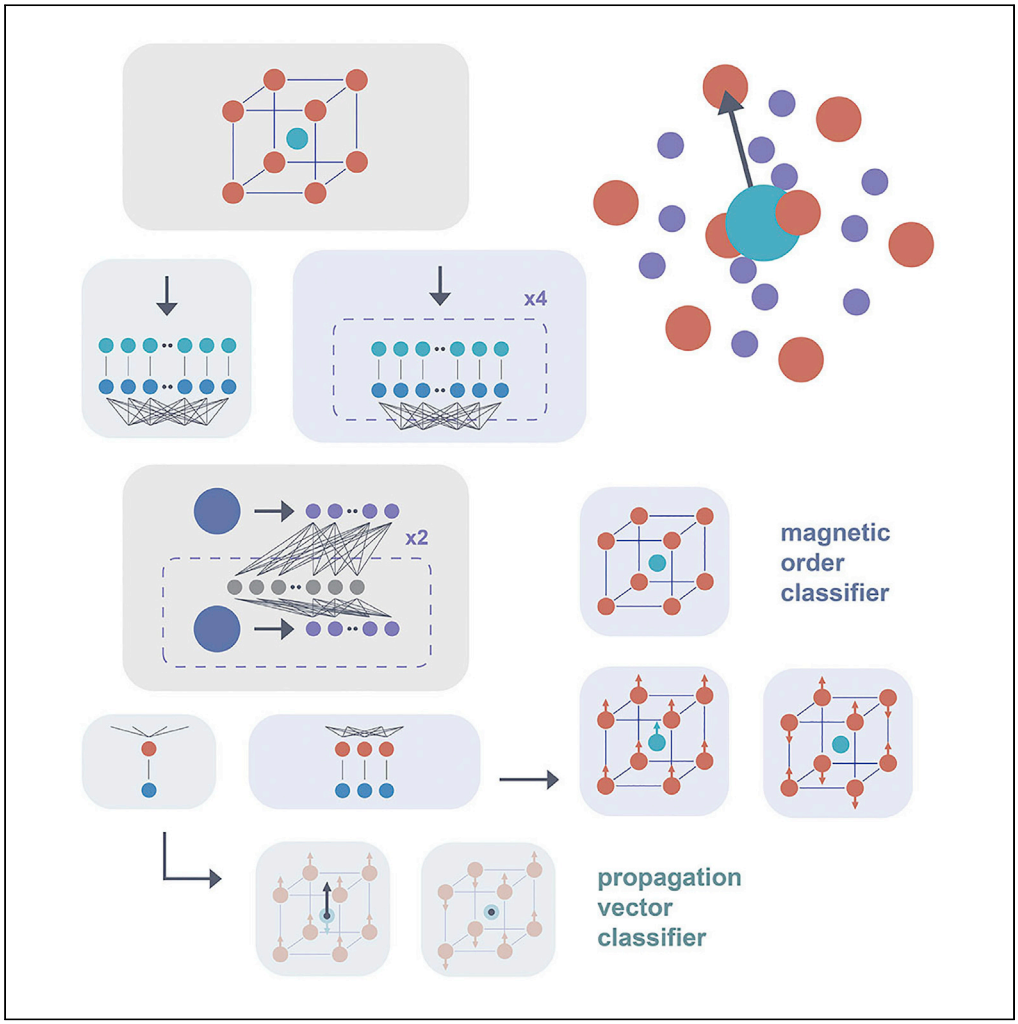
iScience

CellPress
OPEN ACCESS

Article

Machine learning magnetism classifiers from atomic coordinates



Helena A. Merker, Harry Heiberger, Linh Nguyen, ..., Yao Wang, Tess Smidt, Mingda Li

tsmidt@mit.edu (T.S.)
mingda@mit.edu (M.L.)

Highlights

An Euclidean equivariant graph neural network applied to material science

A machine learning model to predict magnetic orders from atomic coordinates

A descriptor to encode general crystal structures of any space group

Merker et al., iScience 25, 105192
October 21, 2022
https://doi.org/10.1016/j.isci.2022.105192

iScience

CellPress
OPEN ACCESS

Article

Machine learning magnetism classifiers from atomic coordinates

Helena A. Merker,[1,2,10] Harry Heiberger,[1,2,10] Linh Nguyen,[1,3,10] Tongtong Liu,[1,3,10] Zhantao Chen,[1,4] Nina Andrejevic,[1,5] Nathan C. Drucker,[1,6] Ryotaro Okabe,[1,7] Song Eun Kim,[2] Yao Wang,[8] Tess Smidt,[2,*] and Mingda Li[1,9,11,*]

SUMMARY

The determination of magnetic structure poses a long-standing challenge in condensed matter physics and materials science. Experimental techniques such as neutron diffraction are resource-limited and require complex structure refinement protocols, while computational approaches such as first-principles density functional theory (DFT) need additional semi-empirical correction, and reliable prediction is still largely limited to collinear magnetism. Here, we present a machine learning model that aims to classify the magnetic structure by inputting atomic coordinates containing transition metal and rare earth elements. By building a Euclidean equivariant neural network that preserves the crystallographic symmetry, the magnetic structure (ferromagnetic, antiferromagnetic, and nonmagnetic) and magnetic propagation vector (zero or non-zero) can be predicted with an average accuracy of 77.8% and 73.6%. In particular, a 91% accuracy is reached when predicting no magnetic ordering even if the structure contains magnetic element(s). Our work represents one step forward to solving the grand challenge of full magnetic structure determination.

INTRODUCTION

As one of the most prominent quantum phenomena, magnetism of materials encompasses a large portion of functional applications such as data storage (Fert, 2008), high-resolution imaging (Vlaardingerbroek and Boer, 2003), spintronic devices (Manchon et al., 2019), high-energy scientific instruments (Artsimovich, 1972; Wiedemann, 2015), and quantum computing (Gershenfeld and Chuang, 1997; Balents, 2010). Particular types of magnetism are also believed to be associated with unconventional quantum phases such as high-T_c and topological superconductivity (Keimer et al., 2015). Unlike small molecules where magnetic structures contain only several high-spin and low-spin configurations, spatial correlations between magnetic moments in sizable materials constitute vast possibilities of different magnetic configurations. With infinite combinations of wavevectors, moments, and correlations lengths, magnetic materials can form a variety of structures such as antiferromagnetism (Shull and Smart, 1949), non-collinear magnetism, skyrmions (Mühlbauer et al., 2009), spin glass (Binder and Young, 1986), and quantum spin liquids (Banerjee et al., 2017; Zhou et al., 2017). Therefore, the determination of magnetic structures, either experimentally or theoretically, is crucial for materials discovery and technological progress in general.

Experimentally, the state-of-the-art neutron scattering (Lovesey, 1986) and more recently resonant X-ray scattering (Ament et al., 2011) have enabled the characterization of magnetic structures with atomic-scale resolution. However, these measurements require large-scale neutron sources or synchrotron X-ray radiation and are highly limited by the capacity and beamtime availability. According to the most comprehensive database, only ~1,500 materials' magnetic structures have been identified through these experimental spectra since the 1950s (Gallego et al., 2016a, 2016b). Therefore, without order-of-magnitude improvements of these facilities' capacity, a pure experimental exploration of magnetic materials is yet to catch up with the rapidly rising demand for materials' discovery of new magnetic materials.

Theoretically, *ab initio* simulations with advanced quantum chemistry and physics methods have been successfully applied to the prediction of magnetism of small molecules (Li et al., 2019). However, the Fock space increases exponentially with the system size, which makes it impractical to be extended to sizable

[1]Quantum Measurement Group, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[5]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[6]Department of Applied Physics, Harvard University, Cambridge, MA 02138, USA

[7]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[8]Department of Physics, Clemson University, Clemson, SC 29634, USA

[9]Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[10]These authors contributed equally

[11]Lead contact

*Correspondence: tsmidt@mit.edu (T.S.), mingda@mit.edu (M.L.)

https://doi.org/10.1016/j.isci.2022.105192

materials beyond nanoscale without any approximation. The first-principles DFT simulations and the associated corrections provide an efficient compromise between accuracy and scalability. Although the delocalization error and the lack of static correlation may underestimate the magnetic moment and correlations (Huang et al., 2016; Liu et al., 2019), DFT-based methods have enabled high-throughput calculations over ~10,000 materials (Jain et al., 2013), allowing for preliminary statistical predictions of materials' properties. Even with substantial acceleration compared to experiments and wavefunction-based methods, the computational complexity of DFT calculations is still non-negligible and hinders the discovery in a huge, possibly infinitely large, parameter space of chemical compositions. Moreover, since electronic structure theory evaluates the energy for a specific electronic configuration, including the magnetic structure, the standard simulation requires traversing all configurations for a single atomic structure and determining the ground-state magnetic configuration. Thus, the large number of possible magnetic configuration forms a "guessing-computing" duo, that is, guessing many possible configurations and computing them one by one. Consequently, most computational efforts are spent on irrelevant magnetic excited states rather than the true ground states. A reliable prediction of the ground-state magnetic structure would greatly accelerate high-throughput calculation and bring us closer to achieving simulation-free materials discovery.

