unlabeled datasets. According to the dimensional reduction analysis, we find that the learned knowledge about elements and magnetism transfers to the generated atomic vector representations. Compared with popular manually constructed descriptors and crystal graph convolutional neural networks, self-supervised material representations can help us to obtain a more accurate and efficient model for Néel temperatures, and the trained model can successfully predict high Néel temperature antiferromagnetic materials. Our self-supervised GNN may serve as a universal pre-training framework for various material properties.

Antiferromagnetic materials are an exciting class of quantum materials with rich physics in condensed matter theory and many practical applications. In theoretical aspects, antiferromagnetic materials are the parent materials of a series of physical phenomena, such as high-temperature superconductivity,[1] spin liquids,[2] quantum anomalous Hall effect,[3] and topological axion insulator.[4] In application aspects, antiferromagnetic materials can be used for implementing spin valves,[5] colossal magnetoresistive effect,[6] room-temperature electrical switching,[7] fast magnetic moment dynamics,[8] and other antiferromagnetic spintronic applications.[9] The most crucial parameter for antiferromagnetic materials is the Néel temperature, which marks the antiferromagnetic ordering transition, similar to the Curie temperature of a ferromagnet. The experimental determination of the Néel temperature often requires long periods and high costs, while the theoretical predictions with analytical or numerical approaches are highly non-trivial. It is usually quite tedious or challenging to specify the microscopic Hamiltonian corresponding to the antiferromagnetic material and to determine the type and strength of magnetic interactions.[10] The most ordinary theoretical methods are based on the mean field theory, which is not very effective due to the divergent length scale near the critical point. The quantum Monte Carlo method suffers the well-known negative sign problem,[11] and the powerful tensor network method bears difficulty in dealing with physical properties at finite temperatures in two and higher dimensions.[12−14] Therefore, it is of great significance to develop methods that can predict the Néel temperature both accurately and efficiently.

Machine learning aims to fit a predictive model or to find patterns from data, which has already achieved great success in several physical scenes, such as detecting phase transitions,[15−17] accelerating numerical simulations,[18] predicting physical quantities,[19,20] and inverse-designing materials.[21] Recently, graph-based neural networks[22−24] demonstrate the state of the art prediction performance in various material properties with a large amount of data. Especially, the crystal graph convolutional neural network (CGCNN)[22] is used to fit a function between the graph representation of materials and the target material properties. Based on the flexible CGCNN framework, one can either encode more physical information into the crystal graph[25,26] or further optimize the edges.[26]

---

*Corresponding author. Email: zhujj@cqupt.edu.cn

In predicting the transition temperature of magnetic materials with machine learning methods, most studies focus only on the Curie temperature of ferromagnetic materials. For example, Nelson *et al.* showed the model of best performance among the random forest, kernel ridge regression (KRR), and neural network model achieves a mean absolute error (MAE), 57 K, on a ferromagnetic material dataset of size 2500.[27] The random forest model trained by Long *et al.* reaches an accuracy of 0.87 to distinguish 1749 ferromagnetic and 1056 antiferromagnetic materials, while the regression analysis of Curie temperature shows that the MAE is about 55 K.[28] Other efforts focus on the impact of different inputs on the prediction accuracy, for instance, constructing 21 descriptive variables as inputs of the KRR algorithm to predict the Curie temperature of transition metal-rare earth compounds.[29] However, there are very few reports on predicting Néel temperature by machine learning, such as a support vector regression (SVR) model trained on perovskite manganese oxides of size 127, which achieves an rms error of 32.3 K on a test set of size 32.[30]

The scarcity of data is the grave difficulty in machine learning for the transition temperature of magnetic materials. There is still no large-scale computational dataset for magnetic materials with calculated transition temperatures. A recent work[31] has mined a large number of published papers for data and constructed a magnetic phase transition temperature dataset of size 40000, utilizing tools from the field of natural language processing and relation extraction techniques. Unfortunately, the complete material structures are absent in the constructed dataset, where the labels also lack quality assurance. Due to the scarcity of magnetic material data, most of the machine learning studies on magnetic materials make use of the ensemble learning algorithms, such as random forests, rather than deep learning algorithms, such as graph neural networks (GNN), on small dataset.[32] Yet, the performance of the former highly relies on the quality of material descriptors.[33]

Self-supervised learning is a new method to train neural networks without the need for expensive labels, utilizing supervised signals from the content or the intrinsic structure of the data itself. As a pre-training strategy for models with many parameters, self-supervised learning has gained significant interests in natural language processing[34] and computer vision.[35] In the graph machine learning field, it is proposed that a pre-training method on large datasets can recover the randomly masked elemental information in molecular graphs and improve prediction performance on various small molecular datasets.[36] Since the crystal graphs representing crystalline materials contain rich physical information, it is believed that different types of self-supervised learning on large-scale crystal graphs can capture prior solid knowledge and lead to performance improvement in downstream tasks.

