Machine Learning Prediction of Magnetic Moments in Transition-Metal Compounds: A Data-Driven Screening Framework for Altermagnetic Candidate Materials

Emon Sarder*[1]

[1]Department of Physics, Pabna University of Science and Technology (PUST), Pabna, Bangladesh

Abstract

Altermagnetism – a collinear magnetic phase combining zero net magnetization with momentum-space spin-splitting – has emerged as a major target for spintronic materials discovery, yet identifying compensated magnetic candidates within large computational databases remains a bottleneck. We present a machine learning framework that predicts the total magnetic moment per formula unit of transition-metal-bearing binary and ternary compounds using structural, thermodynamic, and composition-based descriptors retrieved from the Materials Project database. A dataset of 29,584 compounds containing one or more of nine magnetic transition metals (Mn, Fe, Cr, Co, Ni, Ru, V, Rh, Os) was assembled and expanded to 142 features per compound using Magpie composition descriptors together with structural and thermodynamic properties. Three regression models – a linear (Ridge) baseline, Random Forest, and histogram-based Gradient Boosting – were compared using 5-fold cross-validation and an independent held-out test set. Gradient Boosting achieved the best performance (test $R^2 = 0.778$, MAE $= 1.99\ \mu_B$/f.u., RMSE $= 4.65\ \mu_B$/f.u.), substantially outperforming the linear baseline ($R^2 = 0.466$) and confirming that magnetic moment prediction benefits from non-linear, ensemble-based methods. Permutation importance analysis identifies the number of crystallographically magnetic sites and the mean ground-state magnetic moment of constituent elements as the dominant predictors, with electronic (unfilled d-shell), thermodynamic, and symmetry-related descriptors playing secondary roles. We discuss how such a regression model can support first-pass screening for magnetically compensated, antiferromagnetically-labeled compounds – a necessary precondition for altermagnetic candidacy – while explicitly addressing the model's limitations, including a known ferromagnetic bias in high-throughput collinear DFT databases, a heavily right-skewed target distribution, and the partly definitional role of site-counting features. This work is offered as a screening tool that narrows candidate space for more targeted symmetry-resolved analysis, not as a direct predictor of altermagnetic order itself.

Keywords: Altermagnetism; Machine Learning; Magnetic Moment Prediction; Materials Project; Random Forest; Gradient Boosting; Spintronics; Materials Informatics

1 Introduction

Collinear magnetic order in crystalline solids has traditionally been divided into two classes: ferromagnetism, in which the spin sublattices are inequivalent and produce a macroscopic moment, and antiferromagnetism, in which sublattices related by translation or inversion symmetry cancel to give zero net moment.

*Corresponding author.

Over the past several years this classification has been shown to be incomplete. *Altermagnetism* has been established as a third, symmetry-distinct collinear phase in which opposite-spin sublattices are related by a rotational rather than translational or inversion symmetry, so that the electronic bands acquire an alternating, momentum-dependent spin-splitting even though the crystal carries zero net magnetization [1, 2, 3]. This combination – antiferromagnet-like absence of stray fields together with ferromagnet-like spin-dependent transport – has generated substantial interest for antiferromagnetic spintronics, where it enables spin-splitter effects, large anomalous Hall responses, and giant magnetoresistance without a net moment [9, 8, 10, 11]. Subsequent reviews have surveyed the rapidly expanding list of candidate materials and proposed device architectures [5, 6, 7, 4].

Experimental confirmation has followed quickly. Ruthenium dioxide (RuO_2) was among the first materials in which an anomalous Hall response consistent with altermagnetic spin-splitting was reported despite vanishing net moment [12]. Angle-resolved photoemission spectroscopy (ARPES) has since directly resolved altermagnetic band splittings of order 1 eV in metallic CrSb [13, 14, 15, 16], and Kramers-degeneracy-breaking splittings have been observed in MnTe [17, 18]. In parallel, first-principles work has proposed that chemical doping of the correlated narrow-gap semiconductor $FeSb_2$ – already notable for its colossal thermoelectric response [19, 20, 21] – can stabilize an altermagnetic-type collinear order [22], illustrating that candidate altermagnets are not confined to a single structural family but can emerge across diverse transition-metal chemistries.