Given the challenges in magnetic structure determination from experiments and calculations, significant research effort has recently been dedicated to using machine learning to enhance magnetic structure determination. Some recent studies combine DFT calculations with machine learning (Rhone et al., 2020; Zheng and Zhang, 2021; Katsikas et al., 2021; Frey et al., 2020; Miyazato et al., 2018), in some of which the "guessing" step in the guessing-calculating procedure is obtained with machine learned models. For instance, machine learning has been implemented to reduce the search space of possible magnetic configurations in the "guessing" step. With this approach, the main calculation task is still carried out by the standard first-principles DFT calculations. Some other works are based on model Hamiltonians (Wang et al., 2020; Samarakoon et al., 2020), mostly classical spin models, and use machine learning methods to fit the free parameters in such a model, such as from experimental data that contain the spin information. Even so, the direct prediction of magnetic structure from the more direct atomic structure, aka replacing the "computing" step, is still challenging.

A full description of magnetism (Rodríguez-Carvajal and Villain, 2019) can be nontrivial. In this work, we focus on two different descriptions that use relatively few variables: magnetic order labels and propagation vectors. Magnetic ordering labels (e.g. ferromagnetic (FM) and antiferromagnetic (AFM)) are helpful because they summarize the complexity of magnetic structures into simple classes that are application relevant. Both ferromagnetic and ferrimagnetic materials exhibit a spontaneous magnetization: a non-zero net magnetic moment in the absence of an external magnetic field, while in FM all magnetic dipoles point in the same direction, and in FiM some of the dipoles point in the opposite direction. The antiferromagnetic materials have dipoles that point in the opposite direction in a regular pattern, but they cancel each other and the net magnetic moment is zero. The magnetic dipoles in non-magnetic materials orientate irregularly without a pattern, and the net magnetic moment is zero. A propagation vector is a vector in reciprocal space that describes symmetry breaking due to the presence of a magnetic order (ibid.). A non-zero propagation vector is one indicator for a more involved magnetic structure beyond the FM, AFM, and NM ternary classification. While these descriptions are expressive, they are not comprehensive and we leave more complex descriptions of magnetic order to future work.

We build an ML-based classifier that predicts the magnetic order under a ternary classification (FM/FiM: ferromagnetic/ferrimagnetic, AFM: antiferromagnetic, NM: non-magnetic), and also outputs whether the propagation vector is zero or non-zero. We choose these two in our prediction as a start; in doing so, we acquire partial but valuable information that may accelerate full magnetic structure determination. Some prior efforts of using machine learning to predict magnetic properties, like (Rhone et al., 2020; Miyazato et al., 2018), are restricted to a specific crystal structure. Only the types and ratios of elements to fill in the fixed crystal structure can be changed, chemical formula alone is enough to serve as inputs. There are also studies like (Himanen et al., 2020; Pham et al., 2018), where matrix descriptors are used to encode Coulomb interactions or relative distance between atoms. These models can be generalized to different crystal structures; however, due to the limitation of the matrix form, it becomes harder to encode all structure information for materials with many atoms in each unit cell and longer interaction range. Our prediction model applies to general cases, the inputs are crystal structures of any space group and

some atomic properties; our model can directly output the magnetic order and propagation vector given the inputs encoded as described in section Model architecture. This is realized by applying the Euclidean neural network (E(3)NN) (Geiger et al., 2021), the graph structures enable us to encode crystal structures as inputs; furthermore, E(3)NN preserves 3D rotation, inversion, and translation symmetry in the atomic structures, so that high accuracy is reached without data augmentation (ibid.). The accuracy in test set is about 78% for magnetic order prediction and about 74% for propagation vector prediction. The accuracy varies when the material of interest contains different elements or belongs to different space groups, as we will elaborate in the results section. Since the magnetic order and propagation vector represent two different pieces of magnetic structure information, they are trained separately with different neural network architectures and training data from the Materials Project (Jain et al., 2013) and MAGNDATA (Gallego et al., 2016a, 2016b) databases, respectively. The atomic structure inputs for magnetic order classifiers are from Materials Project, magnetic order labels are obtained using pymatgen's magnetism analyzer given structures with atoms decorated with DFT-calculated magnetic moments as inputs. Both structures and propagation vectors for propagation vector classifiers are from MAGNDATA.

METHODOLOGY

Data assembly

In order to train the magnetism classifier, we assemble a dataset containing both structure and magnetic order information from the Materials Project (Jain et al., 2013). We query structures that contain at least one magnetic elements, including transition metals (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Y, Nb, Mo, Ru, Rh, Re, Os, Ir, and Pt), lanthanides (Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, and Yb), and/or actinides (Th, U, Np, and Pu); however, a material containing a magnetic element does not necessarily host magnetic order. Our search is restricted to calculations using the generalized gradient approximation with Hubbard interaction (GGA + U) (Jain et al., 2013, 2018; Wang and Navrotsky, 2004), which is suitable for magnetic structure calculations and leads to a total of 34,856 structures (Due to recent updates in the Materials Project, if querying in latest pymatgen version with the same magnetic elements, the number of total structures may change. Our models are trained and tested based on pymatgen version 2022.0.8, with 34,856 total structures after querying), among which 30,584 are FiM/FM, 1,790 are AFM, and 2,482 are NM. The magnetic order labels are derived using the Python-based pymatgen analysis code (Ong et al., 2013) and through use of a CollinearMagneticStructureAnalyzer class. The CollinearMagneticStructureAnalyzer class uses the DFT-calculated magnetic moments and total magnetization (the absolute value of the sum of individual magnetic moments) to assign one of the following labels: FM if the total magnetization is greater than zero and all magnetic moments have the same sign, FiM if the total magnetization is greater than zero, AFM if total magnetic moment is zero and max absolute magnetic moment greater than zero, or NM magnetic order if total and max magnetic moment is zero. We then train 20 classifiers with different initial weights, each optimized on a randomly selected subset of 6,086 structures with a AFM:FM:NM ratio of 5:6:6. The size of each class is kept comparable in each selected subset to mitigate the training bias toward non-magnetic examples, since there is a substantially larger fraction of FM/FiM structures in the total dataset and only 1,790 AFM examples. Each subset of 6,086 materials is divided among training, validation, and test sets with a ratio of 0.8:0.1:0.1. Note that we find this further improves our model's ability to differentiate between AFM and FM/FiM classes, and the smaller data size helps reduce the overall training time.