In this Letter, we propose to combine the representation learning capabilities of GNN with the efficiency of the standard machine learning model on the Néel temperature dataset. We extract high-level representations of materials in a self-supervised manner by training the GNN to reproduce elemental information and magnetic moments. As the input of the regression model, self-supervised material descriptors can outperform popular material descriptors and the powerful CGCNN, due to their simplicity, high-relevance, and low computational cost. The trained KRR model possesses the capability to screen magnetic materials with high Néel temperatures from database for spintronic applications.

Self-Supervised Learning on Crystal Graphs. The crystal graph is a multi-graph characterized by node vectors, edge vectors, and adjacent matrix, allowing multiple edges between the same pair of nodes due to the periodicity of crystalline materials. The nodes of the graph encode elemental information of atoms, such as the period number and the group number. The edges of the graph encode the distance between atoms which implicitly capture interactions or bondings. The node vector $\boldsymbol{v}_i$ in the crystal graph is updated by the tth GNN layer as follows:

$$\boldsymbol{v}_i^{(t+1)} = \boldsymbol{v}_i^{(t)} + \sum_{j,k} \sigma\left(\boldsymbol{z}_{(i,j)_k}^{(t)} \boldsymbol{W}_f^{(t)} + \boldsymbol{b}_f^{(t)}\right)$$
$$\odot\, g\left(\boldsymbol{z}_{(i,j)_k}^{(t)} \boldsymbol{W}_s^{(t)} + \boldsymbol{b}_s^{(t)}\right).$$

The node $\boldsymbol{v}_i$ updates itself by aggregating the messages provided by its neighbors, which belongs to the message passing mechanism.[37] $\boldsymbol{W}_f$ and $\boldsymbol{W}_s$ are learnable weight matrices acting on a pair of neighbors $\boldsymbol{z}_{(i,j)_k} = \boldsymbol{v}_i \oplus \boldsymbol{u}_{(i,j)_k} \oplus \boldsymbol{v}_j$ with $\boldsymbol{u}_{(i,j)_k}$ representing the vector corresponding to the kth edge between node i and node j, and $\oplus$ denoting the concatenation of two vectors. $\boldsymbol{b}_f$ and $\boldsymbol{b}_s$ are the corresponding biases of linear transformations. The σ and g are nonlinear activation functions that can increase the expressive power of neural networks and filter out the most important bonds during training. The $\odot$ denotes the element-wise product, or the Hadamard product, between two vectors.

In order to make full use of the sizeable unlabeled dataset merely with information about elements and structures, we propose to extract chemical rules by self-supervised learning on a computational material dataset of size 60000, constructed from the Materials Project.[38] More specifically, as shown in Fig. 1(a), we randomly mask the elemental information, the group number and the period number, of a pre-defined proportion of atoms in the unit cell during training.

A GNN is then trained to recover the information based on the surrounding crystal environments of the masked atoms. Also, we randomly mask the edge information, the distance between atoms, of the masked nodes, and train the neural network to correctly predict the distance information. We expect the GNN to learn high-level knowledge of crystal structures and chemical properties by self-supervised training and save it in the form of neural network weights. The trained GNN is denoted as Elem-GNN.
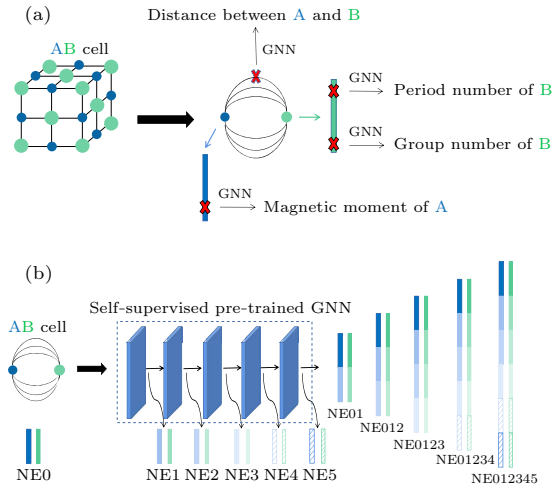

Fig. 1. (a) Self-supervised training on crystal graphs (AB cell as an illustrative example). The green (blue) balls represent B (A) atoms, the green (blue) bar is the node vector corresponding to B (A) atom, and the red crosses denote randomly masked node or edge information during training. A GNN is trained to recover the masked information based on the surrounding crystal environment. (b) The 64-dimensional atomic vectors generated from the self-supervised pre-trained GNN, given a crystal graph as input (AB cell as an illustrative example). The green (blue) bars at the bottom are single-scale atomic vectors of different layers, while those on the right are multi-scale atomic vectors.