A practical consequence of this diversity is that discovering new altermagnetic candidates increasingly resembles a materials-screening problem: among the tens of thousands of transition-metal compounds catalogued in high-throughput density functional theory (DFT) databases, which ones are magnetically *compensated* – carrying negligible net moment despite being locally magnetic – and therefore worth subjecting to the symmetry-resolved (spin space group) analysis needed to confirm altermagnetic order? Directly computing spin-resolved band structures for every candidate is computationally prohibitive at database scale. Machine learning (ML) trained on inexpensive, already-tabulated structural and compositional descriptors offers a way to triage this space, in the same spirit as ML-accelerated screening efforts for band gaps, formation energies, and other DFT-derived properties [27, 26].

Predicting magnetic moment from such descriptors is nonetheless known to be a comparatively difficult regression target. Kaba *et al.* showed, using both graph neural networks and random forests trained on the Materials Project database, that magnetic properties are intrinsically harder to predict than other DFT-computed quantities such as formation energy or band gap [34]. Several groups have nonetheless demonstrated useful accuracy on more restricted chemistries: Verma *et al.* developed a structure-based descriptor achieving high accuracy for magnetic ordering and magnetic moment per atom on a curated set of stable ferromagnetic and ferrimagnetic binary/ternary compounds [37]; Pham *et al.* showed that an orbital-field-matrix representation predicts local magnetic moments in lanthanide-transition-metal alloys [38]; and Sanspeur and Kitchin specifically addressed the class imbalance that arises when most compounds in a magnetic database have small or zero moment [39] – an issue that, as we show below, is directly relevant to the present dataset. At the classification level, Frey, Horton *et al.* trained ML classifiers to predict magnetic ground states across transition-metal oxides in the Materials Project [36], while Fahmy *et al.* more recently demonstrated that Materials Project's collinear DFT workflows carry a systematic bias toward ferromagnetic convergence relative to experimentally validated magnetic structures [35] – a caveat we return to in our discussion of limitations.

In this work we build a regression framework that predicts the net magnetic moment per formula unit for a broad, multi-element set of transition-metal compounds (rather than a single family), using structural, thermodynamic, and Magpie composition-based descriptors [26, 25]. Our goal is not to predict altermagnetic order directly – that requires spin-resolved and symmetry-resolved analysis beyond the scope of a scalar regression target – but to provide a fast, physically grounded first-pass screen: compounds whose predicted and database-reported moments are near zero, yet whose underlying magnetic-site count and elemental mag-

netism suggest locally ordered spins, are natural candidates for closer inspection as potential compensated (and possibly altermagnetic) systems. We report cross-validated and held-out test accuracy for three model classes, use permutation importance to identify which descriptors actually drive the model's predictions, and discuss both the physical plausibility and the limitations of the resulting picture in some detail, since an honest accounting of where such a screening tool is and is not reliable is essential to its practical use.

2 Materials and Methods

2.1 Data Collection

Compound-level data were retrieved programmatically from the Materials Project database via its Python API (mp-api) [23], restricting the query to binary and ternary compounds containing at least one of nine transition metals commonly associated with magnetism and with proposed or confirmed altermagnetic candidacy: Mn, Fe, Cr, Co, Ni, Ru, V, Rh, and Os. For each compound we retrieved structural descriptors (unit-cell volume, density, atomic density), thermodynamic descriptors (formation energy per atom, energy above the convex hull), electronic descriptors (band gap), symmetry information (space group number and symbol, crystal system), and magnetic descriptors (magnetic ordering classification, total magnetization, and the number of magnetic and symmetry-unique magnetic sites) as reported by Materials Project's collinear spin-polarized DFT workflow. After removing duplicate entries and compounds with missing thermodynamic or magnetic data, the working dataset comprised 29,584 compounds.

2.2 Feature Engineering

Each compound's chemical formula was parsed into a `pymatgen` composition object [24], from which 132 Magpie composition-based descriptors were generated using the `matminer` toolkit's `ElementProperty` featurizer [26, 25]. These descriptors summarize elemental properties – including electronegativity, atomic and covalent radii, melting temperature, valence electron counts, unfilled orbital counts, Mendeleev number, and ground-state magnetic moment – as the mean, standard deviation, range, minimum, maximum, and mode across the elements present in each compound. Combined with the ten structural, thermodynamic, and symmetry descriptors described above (with crystal system one-hot encoded), the final feature matrix contained 142 numeric columns. Compounds missing more than 10% of feature values were dropped; remaining sparse values were imputed with the column median. The regression target was the total magnetization normalized per formula unit, as computed by Materials Project's DFT workflow.