Magnetic materials may host more complex magnetic structures beyond AFM/FiM which are described by other non-zero magnetic propagation vectors. In order to capture this complexity, we further train a binary classifier of the propagation vector magnitude (zero or non-zero). The zero propagation vector represents the prototypical FM order, while the non-zero propagation vector represents AFM/FiM orders and beyond. We obtain the structure and magnetic propagation vector information from the MAGNDATA database (Gallego et al., 2016a, 2016b), which to date contains the comprehensive experimentally determined magnetic structures of 1,562 compounds. The data used in this study were restricted to commensurate magnetic structures, as incommensurate magnetic structures always have non-zero propagation vectors. This yielded 1,134 total structures, of which 552 (582) have zero (non-zero) propagation vectors. We again train 20 classifiers with different initial weights, independently dividing the 1,134 total structures at random among training, validation, and test sets with a ratio of 0.8:0.1:0.1 for each model. To link the propagation vector classification with that of magnetic order, we show the statistics of structures appearing in the propagation vector classification datasets in Figure 1. Some structures' magnetic orders are unknown to

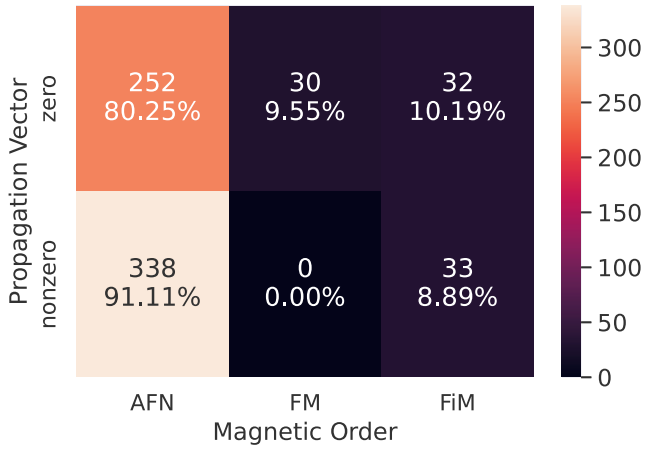


Figure 1. Number of examples with zero and non-zero propagation vector in each magnetic ordering class

pymatgen's magnetism analyzer which we use for determining magnetic order, so those are not included in the figure. Most structures with non-zero propagation vector are AFM, some are FiM, and none are FM or NM, as expected. The propagation vector classifier can further divide AFM and FiM classes into sub-classes with zero/non-zero propagation vector, giving us more information about the magnetic structure of a material. The hyperparameters of the model are extensively optimized by random search, and the initial set of hyperparameters is close to those presented in a previous relevant study using E(3)NN for phonon DOS prediction (Chen et al., 2021).

Model architecture

The architecture of both classifiers is based on Euclidean Neural Networks (E(3)NNs) (Geiger et al., 2021), a class of 3D Euclidean group (E(3))-equivariant neural networks. Any space group that describes the crystal geometric symmetries in three dimensions is a subgroup of E(3), and thus E(3)NNs preserve all geometric symmetries of the crystal structure, which removes the need for extensive data augmentation needed to consider arbitrary translations or rotations of the input structures. The neural network inputs consist of a material's crystal structure and one or more descriptors of each constituent atom. Specifically, the unit cell of each example is first converted into a periodic graph, where each node a represents an atom described by a feature vector x_a. A single convolutional layer operates on the input x_a and the radial distance vectors $\vec{r}_{ab}$ between atoms a and b in the neighborhood of a up to a radial cutoff $r_{max} = 5$Å, as shown in Figure 2A. The feature vector x_a associated with each node is constructed by a property-weighted one-hot encoding of selected atomic properties, illustrated in Figure 2B. For the propagation vector classifier, each feature vector is an array of 125 scalars, with the Z-th scalar being the atomic mass in amu (atomic mass units), where Z denotes the atomic number. Additional properties can be considered by simple concatenation of several such feature vectors, each weighted by the appropriate value of the property of interest. For example, each of the input feature vectors used for the magnetism classifier is an array of 3× 118 scalars, formed by concatenating three arrays of 118 scalars that encode the atomic radius (pm), electronegativity on a Pauling scale, and dipole polarizability (a.u.). The motivation for choosing these descriptors is they are related to the screened Coulomb interactions between atoms, and the magnetic properties can be solved in a lattice model if given the atomic structures and interactions between atoms. The input feature vector used for the propagation vector classifier only encodes the atomic mass, because given the very limited number of training data for this task, a simpler mapping that takes low-dimensional feature vectors as inputs would likely to have a less complicated loss landscape and be more favorable for training. Among the feature vectors that encode only one property, the ones with atomic mass have better performance compared to unit-weighted one-hot encoding and other properties we used for the magnetic order classifier.

The architecture of both classifiers consists of three principal parts, as shown in Figure 3. First, the input feature vectors x_a are passed to an embedding layer and mapped to a lower dimensional feature vector f_a. The E(3)NN layers are then applied to the resulting hidden feature f_a and consist of alternating convolution and gated block operations (dashed rounded rectangle). The convolution signifies the tensor

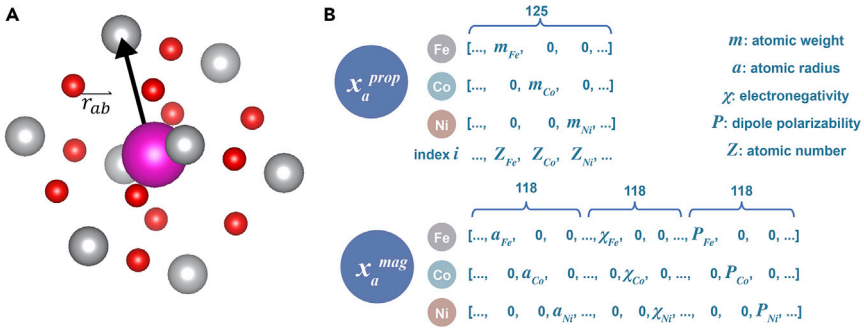


Figure 2. Illustration of input data structures

(A) A representative periodic graph constructed from the crystal structure in the neighborhood of a given atom. Each atom (node) carries a feature vector x_a, and each edge connecting node a to a neighboring node b is characterized by the relative distance vector $\vec{r}_{ab}$.

(B) Each atom is represented by a node in the periodic graph, and the atom type is expressed by a property-weighted one-hot feature vector. The top row shows a set of representative feature vectors used for the propagation vector classifier. Each is an array of 125 scalars, with the Z-th scalar being the atomic mass in amu (atomic mass unit), where Z denotes the atomic number. The bottom row shows a set of representative feature vectors used for the magnetism classifier. Each is an array of 3 × 118 scalars, formed by concatenating three arrays of 118 scalars which encode the atomic radius(pm), electronegativity on a Pauling scale, and dipole polarizability(a.u.) of a given atom employing the same property-weighted one-hot encoding scheme used for the propagation vector classifier input.

product between input feature vectors and symmetry-constrained convolutional kernels. The convolution step is implemented as:

$$f_a' = \frac{1}{\sqrt{z}} \sum_{b:\left|\vec{r}_{ab}\right|<r_{\max}} f_b \otimes \left(h(\| \vec{r}_{ab} \|)\right) Y\left(\vec{r}_{ab} / \| \vec{r}_{ab} \|\right). \qquad \text{(Equation 1)}$$

where the node f_a' is the output feature for the atom a. Each summand in (1) is a tensor product between the spherical harmonics $Y(\vec{r}_{ab} / \| \vec{r}_{ab} \|)$ and the feature f_b of a neighboring atom b, with $\vec{r}_{ab}$ being the positional vector pointing from atom a to b. Every tensor product could have multiple allowed paths(For example, two vector-like features $\vec{u}$ and $\vec{v}$ are able to generate a scalar w by inner product, a vector $\vec{w}$ by cross product, and a matrix W by outer product, thus leading to an output feature as the $u_1 w \oplus u_2 \vec{w} \oplus u_3 W$, where$u_i$'s denotes scalar weights that one can optionally adopt), where each path is multiplied by a scalar weight given by the output of the neural network $h(\| \vec{r}_{ab} \|)$. Intuitively, the spherical harmonics $Y(\vec{r}_{ab} / \| \vec{r}_{ab} \|)$ and neural network $h(\| \vec{r}_{ab} \|)$ take into account of direction and distance information of $\vec{r}_{ab}$, respectively. In addition, the z denotes the average number of neighboring atoms based on training dataset; thereby, the prefactor $1/\sqrt{z}$ can properly normalize the summed tensor products to increase numerical stability after several such layers.

The gated block step denotes a rotation-equivariant nonlinear activation function, which takes a set of input irreducible representations, x_i (scalar-like features) and y_i (higher degree features such as vectors and matrices), and outputs a set of activated features:

$$(\oplus_i \varphi_i(x_i)) \oplus \left(\oplus_j \varphi_j(g_j) y_j\right), \qquad \text{(Equation 2)}$$

where the symbol $\oplus$ denotes the direct sum(Intuitively, this means components of different features are "concatenated" instead of summed (with the $\sum$ sign), i.e., $w \oplus \vec{w}$ can be stored into a single array $[w, w_1, w_2, w_3]$. Thew_i's are components of $\vec{w}$, namely $[\vec{w}] = [w_1, w_2, w_3]$). More specifically, the scalar features x_i are directly passed into activation functions φ_i, while higher degree ones y_j are activated indirectly by passing additional learnable scalar features g_j to activation functions φ_j, which then serve as multiplying pre-factors of y_j.