We also perform self-supervised learning on a computational magnetic material dataset of size 50000, constructed from the Materials Project, to extract knowledge about magnetism. Similar to the self-training process described above, we still randomly mask the elemental and distance information of each material. Some of the masked atoms have nonzero magnetic moments in this case. Therefore, we can train another GNN, which we refer to as Mag-GNN, to reproduce the size of the magnetic moments of masked atoms.

Given the original crystal graphs as inputs, we can generate 64-dimensional node embeddings (NE) by utilizing the self-supervised pre-trained GNN, i.e., Elem-GNN and Mag-GNN, as shown in Fig. 1(b). We name the atomic vectors Elem-NEs and Mag-NEs from the Elem-GNN and Mag-GNN, respectively. We can transfer the knowledge obtained by self-supervised training from GNN to the generated atomic repre-

sentations. By averaging over the NEs in the same unit cell, we can obtain the vector representation, or the graph embeddings (GE), of the materials, which can be directly taken as the input vectors of machine learning models for studying material properties. We use a five-layer GNN in the self-supervised training and output the corresponding self-supervised atomic vector NEi ($i = 0, 1, 2, 3, 4, 5$) from each layer. Then we obtain multi-scale atomic representations NE01, NE012,…, and NE012345 by concatenating atomic representations of different layers and thus obtain multi-scale material representations GE01, GE012,…, and GE012345. The necessities of the multi-scale representations are two-fold: (1) The GNN updates the atomic representations through the message passing mechanism, and the atomic vectors of the deeper GNN layer receive a more extensive range of environmental information. (2) The deep GNN, however, suffers the so-called over-smoothing problem,[39] which results in the degradation of the performance from the similarities of the node representations of deep GNN layer. The CGCNN is also perplexed by the same problem due to the sizeable effective range of single message passing for unit cells with a few atoms. Multi-scale material representations can reduce the similarities between the atomic representations and offer us additional freedom by controlling the amount of environmental information in descriptors.
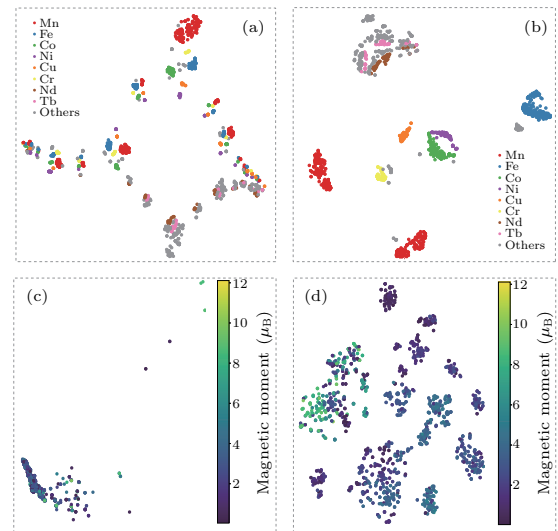

Fig. 2. The t-SNE visualization of self-supervised atomic vectors Elem-NEs, Mag-NEs, and Random-NEs labeled by element types and magnetic moments. Different colors represent different elements or moments. (a) Two-dimensional distribution of Random-NEs under element labels. (b) Two-dimensional distribution of Elem-NEs under element labels. (c) Two-dimensional distribution of Random-NEs labeled by magnetic moments. (d) Two-dimensional distribution of Mag-NEs labeled by magnetic moments.

Dimensional Reduction Visualization. To examine effects of self-supervised training and to make

sure that the learned knowledge is transferred to atomic vectors, we visualize the distribution of two-dimensional (2D) points corresponding to the 64-dimensional self-supervised atomic vectors Elem-NEs, Mag-NEs, and randomly initialized atomic vectors Random-NEs by utilizing the t-SNE (t-distributed stochastic neighbor embedding) technique.[40] The 2D data points obtained by dimensional reduction of Elem-NEs are labeled with element types, and those of Mag-NEs are marked with the size of magnetic moments. We expect that the patterns formed by the self-supervised atomic vectors are more regular than the random atomic vectors, which highlights the benefits of self-supervised learning. The magnetic moment dataset used for visualization is constructed from MAGNDATA,[41,42] containing 1816 magnetic atoms from 29 elements, including 13 transition metals and 12 lanthanides.

From Figs. 2(a) and 2(b), we can see that the atomic vectors Elem-NEs form more well-clustered patterns under element labels than Random-NEs, especially for transition metals. The most abundant Mn atoms form several smaller clusters, which may result from competition between different local crystal environments. Most lanthanides, such as Nd and Tb, are also well clustered in the upper region, as shown in Fig. 2(b). In contrast, Random-NEs shown in Fig. 2(a) have no clear organizing patterns, which indicates that Elem-NEs indeed contain richer chemical information. The distribution of magnetic moments is shown in Figs. 2(c) and 2(d): The points of Random-NEs shrink into a small region, and the magnetic moments of different sizes are mixed. The points of Mag-NEs are more uniformly distributed, and the magnetic moments of different sizes are distinguished. For example, the magnetic moments in the range $6\mu_{\mathrm{B}}$–$12\mu_{\mathrm{B}}$ are primarily distributed in the left middle region, while the magnetic moments in the range $4\mu_{\mathrm{B}}$–$6\mu_{\mathrm{B}}$ are distributed in the right and bottom region. The dimensional reduction analysis verifies that Mag-NEs contain more knowledge about magnetism, and we can indeed extract useful physical information from the self-supervised pre-training process.