Finally, the E(3)NN output is converted to a class label by first adding together the output vectors for all atoms in a given material, and then applying a final nonlinear activation. For the propagation vector classifier, the output is a sigmoid-activated scalar, while for the magnetism classifier, it is a softmax-activated array of three scalars giving the probability of exhibiting one of three magnetic orders (AFM, FM/FiM, and NM). Intuitively, the model is appropriate for magnetic properties prediction for the analogy between E(3)

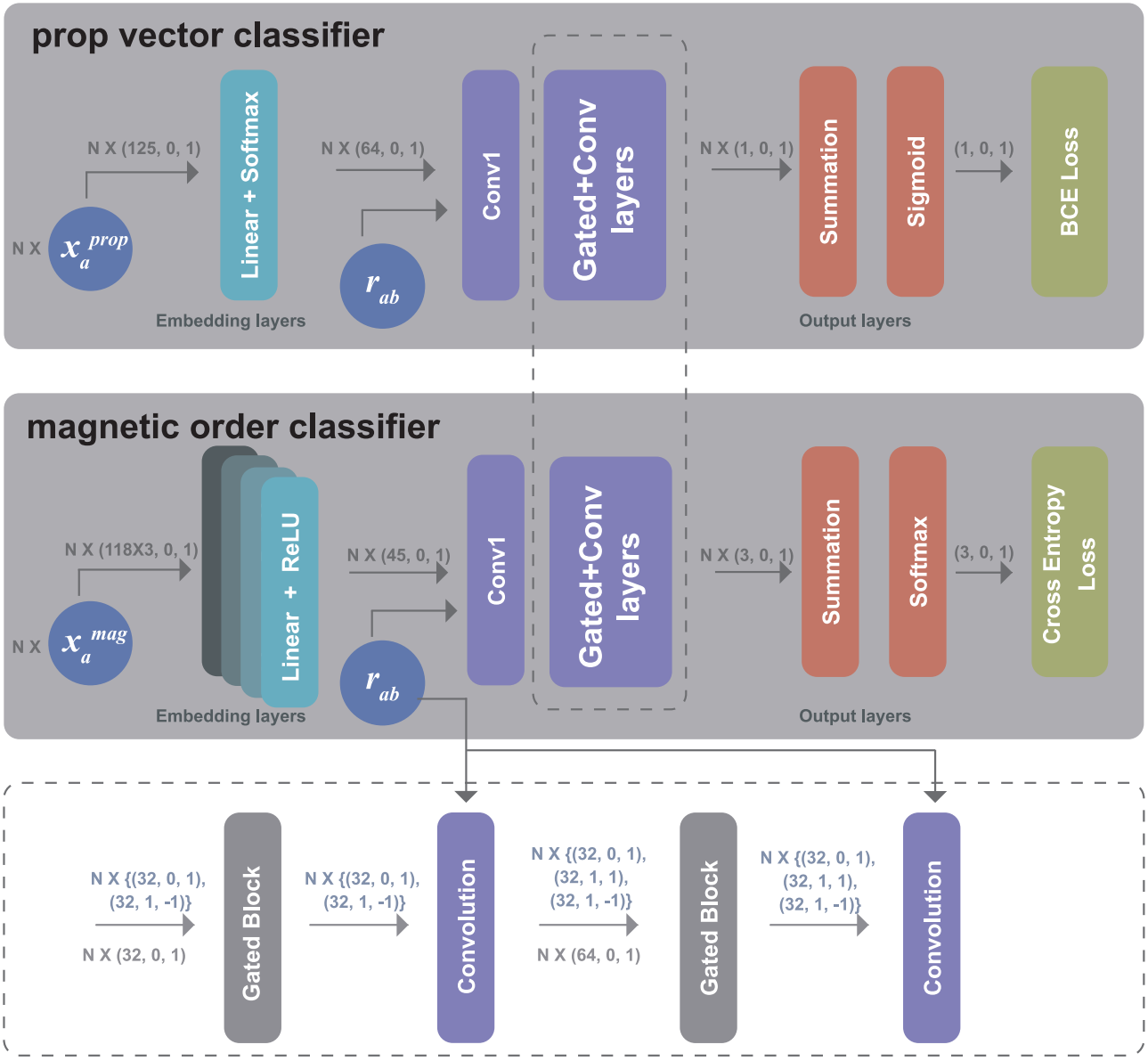


Figure 3. Illustration of the neural network architectures for propagation vector (top block) and magnetic order classification (middle block)

The models each consist of three principal layers: embedding layers, convolution and gated layers based on E(3)NN, and the output layers which generate the predicted classes. The convolution and gated layers of both models share the same architecture (detailed architecture shown in the bottom block). Even though the loss functions and embedding layers are optimized separately, by adopting this approach, one atomic structure can lead to a simultaneous prediction of both magnetic order and propagation vector.

NN and a lattice model for magnetic materials. In a lattice model, the magnetic configurations with the lowest energy can be derived given the screened Coulomb interactions between atoms(lattice sites). Equivalently in E(3)NN, it has the same graph structure as a lattice model, and information that is related to screened Coulomb interactions, including atomic properties and distance between atoms is encoded. The step of adding together the outputs vectors and applying a final nonlinear activation is equivalent to deriving magnetic orders from the net magnetic moment among atoms.

RESULTS

To quantify the consistency of predictions made by the magnetism and propagation vector classifiers, we independently train 20 models for each task using randomly drawn data subsets as described in section

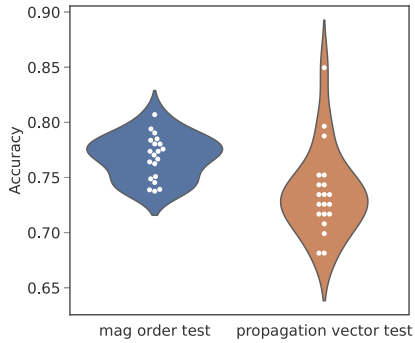


Figure 4. Prediction accuracies in test sets for magnetic orders (left) and propagation vectors (right) collected from 20 independently trained models

data assembly above. For the magnetic order classification, we show the test set accuracies of the 20 models in the left panel of Figure 4, which range from 73.8% to 80.7% and have a mean accuracy of 77.8%. For the propagation vector classification, the test accuracies obtained from 20 models are displayed in the right panel, ranging from 68.1% to 85.0% with a mean of 73.6%; the thresholds of this binary classification are chosen separately for each model to guarantee maximal accuracy. In Tables 1 and 2, we summarize the averaged precision, recall, and F1-scores over all trained models for the two tasks, respectively. We note that the larger spread of accuracies in propagation vector classification can possibly be attributed to fewer training examples (1,134 total structures taken 907 at a time for training), which are not sufficient for our model to learn all complex connections between crystal structures and propagation vectors. Besides, it may also suggest that the propagation vector contains rich strong correlation effect that cannot be fully characterized by only atomic structures. More details about the performance of both classifiers are apparent in the confusion matrices (CM) shown in Figure 5. For the magnetic order classifier, we observe excellent separation of the NM class from the two magnetic classes (Figure 5A). Since all calculations are performed at $T = 0K$, we believe that this non-magnetic separation is important in rapidly screening and excluding materials that do not host any magnetism, without having to experimentally cool down to the lowest measurable temperature. More ambiguities appear between classifications of the AFM and FM/FiM classes, where magnetic orders exist in both classes but in different formats. This is possibly due to the difficulty of distinguishing between FiM and AFM from atomic structures, and the energy difference between magnetic structures with AFM and FM/FiM orders can be small. The overall CM suggests good capability of recognizing potential magnetic orders but slightly weaker ability to identify the exact class. Figure 5B depicts the CM for the propagation vector classifier. Although the overall performance is hard to be considered satisfactory, as already mentioned above, the model has the better precision versus the recall for the non-zero propagation vector.