Prediction Performance on Experimental Néel Temperature Dataset. To further understand the performance of self-supervised material representations, GEs, we train and evaluate a standard machine learning model with different material representations on the experimental Néel temperature dataset. The dataset contains a total of 1007 antiferromagnetic materials in the MAGNDATA database with corresponding magnetic structures and experimentally measured Néel temperatures. However, we can find that the distribution of materials with different Néel temperatures is very unbalanced, as shown in Fig. 3. Most materials are in the low-temperature zone, and 881 materials are below 200 K, accounting for 80% of the total materials. This imbalance of distribution poses tough challenges to machine learning algorithms.
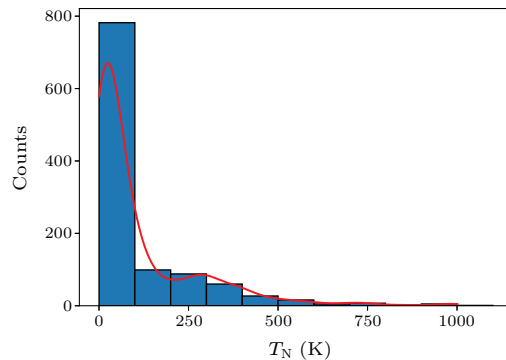

Fig. 3. Distribution of experimental antiferromagnetic materials in different Néel temperature ranges.

Table 1. The prediction performance of the KRR model on the Néel temperature sub-dataset when the manually constructed descriptors ESM, SM, and OFM are taken as inputs.

	ESM	SM	OFM
MAE	95.05 ± 10.20	86.53 ± 4.97	75.83 ± 6.56
R2	0.13 ± 0.09	0.21 ± 0.11	0.19 ± 0.15

The machine learning model we employed is KRR, a nonlinear regression algorithm that combines kernel trick and ridge regression, which is widely used in material science. We compare the popular material representations, such as sine matrix (SM), Ewald sum matrix (ESM),[43] and orbital field matrix (OFM),[44] with our proposed multi-scale self-supervised material representations Elem-GEs and Mag-GEs. The SM and ESM can capture Coulomb interactions of crystalline materials, and their lengths are square of the maximum number of atoms in the unit cell. The OFM can encode orbital interactions of materials, and its length is 1056, independent of the dataset. In contrast, the length of self-supervised material representations is only 64 per scale, independent of the dataset as well. Therefore, we construct a dataset of 748 materials keeping magnetic materials of a unit cell with fewer than 60 atoms for comparison, where SM and ESM are 3600-dimensional vectors for this sub-dataset. In order to evaluate the performance of the model, we divide the dataset into a training set and a test set by 8 : 2 and then perform 10-fold cross-validation and grid search on the training set for the best hyperparameters. Given the optimal hyperparameters, we re-train the model on the whole training set and find the performance scores on the test set. The above procedure repeats five times with random train-test splits of dataset. The final model performance is given by the mean and standard deviation of scores on five different test sets, and the evaluation process is more reliable for a small dataset.

From the results shown in Table 1, we can see that the traditional, manually constructed material descriptors ESM and SM have poor prediction performance on Néel temperature, not only the large MAE but also the small R2 score. The reason for the failure may lie in the fact that the lengths of SM and ESM are too long on a small dataset, whose information is not relevant enough to the Néel temperature. The prediction error of OFM is lower than that of SM and ESM, which may be attributed to the orbital interactions encoded in OFM. The vector representation of OFM is also more compact than SM and ESM.
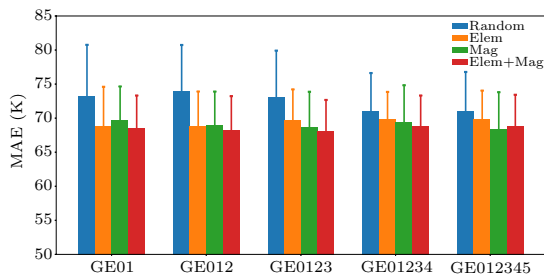

Fig. 4. Prediction performance of the KRR model on the Néel temperature sub-dataset when Random-GEs, Elem-GEs, Mag-GEs, and Elem+Mag-GEs are taken as inputs.

Next, we examine the prediction performance of the KRR model on the Néel temperature sub-dataset, taking the Elem-GEs and Mag-GEs as inputs. In order to further verify the effects of self-supervised learning, we denote the multi-scale material descriptors generated by the randomly initialized GNN as Random-GE. We have a few remarks on the effects of self-supervised learning shown in Fig. 4. (1) Random-GEs already have relatively good performance despite the absence of a specific type of knowledge through self-supervised training. The MAE obtained by Random-GE01 is 73.23 K, which is slightly better than OFM. However, the length of Random-GE01 is only 128, which is much shorter than 1056, the length of OFM. The shorter length of the former indicates that Random-GE01 contains rich physical information more compactly. The reason is evident since the original crystal graph already contains sufficient information about materials, and atomic and material

representations obtained by message passing on the crystal graph can naturally acquire information about elements and structures. (2) For material descriptors at each scale, the prediction error of self-supervised representations is systematically lower than that of random representations, i.e., the MAE of Elem-GEs or Mag-GEs is smaller than the MAE of Random-GEs, which proves that self-supervised pre-training is indeed beneficial. Specifically, we regularize the random weights in a more significant way for material properties by performing self-supervised training. Then we transfer it to the vector representation of the material and gain a more accurate machine learning model. (3) As we can see, by virtue of the combination of two types of knowledge, i.e., chemical rules and magnetism, combining the Elem-GEs and Mag-GEs (Elem+Mag-GEs) achieves the lowest prediction error on all scales except GE012345. The exception of GE012345 is due to the competition between the physical information in the descriptor and the vector length of the descriptor. (4) The self-supervised representations, Elem-GE012, Mag-GE012345, and Elem+Mag-GE0123, achieve the lowest prediction errors of 68.78 K, 68.34 K, 68.10 K, respectively. Their R2 scores are 0.32, 0.34, 0.32, respectively. The performance of the self-supervised representations is much better than popular manually constructed descriptors, e.g., ESM, SM, and OFM, at a similar overall training cost of time, as shown in Fig. 5.
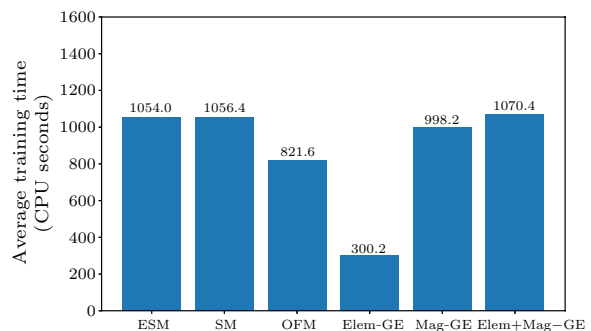

Fig. 5. Average training time of the KRR model on five training sets, given different material representations as inputs.

Table 2. Prediction performance of CGCNN and the KRR model on the full Néel temperature dataset. OFM, Mag-GE01, Elem-GE01, and Elem+Mag-GE01 are taken as inputs of the KRR model.

	OFM	CGCNN	Mag–GE01	Elem–GE01	Elem+Mag–GE01
MAE	65.14 ± 5.95	63.85 ± 6.85	59.72 ± 6.15	58.32 ± 4.70	58.23 ± 5.08
R2	0.42 ± 0.06	0.42 ± 0.06	0.50 ± 0.04	0.54 ± 0.05	0.54 ± 0.05

After demonstrating the advantages of self-supervised multi-scale material representations over traditional descriptors SM and ESM on a sub-dataset of size 748, we train a better model by utilizing the whole Néel temperature dataset of size 1007 and compare it with OFM and CGCNN. As shown in Table 2,

the performance of OFM is the worst, which is consistent with our previous observations on the subset, illustrating the limitation of manually constructed descriptors. The powerful CGCNN also fails to achieve good prediction performance due to the limitation of the size of the dataset, and the sophisticated tuning

procedure of hyperparameters makes CGCNN less efficient compared with the KRR model. However, taking the self-supervised multi-scale material representations as inputs of the KRR model, the prediction accuracy is better than CGCNN. The training is also more efficient and flexible.
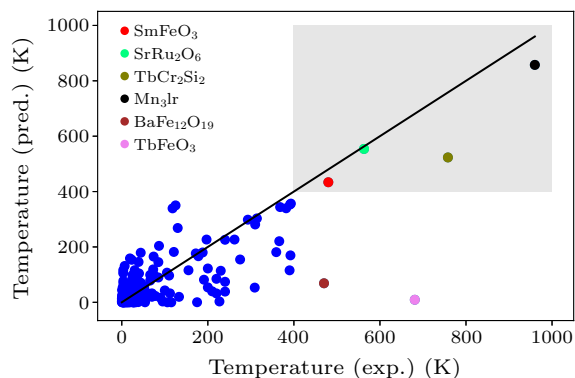

Fig. 6. Parity plot of the prediction performance of the trained KRR model on the test set when Elem-GE01 is taken as the input. The six antiferromagnetic materials in the high Néel temperature regime (exp. > 400 K) are colored differently, and the four successfully predicted materials are shown in the gray zone.