To analyze the performance of the magnetism classifier in more detail, we visualize the element-specific test set accuracies in Figure 6A. We observe highest classification accuracy on examples containing elements commonly found in ferromagnetic materials, such as Fe, Co, and Ni, with accuracies exceeding 76%. In addition, materials containing certain rare earth elements such as Tb, Dy, and Ho are classified with similar level of accuracy.

To further understand varying accuracies across different elements from the aspect of data abundances, we show the appearance frequency of each element inside the training set in Figure 6B. The correlations between high accuracies of some elements and large numbers of training samples containing those elements, including Mn, Fe, Co, Ni, and Cu, can be readily found. On the other hand, the elements with lower

Table 1. Averaged metrics of the 20 magnetic order classification models in test sets

class	Precision	recall	f1-score
NM	0.91	0.92	0.91
AFM	0.70	0.68	0.69
FM	0.68	0.70	0.69

Table 2. Averaged metrics of the 20 propagation vector classification models in test sets

class	Precision	recall	f1-score
Zero	0.70	0.83	0.76
Non-zero	0.79	0.64	0.71

prediction accuracies are typically less common, such as Ga, Lu, and Re. However, it is worth highlighting the great performance of some rare earth elements (e.g. Tb, Dy, and Ho) given the small amount of training samples, this is because they usually coexist with other abundant elements, for example, 65.9% of structures that contain Tb, Dy, and Ho also contain elements Mn, Fe, Mo, Co, and Ni.

Similar connections between the classification accuracy and the number of training samples can be made for different crystal systems. In Figure 7A, we show the number of examples of each magnetic order class as a function of the crystal system of the corresponding structure. Figure 7B indicates that higher appearance frequency in training data in general leads to higher classification accuracy in test data. Such relationships suggest that the predictions made by our model are based on not only the atomic species but also their coordinates and the crystal structure.

DISCUSSION

From the intuition of physics, E(3)NN is an appropriate model to predict magnetic properties for its ability to encode general atomic structures as input for a graph neural network without losing any spatial information and preserve the crystal symmetries. Early works like (Landrum and Genin, 2003) used decision trees to classify magnetic properties and atomic properties like valance electron orbits of the materials are taken as inputs. Many prior works like (Nelson and Sanvito, 2019; Bassman et al., 2018; Xie et al., 2021) only use chemical composition as inputs, the difference is their models only predict properties for two-dimensional materials with some specific crystal structures, while for most three-dimensional materials, there are much more space group and various structures. There is one exception (Nelson and Sanvito, 2019) that only encodes chemical composition but applies to general crystal structures in three dimension. The paper makes a discussion about adding crystal structures as inputs and trains comparison models containing only a limited description of the structural information of a compound. The efforts of encoding crystal structures do not improve performance significantly in their case, one reason as they discussed is feature vectors of crystal structures being too complicated compared to a limited training dataset. There are also some works (Lu et al., 2020; Pham et al., 2018; Himanen et al., 2020) that encode partial information of crystal structures with a simpler descriptor like the atomic adjacent matrix or the orbital field matrix. The matrix descriptors elegantly encode information related to the pairwise distances but not the relative angles and symmetry group between atoms. Among these efforts, E(3)NN can encode any crystal structures without losing the angle and symmetry information as input (e.g. input an *.cif file), thus our model is more general than most of the previous studies that can predict magnetic order for any crystal structures. The necessity

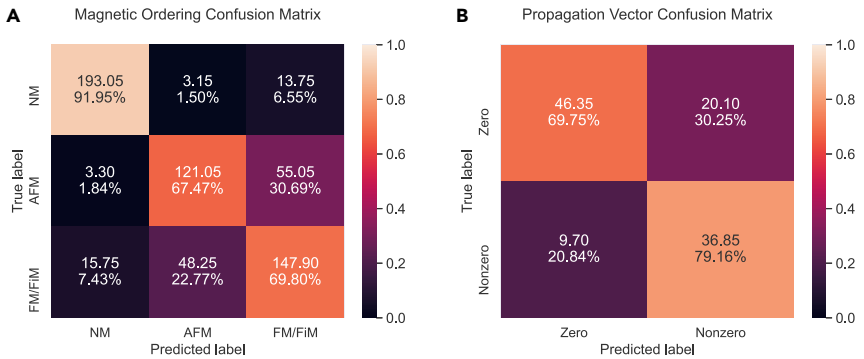


Figure 5. Confusion matrices for prediction results

(A) Magnetic orders confusion matrix (B) Propagation vectors confusion matrix. The color represents the percentage.

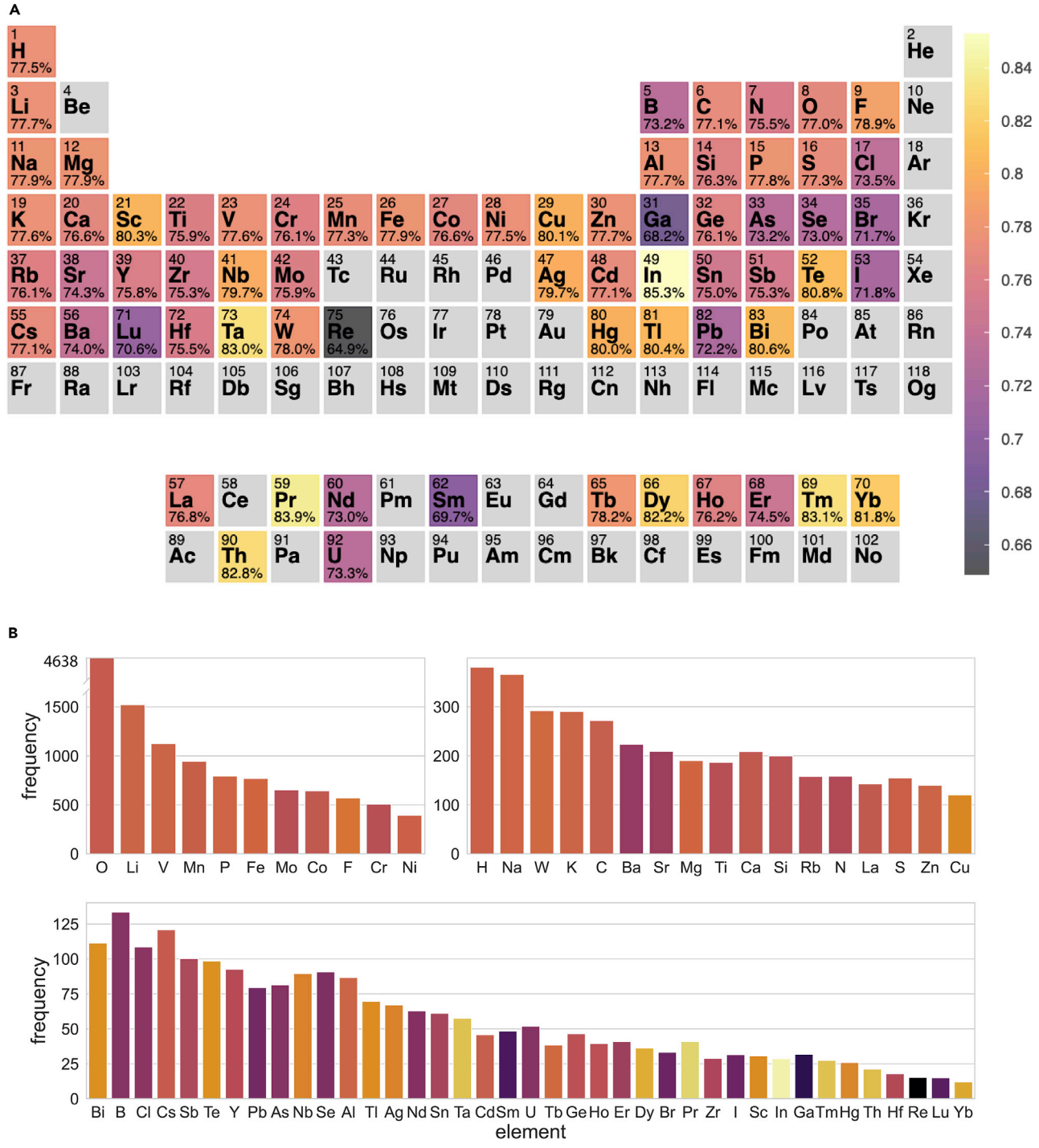


Figure 6. Performance on the testing set and sample statistics of the training set for the magnetic order predictions

(A) The testing accuracy of magnetic order classification per element shown in a color map on a periodic table, gray color indicates the element is absent or insufficient (mean frequency less than 1) in testing sets.