We can further analyze the trained model by drawing its parity plot, i.e., experimental values versus predicted values on the test set. For instance, the parity plot of Elem-GE01 on one of five test sets is shown in Fig. 6. We focus on the predictions of antiferromagnetic materials with the experimental Néel temperatures in the high-temperature regime, i.e., above 400 K, since it is more relevant for applications in the field of antiferromagnetic spintronics.[45] We find that the predicted high Néel temperatures of most materials agree with the experimental high Néel temperatures, which is shown in the grey zone of Fig. 6. For example, the predicted Néel temperatures are 433.65 K for $SmFeO_3$ and 553.99 K for $SrRu_2O_6$, respectively, with relative errors of only 9.6% to the experimental value 480 K of $SmFeO_3$[46] and 1.6% to 563 K of $SrRu_2O_6$.[47] Particularly, the antiferromagnetic material Mn_3Ir[48] has the highest Néel temperature of 960 K in the test set, and the trained KRR model also finds the highest predicted value of 857 K, with a relative error of only 10.6%, showing that the trained model owns the ability to screen high-temperature antiferromagnetic materials in a highthroughput way. However, for $BaFe_{12}O_{19}$[49] and $TbFeO_3$,[50] the predictions are completely failed. The dataset lacks sufficient materials covering more diverse chemical compositions and crystal structures. If we can expand the dataset by including more types of antiferromagnetic materials, exceptional failures like $BaFe_{12}O_{19}$ and $TbFeO_3$ would be erased.

Conclusion and Outlook. We have proposed a self-supervised training strategy of CGCNN to extract material representations with rich physical information. The combination of the self-supervised material representations and the KRR model outperforms popular manually constructed descriptors as well as CGCNN on the Néel temperature dataset. The self-supervised GNN may also serve as a universal pre-training framework for various material properties. Although the trained model shows the ability to screen high Néel temperature materials, there are still challenges to gain a more reliable and accurate model. First, the proper encoding of the information of magnetic structures, i.e., the value and direction of magnetic moments, is of great significance since the magnetic materials with different Néel temperatures could have the same chemical composition and crystal structures but different magnetic structures. The self-supervised material descriptors cannot distinguish them, and geometric deep learning may be a promising solution.[51] Second, considering the absence of a large-scale, high-quality computational dataset of magnetic materials, we can further perform transfer learning[52−54] on low-fidelity datasets, which is also an effective strategy to improve the prediction performance of the model on small datasets.

References

[1] Lee P A, Nagaosa N and Wen X G 2006 *Rev. Mod. Phys.* **78** 17

[2] Zhou Y, Kanoda K and Ng T K 2017 *Rev. Mod. Phys.* **89** 025003

[3] Qiao Z, Ren W, Chen H, Bellaiche L, Zhang Z, MacDonald A H and Niu Q 2014 *Phys. Rev. Lett.* **112** 116404

[4] Liu C, Wang Y, Li H, Wu Y, Li Y, Li J, He K, Xu Y, Zhang J and Wang Y 2020 *Nat. Mater.* **19** 522

[5] Park B G, Wunderlich J, Martí X, Holỳ V, Kurosaki Y, Yamada M, Yamamoto H, Nishide A, Hayakawa J, Takahashi H, Shick A B and Jungwirth T 2011 *Nat. Mater.* **10** 347

[6] Qiu Z, Hou D, Barker J, Yamamoto K, Gomonay O and Saitoh E 2018 *Nat. Mater.* **17** 577

[7] Wadley P, Howells B, Železnỳ J, Andrews C, Hills V, Campion R P, Novák V, Olejník K, Maccherozzi F, SS D, Martin S Y, Wanger T, Wunderlich J, Freimuth F, Mokrosov Y, Kuneš J, Chauhan J S, Grzybowski M J, Rushforth A W, Edmonds K W, Gallagher B L and Jungwirth T 2016 *Science* **351** 587

[8] Jungwirth T, Marti X, Wadley P and Wunderlich J 2016 *Nat. Nanotechnol.* **11** 231

[9] Baltz V, Manchon A, Tsoi M, Moriyama T, Ono T and Tserkovnyak Y 2018 *Rev. Mod. Phys.* **90** 015005

[10] Li X, Hongyu Y, Feng L, Feng J S, Whangbo M H and Xiang H 2021 *Molecules* **26** 803

[11] Loh E Y, Gubernatis J E, Scalettar R T, White S R, Scalapino D J and Sugar R L 1990 *Phys. Rev. B* **41** 9301
[12] Didier P, Matthieu M and Fabien A 2021 *SciPost Phys.* **10** 19
[13] Li W, Ran S J, Gong S S, Zhao Y, Xi B, Ye F and Su G 2011 *Phys. Rev. Lett.* **106** 127202
[14] Czarnik P, Cincio L and Dziarmaga J 2012 *Phys. Rev. B* **86** 245101
[15] Rao W J 2020 *Chin. Phys. Lett.* **37** 080501
[16] Cheng Z and Yu Z 2021 *Chin. Phys. Lett.* **38** 070302
[17] Zhang R, Wei B, Zhang D, Zhu J J and Chang K 2019 *Phys. Rev. B* **99** 094427
[18] Lu H, Li C, Chen B B, Li W, Qi Y and Meng Z Y 2022 *Chin. Phys. Lett.* **39** 050701
[19] Ouyang Y, Zhang Z, Yu C, He J, Yan G and Chen J 2020 *Chin. Phys. Lett.* **37** 126301
[20] Schmidt J, Marques M R G, Botti S and Marques M A L M 2019 *npj Comput. Mater.* **5** 83
[21] Noh J, Gu G H, Kim S and Jung Y 2020 *Chem. Sci.* **11** 4871
[22] Xie T and Grossman J C 2018 *Phys. Rev. Lett.* **120** 145301
[23] Chen C, Ye W, Zuo Y, Zheng C and Ong S P 2019 *Chem. Mater.* **31** 3564
[24] Schütt K T, Sauceda H E, Kindermans P J, Tkatchenko A and Müller K R 2018 *J. Chem. Phys.* **148** 241722
[25] Karamad M, Magar R, Shi Y, Siahrostami S, Gates I D and Farimani A B 2020 *Phys. Rev. Mater.* **4** 093801
[26] Park C W and Wolverton C 2020 *Phys. Rev. Mater.* **4** 063801
[27] Nelson J and Sanvito S 2019 *Phys. Rev. Mater.* **3** 104405
[28] Long T, M F N, Zhang Y, Gutfleisch O and Zhang H 2021 *Mater. Res. Lett.* **9** 169
[29] Nguyen D N, Pham T L, Nguyen V C, Nguyen A T, Kino H, Miyake A and Dam H C 2019 *J. Phys.: Conf. Ser.* **1290** 012009
[30] Lu K, Chang D, Lu T, Ji X, Li M and Lu W 2021 *J. Supercond. Novel Magn.* **34** 1961
[31] Court C and Cole J 2020 *npj Comput. Mater.* **6** 18
[32] Dunn A, Wang Q, Ganose A, Dopp D and Jain A 2020 *npj Comput. Mater.* **6** 138
[33] Ghiringhelli L M, Vybiral J, Levchenko S V, Draxl C and Scheffler M 2015 *Phys. Rev. Lett.* **114** 105503
[34] Devlin J, Chang M W, Lee K and Toutanova K 2018 arXiv:1810.04805 [cs.CL]
[35] He K, Fan H, Wu Y, Xie S and Girshick R 2019 arXiv:1911.05722 [cs.CV]
[36] Hu W, Liu B, Gomes J, Zitnik M, Liang P, Pande V and Leskovec J 2019 arXiv:1905.12265 [cs.LG]
[37] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E 2017 arXiv:1704.01212 [cs.LG]
[38] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K A 2013 *APL Mater.* **1** 011002
[39] Li Q, Han Z and Wu X M 2018 arXiv:1801.07606 [cs.LG]
[40] van der Maaten L and Hinton G 2008 *J. Mach. Learn. Res.* **9** 2579
[41] Gallego S V, Perez-Mato J M, Elcoro L, Tasci E S, Hanson R M, Momma K, Aroyo M I and Madariaga G 2016 *J. Appl. Crystallogr.* **49** 1750
[42] Gallego S V, Perez-Mato J M, Elcoro L, Tasci E S, Hanson R M, Momma K, Aroyo M I and Madariaga G 2016 *J. Appl. Crystallogr.* **49** 1941
[43] Faber F, Lindmaa A, von Lilienfeld O A and Armiento R 2015 *Int. J. Quantum Chem.* **115** 1094
[44] Pham T L, Kino H, Terakura K, Miyake T T I, Tsuda K and Dam H C 2017 *Sci. Technol. Adv. Mater.* **18** 756
[45] Jenkins S, Chantrell R W and Evans R F L 2021 *Phys. Rev. B* **103** 014424
[46] Kuo C Y, Drees Y, Fernández-Díaz M T, Zhao L, Vasylechko L, Sheptyakov D, Bell A M T, Pi T W, Lin H J, Wu M K, Pellegrin E, Valvidares S M, Li Z W, Adler P, Todorova A, Küchler R, Steppke A, Tjeng L H, Hu Z and Komarek A C 2014 *Phys. Rev. Lett.* **113** 217203
[47] Hiley C I, Scanlon D O, Sokol A A, Woodley S M, Ganose A M, Sangiao S, De Teresa J M, Manuel P, Khalyavin D D, Walker M, Lees M R and Walton R I 2015 *Phys. Rev. B* **92** 104413
[48] Tomeno I, Fuke H N, Iwasaki H, Sahashi M and Tsunoda Y 1999 *J. Appl. Phys.* **86** 3853
[49] Collomb A, Wolfers P and Obradors X 1986 *J. Magn. Magn. Mater.* **62** 57
[50] EF Bertaut, Chappert J, Mareschal J, Rebouillat J P and Sivardière J 1967 *Solid State Commun.* **5** 293
[51] Bronstein M M, Bruna J, Cohen T and Veličković P 2021 arXiv:2104.13478 [cs.LG]
[52] Xie T, Bapst V, Gaunt A L, Obika A, Back T, Demis H, Kohli P and Kirkpatrick J 2021 arXiv:2103.13795 [cond-mat.mtrl-sci]
[53] Chen C, Zuo Y, Ye W, Li X G and Ong S 2021 *Nat. Comput. Sci.* **1** 46
[54] Lee J and Asahi R 2021 *Comput. Mater. Sci.* **190** 110314