(B) Histogram of frequency each element appears in the training sets. The following elements are grayed out due to their absences or insufficiency in test compounds: Pd, Th, Re, Yb, Ce, Eu, Os, Ir, Pt, Rh, Ru, and Np.

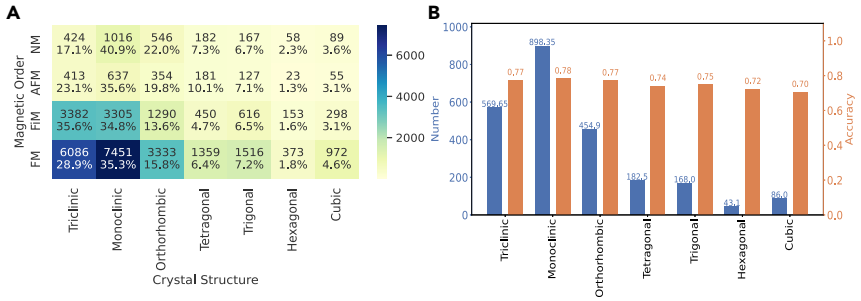


Figure 7. Relation between crystal system and prediction accuracy

(A) Number of examples in each magnetic order class as a function of the crystal system.

(B) Comparison between the number of training samples and testing classification accuracy for each crystal system, the values are averaged over 20 independent models.

to include atomic structures for prediction can be argued by the facts that some materials are described by the same chemical formulas but have different atomic structures, thus completely different magnetic structures. To further elaborate, it is hard to compare the performance of our model with some prior works that did not encode crystal structures or used other descriptors to encode, because the training data are different (most prior works focused on 2D materials (Nelson and Sanvito, 2019; Bassman et al., 2018; Xie et al., 2021)), and the prediction outputs are different(some prior works predict critical temperatures of ferromagnetic (Nelson and Sanvito, 2019) or superconducting (Bassman et al., 2018; Konno et al., 2021), or classify magnetic orders but only between two classes FM and AFM (Acosta et al., 2022)). In order to make a comparison and show the necessity to include atomic structures, we performed some comparable models, with only chemical composition as inputs.

The input is an array with 118 or 119 scalars; different models use different inputs between the two choices as shown in Table 3. The array of 118 scalars is a weighted average of one-hot encoding; the non-zero Z-th scalar is the ratio of the element with atomic number Z contained in the material. The array of 119 scalars has an additional last 1 scalar, which is the weighted average of atomic radius; the weight factors are ratios of each element. We randomly sample 6,147 structures with the ratios AFM:NM:FM/FiM = 1.2:1.0:1.2, and divide them into training:validation:testing = 0.8:0.1:0.1. We tried different neural networks including a simple linear model, two dense models, and one convolutional neural network(CNN) model; among all models, after optimizing the parameters, the best averaged accuracy we get is 69.0%, while the E(3)NN model that encodes crystal structures has an average accuracy of 77.8%. This may suggest crystal structures have a significant influence in predicting magnetic structures. Another work (Nelson and Sanvito, 2019) mentioned above also did a comparison between models with/without crystal structures, where encoding crystal structures does not improve accuracy significantly. The possible reasons are (1)our training dataset is large enough to support complicated feature vectors for encoding crystal structures, and (2)E(3)NN has a better performance than other descriptors of encoding crystal structures in (ibid.). In summary, crystal structures are appropriate inputs for magnetic structures prediction, other inputs are atomic properties related to screened Coulomb interactions, such as atomic radius, electronegativity, and dipole polarizability that are used in our model, as well as electron orbit configurations and pseudopotentials.

Table 3. The models with chemical composition and the averaged atomic property as inputs

Models	inputs	Structures	Activation	average accuracy
Linear model	(119,1)	One linear layer: (119, 3)	None	59.88%
Dense model 1	(119,1)	Linear layers: $(119, 64) \rightarrow (64, 32) \rightarrow (32, 3)$	LeakyReLU	53.0%
Dense model 2	(118,1)	Linear layers: $(118, 64) \rightarrow (64, 32) \rightarrow (32, 3)$	LeakyReLU	69.0%
CNN model	(118,1)	two convolutional layers and two linear layers	LeakyReLU	67.2%

In the CNN model, a max-pooling layer and a LeakyReLU activation follow after each convolution layer, then comes a flattened layer, 2 linear layers with dimensions $(116, 32) \rightarrow (32, 3)$. The parameters of two convolutional layers are (in channels, out channels, kernel size, padding)=(1,2,3,1), (2,4,3,1), the kernel size of max-pooling is 2.

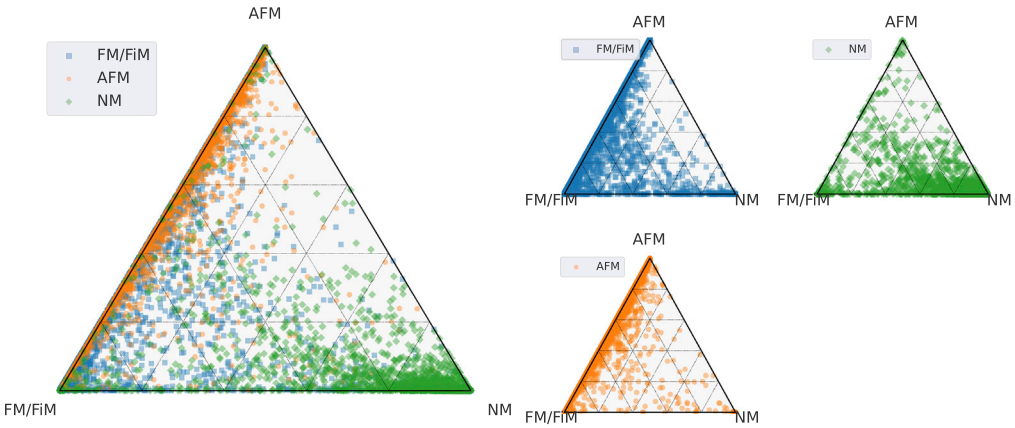


Figure 8. Outputs of the neural network with testing data as inputs shown in ternary plots, points represent 3-element vector outputs and are colored by their true label: FM/FiM(blue square), AFM(orange round), and NM(green diamond)

Overall, we present a machine-learning-based magnetism classifier that takes crystal structure as input and basic magnetic structure information as output. Despite extensive training, it seems that the prediction of magnetism hits a performance barrier, except for the non-magnetic class. This brings up the fundamental query on magnetic representation: besides the atomic species, the interatomic bonding, and the crystal symmetries, there might be some additional information still at large that can capture the essence of magnetism. We anticipate that finding proper magnetic representation can significantly boost the machine learning on magnetic materials research and shed light on strongly correlated electronic materials in general.

Limitations of the study

As mentioned above, our magnetic order classifiers show excellent separation between magnetic (including FM/FiM and AFM) and non-magnetic orders, but more ambiguities between classification of the AFM and FM/FiM classes. To discuss this further, particularly to see whether the FiM is preferably assigned to the FM or AFM class, Figures 8 and 9 show the 3-element vector outputs of the neural network before taking argmax function in a ternary form, with testing data as inputs. The closer a point is to a corner of the triangle, the more likely the predicted label is the corresponding class of that corner. We color the points with their true label to show distinction. Figure 8 explains that the testing outputs with true labels of AFM and FM/FiM distribute along the edge connecting AFM and FM/FiM classes and do not distinctly separate from each other, which causes the ambiguities in the classification. The data with true label of