Supplementary Material for "Self-supervised graph neural networks for accurate prediction of Néel temperature"

Jian-Gang Kong[1], Qing-Xu Li[1,2], Jian Li[1,2,3], Yu Liu[4], Jia-Ji Zhu[1,2,3]

[1]School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]Institute for Advanced Sciences, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[3]Southwest Center for Theoretical Physics, Chongqing University, Chongqing 401331, China
[4]Inspur Electronic Information Industry Co., Ltd, Beijing 100085, China

In the supplementary material, we provide more detailed information about magnetic material datasets used in this work, and the equivalence between the self-supervised training of graph neural networks(GNN) and classification tasks.

1   Datasets

MAGNDATA[1, 2] is a high-quality experimental database with more than 1500 published commensurate and incommensurate antiferromagnetic materials, which contains rich information on the magnetic structure and experimentally measured Néel temperatures.

The Néel temperature full dataset constructed from MAGNDATA, excluding the disordered and incommensurate magnetic structures which can't be handled appropriately by pymatgen[3], contains only 1,007 entries, which is relatively small. While the Néel temperature sub-dataset of size 748 is obtained by keeping the materials with a unit cell of fewer than 60 atoms in the entire dataset.

The magnetic moment dataset of size 1816 is constructed from the Néel temperature dataset. Transition metals account for 67% of the total number of atoms in the dataset, and the most frequent elements of transition metals are Mn, Fe, Co, Ni, Cu, and Cr. In contrast, lanthanides account for 25% of the total atoms in the dataset, and the most frequent lanthanide elements are Nd and Tb. The average size of magnetic moments in the dataset is 3.05 μ_B, while the minimum is 0.04 μ_B and the maximum is 12.11μ_B, and 90% of the magnetic moments are smaller than 6 μ_B.

2   The self-supervised training of GNN

In CGCNN[4], the element information and distance between atoms are one-hot encoded in the node vectors and edge vectors, respectively. More specifically, for properties represented by discrete values, such as the period number and group number of elements, the vector is encoded according to the category to which the value belongs. On the other hand, for a property represented by continuous value, for example, the electronegativity of an element, the value will be discretized into ten bins, and the vector will be constructed accordingly.

Therefore, the self-supervised learning tasks of reproducing period number and group number of randomly masked atoms correspond to classification problems of 8 and 18 categories, respectively. We also randomly mask the edge information, the distance between atoms of the masked nodes, and train the neural network to predict the distance information, which corresponds to a classification problem of 41 categories.

We group the magnetic moments represented by continuous values into 64 discrete data bins to reduce the difficulty of the training and extract more information. As a result, the prediction of magnetic moments can also be regarded as a classification problem of 64 categories. The numbers of training epochs, mask ratio, and discrete bins are hyperparameters that can be fine-tuned for the best model.

References

[1] Gallego S V, Perez-Mato J M, Elcoro L, Tasci E S, Hanson Robert M, Momma K, Aroyo M I and Madariaga G 2016 J. Appl. Crystallogr. 49 1750

[2] Gallego S V, Perez-Mato J M, Elcoro L, Tasci E S, Hanson Robert M, Momma K, Aroyo M I and Madariaga G 2016 J. Appl. Crystallogr. 49 1941

[3] Ong S P, Richards W D, Jain A, Hautier G, Kocher M, Cholia S, Gunter D, Chevrier V L, Persson K A and Ceder G 2016 Comput. Mater. Sci. 68 314–319

[4] Xie T and Grossman J C 2018 Phys. Rev. Lett. 120 145301