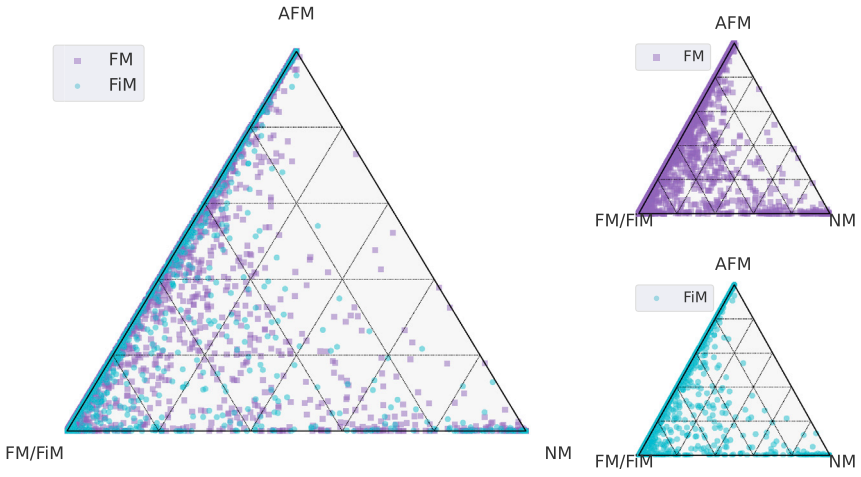


Figure 9. Outputs of the neural network with testing data as inputs shown in ternary plots, points represent 3-element vector outputs and are colored by their true label: FM(purple square) and FiM(blue diamond) data

Table 4. $m_\uparrow$ is the up total magnetization, we set the threshold to be the ratio between net magnetization and up magnetization

Threshold $\left(\left\lvert\frac{m_\uparrow - m_\downarrow}{m_\uparrow}\right\rvert\right)$	0%	1%	2%	3%	5%
Accuracy	0.76	0.69	0.65	0.68	0.67

All training is performed with a fixed ratio between the number of data in each label: NM:AFM:FM = 1:1:1.1(2537:2537:2791).

NM stay close to the NM corner and away from the other two classes. Figure 9 suggests that our choice to consider FM and FiM order jointly in the class FM/FiM is reasonable, since the distribution of FM and FiM in the ternary plots is similar, and our classifiers treat atomic structures with FM and FiM orders similarly.

Another factor that could cause ambiguity between AFM and FM/FiM classes is the magnetization cutoff between AFM and FiM classes. The magnetic order labels are assigned based on the DFT-calculated magnetic moments and total magnetization; while both FiM and AFM hold magnetic moments non-zero and with different signs, there is a cutoff of total magnetization to distinguish the two classes: zero for AFM, greater than zero for FiM. The models in the main text are trained with this cutoff strictly zero; however, there are materials in our dataset with small but non-zero magnetization. To compare how the change of this threshold affects accuracy, we changed the threshold for labeling FiM/AFM to be non-zero and re-trained our model; the prediction accuracy is shown in Table 4. When the threshold is set at zero, our model reaches maximal accuracy.

Another limitation of magnetic order prediction comes from Materials Project, first, it is a DFT database, so it has its own limitation; second, given the limited amount of AFM structures and our intention to keep a balanced ratio between different labels for training, the number of materials for each training is limited to around 4,869. By applying a Hubbard-like correction to localized d orbitals to correct self-interaction error, GGA + U has better performance (fits better with experiments) than generalized gradient approximation (GGA). There is currently no perfect DFT functional as they are all approximations to the complete set of physics that define materials phenomena, so does the GGA + U calculations. Our model is trained based on the DFT results from Materials Project; there is no guarantee that DFT results are the same as experiments. Admittedly, methods beyond DFT + U, such as QMC, DMRG, DMFT, or exact diagonalization, do exist; however, those methods have exceedingly high computational costs with at least N^6 to $exp(N)$ complexity, thereby still hard to compute the large volume of real materials. In light of this, DFT + U balances the accuracy with the computational cost, which can generate large training data to be used for machine learning. To show the effect of the number of training data on model performance, we trained our model with less training data. The average accuracy with the current amount of training data is 77.8%, if reducing the amount to $3/4, 1/2, 1/4$ of the current choice while keeping the ratios between labels the same, the average accuracy reduces to 70.8%, 66.7%, and 65.4%. It is likely that if we have more training data available, a higher accuracy could be reached.

For the propagation vectors, we ultimately aim to predict the number of propagation vectors and the magnitude and direction for each. We only achieve partially this by predicting whether the propagation vector is zero or non-zero. The challenge comes from the constitution of the MAGNDATA database, almost half of materials in MAGNDATA have a single zero(null) propagation vector. Among these with non-zero propagation vectors, the number of propagation vectors can be one or multiple, and each propagation vector can point at a high-symmetry point or just a general position in the Brillouin zone. The situations for non-zero propagation vector are complicated and there are not enough data to perform training and get an accurate prediction. Besides, we exclude incommensurable magnetic structures in training and prediction. The magnetic unit cells for incommensurable magnetic structures are infinitely large, and in E(3)NN graphs, each node represents an atom and is expected to carry information about magnetism on that atom after training, thus it is unlikely a finite graph model can characterize incommensurable structures with infinitely large unit cells.

Given the above limitation, we think our model will be improved in future studies if (1) beyond the properties shown in Figure 2B, a thorough search over more atomic properties to be encoded as inputs is performed, some candidates are electron configurations of atoms and pseudopotentials. (2) Instead of training the model with one output label like the magnetic order or the propagation vector, one can use the local magnetization of each atom as training labels and define a loss function to compare magnetization on every atom.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability

ACKNOWLEDGMENTS

H.A.M., H.H., and L.N. thank the support from MIT Experiential Learning Opportunities (ELO) Program for first-year students and MIT Undergraduate Research Opportunities Program (UROP) and the Baruch (1947) Fund. T.L. acknowledges the support from MathWorks Science Fellowship. Z.C., N.A., and M.L. acknowledge the support from U.S. DOE, Office of Science (SC), Basic Energy Sciences (BES), award No. DE-SC0021940 and DE-SC0020148. M.L. acknowledges the Norman C. Rasmussen Career Development Chair. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award BES-ERCAP0020159.

AUTHOR CONTRIBUTIONS

H.A.M., H.H., and L.N. built and trained the machine learning models and performed optimization with support of T.L. T.L. analyzed the results with support from H.A.M., H.H., and L.N. T.L. and other authors., Z.C. and N.A. wrote the manuscript with input from all authors. M.L. and T.S. supervised the research.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 7, 2022
Revised: August 24, 2022
Accepted: September 20, 2022
Published: October 21, 2022

REFERENCES

Acosta, C.M., Ogoshi, E., Souza, J.A., and Dalpian, G.M. (2022). Machine learning study of the magnetic ordering in 2D materials. ACS Appl. Mater. Interfaces *14*, 9418–9432. https://doi.org/10.1021/acsami.1c21558.

Ament, L.J.P., van Veenendaal, M., Devereaux, T.P., Hill, J.P., and van den Brink, J. (2011). Resonant inelastic x-ray scattering studies of elementary excitations. Rev. Mod. Phys. *83*, 705–767.

Artsimovich, L. (1972). Tokamak devices. Nucl. Fusion *12*, 215–252.

Balents, L. (2010). Spin liquids in frustrated magnets. Nature *464*, 199–208.

Banerjee, A., Yan, J., Knolle, J., Bridges, C.A., Stone, M.B., Lumsden, M.D., Mandrus, D.G., Tennant, D.A., Moessner, R., and Nagler, S.E. (2017). Neutron scattering in the proximate quantum spin liquid alpha-RuCl3. Science *356*, 1055–1059.

Bassman, L., Rajak, P., Kalia, R.K., Nakano, A., Sha, F., Sun, J., Singh, D.J., Aykol, M., Huck, P., Persson, K., and Vashishta, P. (2018). Active learning for accelerated design of layered materials. NPJ Comput. Mater. *4*, 74–79.

Binder, K., and Young, A.P. (1986). Spin glasses: experimental facts, theoretical concepts, and open questions. Rev. Mod. Phys. *58*, 801–976.

Chen, Z., Andrejevic, N., Smidt, T., Ding, Z., Xu, Q., Chi, Y.T., Nguyen, Q.T., Alatas, A., Kong, J., and Li, M. (2021). Direct prediction of phonon density of states with euclidean neural networks. Adv. Sci. *8*, 2004214.

Fert, A. (2008). Nobel Lecture: origin, development, and future of spintronics. Rev. Mod. Phys. *80*, 1517–1530.

Frey, N.C., Horton, M.K., Munro, J.M., Griffin, S.M., Persson, K.A., and Shenoy, V.B. (2020). High-throughput search for magnetic and topological order in transition metal oxides. Sci. Adv. *6*, eabd1076.

Gallego, S.V., Perez-Mato, J.M., Elcoro, L., Tasci, E.S., Hanson, R.M., Aroyo, M.I., and Madariaga, G. (2016a). MAGNDATA: towards a database of magnetic structures. II. The incommensurate case. J. Appl. Crystallogr. *49*, 1941–1956.

Gallego, S.V., Perez-Mato, J.M., Elcoro, L., Tasci, E.S., Hanson, R.M., Momma, K., Aroyo, M.I., and Madariaga, G. (2016b). MAGNDATA: towards a database of magnetic structures. I. The commensurate case. J. Appl. Crystallogr. *49*, 1750–1776.

Geiger, M., Smidt, T., Alby, M., Miller, B.K., Boomsma, W., Dice, B., Lapchevskyi, K., Weiler, M., Tyszkiewicz, M., Batzner, S., et al. (2021). e3nn/e3nn: 2021-08-27. Version 0.3.5.

Gershenfeld, N.A., and Chuang, I.L. (1997). Bulk spin-resonance quantum computation. Science *275*, 350–356.

Himanen, L., Jäger, M.O., Morooka, E.V., Federici Canova, F., Ranawat, Y.S., Gao, D.Z., Rinke, P., and Foster, A.S. (2020). DScribe: library of descriptors for machine learning in materials science. Comput. Phys. Commun. *247*, 106949.

Huang, W., Xing, D.-H., Lu, J.-B., Long, B., Schwarz, W.H.E., and Li, J. (2016). How much can density functional approximations (DFA) fail? The extreme case of the FeO4 species. J. Chem. Theory Comput. *12*, 1525–1533.

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). Commentary: the Materials Project: a materials genome approach to accelerating materials innovation. Apl. Mater. *1*, 011002.

Jain, A., Montoya, J., Dwaraknath, S., Zimmermann, N.E.R., Dagdelen, J., Horton, M., Huck, P., Winston, D., Cholia, S., Ong, S.P., and Persson, K. (2018). The materials Project: accelerating materials design through theory-driven data and tools. In Handbook of Materials Modeling : Methods: Theory and Modeling, W. Andreoni and S. Yip, eds. (Springer International Publishing), pp. 1–34.

Katsikas, G., Sarafidis, C., and Kioseoglou, J. (2021). Machine learning in magnetic materials. Phys. Status Solidi B *258*, 2000600.

Keimer, B., Kivelson, S.A., Norman, M.R., Uchida, S., and Zaanen, J. (2015). From quantum matter to hightemperature superconductivity in copper oxides. Nature *518*, 179–186.

Konno, T., Kurokawa, H., Nabeshima, F., Sakishita, Y., Ogawa, R., Hosako, I., and Maeda, A. (2021). Deep learning model for finding new superconductors. Phys. Rev. B *103*, 014509.

Landrum, G.A., and Genin, H. (2003). Application of machine-learning methods to solid-state chemistry: ferromagnetism in transition metal alloys. In Journal of Solid State Chemistry 176.2. Special issue on The Impact of Theoretical Methods on Solid-State Chemistry, pp. 587–593.

Li, Z., Guo, S., Sun, Q., and Chan, G.K.-L. (2019). Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations. Nat. Chem. *11*, 1026–1033.

Liu, F., Yang, T., Yang, J., Xu, E., Bajaj, A., and Kulik, H.J. (2019). Bridging the homogeneous-heterogeneous divide: modeling spin for reactivity in single atom catalysis. Front. Chem. *7*, 219.

Lovesey, S.W. (1986). Theory of Neutron Scattering from Condensed Matter: Volume I: Nuclear Scattering. International Series of Monographs on Physics (Oxford University Press).

Lu, S., Zhou, Q., Guo, Y., Zhang, Y., Wu, Y., and Wang, J. (2020). Coupling a crystal graph multilayer descriptor to active learning for rapid discovery of 2D ferromagnetic semiconductors/half-metals/metals. Adv. Mater. *32*, 2002658. https://doi.org/10.1002/adma.202002658.

Manchon, A., Železný, J., Miron, I., Jungwirth, T., Sinova, J., Thiaville, A., Garello, K., and Gambardella, P. (2019). Current-induced spin-orbit torques in ferromagnetic and antiferromagnetic systems. Rev. Mod. Phys. *91*, 035004.

Miyazato, I., Tanaka, Y., and Takahashi, K. (2018). Accelerating the discovery of hidden two-dimensional magnets using machine learning and first principle calculations. J. Phys. Condens. Matter *30*, 06LT01.

Mühlbauer, S., Binz, B., Jonietz, F., Pfleiderer, C., Rosch, A., Neubauer, A., Georgii, R., and Böni, P. (2009). Skyrmion lattice in a chiral magnet. Science *323*, 915–919.

Nelson, J., and Sanvito, S. (2019). Predicting the Curie temperature of ferromagnets using machine learning. Phys. Rev. Mater. *3*, 104405.

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. (2013). Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. Comput. Mater. Sci. *68*, 314–319.

Pham, T.-L., Nguyen, N.-D., Nguyen, V.-D., Kino, H., Miyake, T., and Dam, H.-C. (2018). Learning structure-property relationship in crystalline materials: a study of lanthanide-transition metal alloys. J. Chem. Phys. *148*, 204106. https://doi.org/10.1063/1.5021089.

Rhone, T.D., Chen, W., Desai, S., Torrisi, S.B., Larson, D.T., Yacoby, A., and Kaxiras, E. (2020). Data-driven studies of magnetic two-dimensional materials. Sci. Rep. *10*, 15795.

Rodríguez-Carvajala, J., and Villain, J. (2019). Magnetic structures. Compt. Rendus Phys. *11*, 770–802.

Samarakoon, A.M., Barros, K., Li, Y.W., Eisenbach, M., Zhang, Q., Ye, F., Sharma, V., Dun, Z., Zhou, H., Grigera, S.A., et al. (2020). Machine-learning-assisted insight into spin ice Dy2Ti2O7. Nat. Commun. *11*, 892.

Shull, C.G., and Smart, J.S. (1949). Detection of antiferromagnetism by neutron diffraction. Phys. Rev. *76*, 1256–1257.

Vlaardingerbroek, M.T., and Boer, J.A. (2003). Magnetic Resonance Imaging: Theory and Practice (Springer Science & Business Media).

Wang, D., Wei, S., Yuan, A., Tian, F., Cao, K., Zhao, Q., Zhang, Y., Zhou, C., Song, X., Xue, D., and Yang, S. (2020). Machine learning magnetic parameters from spin configurations. Adv. Sci. *7*, 2000566.

Wang, M., and Navrotsky, A. (2004). Enthalpy of formation of LiNiO2, LiCoO2 and their solid solution, LiNi1- xCoxO2. Solid State Ion. *166*, 167–173.

Wiedemann, H. (2015). Particle Accelerator Physics (Springer Nature).

Xie, Y., Tritsaris, G.A., Grånäs, O., and Rhone, T.D. (2021). Data-Driven studies of the magnetic anisotropy of two-dimensional magnetic materials. J. Phys. Chem. Lett. *12*, 12048–12054. https://doi.org/10.1021/acs.jpclett.1c03783.

Zheng, F., and Zhang, P. (2021). MagGene: a genetic evolution program for magnetic structure prediction. Comput. Phys. Commun. *259*, 107659.

Zhou, Y., Kanoda, K., and Ng, T.-K. (2017). Quantum spin liquid states. Rev. Mod. Phys. *89*, 025003.

STAR★METHODS

KEY RESOURCES TABLE

.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
The Materials Project	Jain et al. (2013)	https://doi.org/10.1063/1.4812323
MAGNDATA	Bilbao Crystallographic Server	https://doi.org/10.1107/S1600576716012863
Software and algorithms		
e3nn version 0.0.0	Geiger et al., 2021	https://e3nn.org
pymatgen version 2022.0.8	Ong et al. (2013)	https://doi.org/10.1016/j.commatsci.2012.10.028
Magnetism Classifiers from Atomic Coordinates	this paper; Zenodo	https://doi.org/10.5281/zenodo.7005888

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Prof. Mingda Li (mingda@mit.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- This paper analyzes existing, publicly available data. The accession numbers for these datasets are listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.