2.3 Model Architecture and Training

Three regression models were compared: (i) a Ridge-regularized linear model as a baseline, (ii) a Random Forest regressor with 200 trees [28], and (iii) a histogram-based Gradient Boosting regressor with 200 boosting iterations, implemented via scikit-learn's `HistGradientBoostingRegressor` [29]. The histogram-based approach bins continuous features before searching for split points, following the strategy introduced for LightGBM [30] and generalizing the original gradient boosting formulation of Friedman [31]; this makes it substantially more scalable than conventional gradient boosting on datasets of the size used here ($\sim 3 \times 10^4$ compounds, 142 features) while retaining comparable predictive accuracy. All models were trained on an 80% training split (23,667 compounds) with the remaining 20% (5,917 compounds) held out for final testing, using a fixed random seed for reproducibility. Model selection and hyperparameter comparison were performed using 5-fold cross-validation on the training set only, with R^2, mean absolute error (MAE), and root-mean-square error (RMSE) as evaluation metrics; the held-out test set was used exactly once, after model selection, to report final performance.

2.4 Feature Importance

Rather than relying on the built-in impurity-based feature importance of tree ensembles, which is known to be biased toward high-cardinality and continuous features [28], we computed permutation importance on the held-out test set [32]: each feature is independently shuffled ten times, and the resulting mean drop in R^2 is taken as that feature's importance. This model-agnostic approach measures the actual contribution of a feature to predictive performance rather than a training-time heuristic, and is applicable regardless of the underlying model class [33].

3 Results

3.1 Target Distribution

Figure 1 shows the distribution of the regression target across the full dataset. The distribution is heavily right-skewed: the large majority of compounds carry a small or vanishing net magnetic moment (0–5 μ_B per formula unit), consistent with the fact that most transition-metal compounds in this composition space are non-magnetic, weakly magnetic, or antiferromagnetically compensated, while a comparatively small number of strongly ferromagnetic compounds extend the tail out to moments exceeding 100–200 μ_B per formula unit. This skew has two consequences that we return to in the Discussion: it means that aggregate accuracy metrics such as R^2 are dominated by performance on the densely populated low-moment region, and it creates a class-imbalance-like challenge for the high-moment tail that mirrors observations made elsewhere in the magnetic-moment ML literature [39].

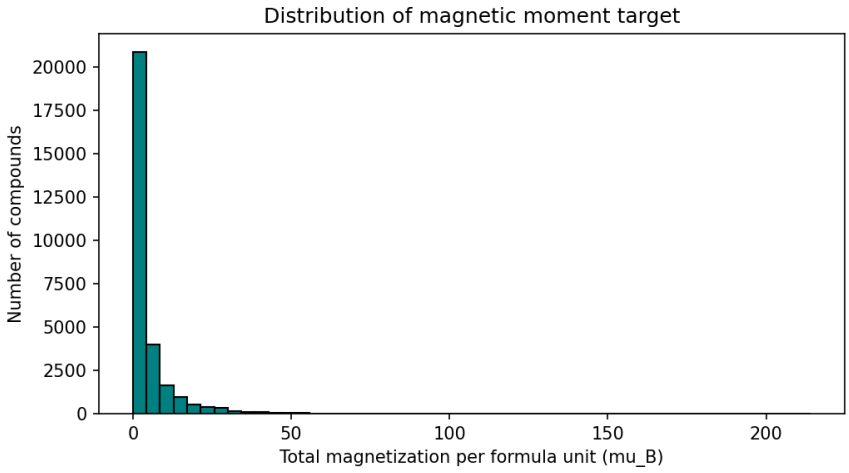


Figure 1: Distribution of the total magnetic moment per formula unit across the 29,584-compound dataset. The distribution is strongly right-skewed, with the majority of compounds carrying a small net moment and a long tail of strongly ferromagnetic compounds.

3.2 Model Performance

Table 1 reports 5-fold cross-validated performance on the training set. The Ridge baseline achieves a modest R^2 of 0.448, while both ensemble methods substantially outperform it: Random Forest reaches a mean R^2 of 0.730, and Gradient Boosting reaches 0.753 with a notably smaller fold-to-fold standard deviation (0.0035 versus 0.014 for Random Forest), indicating more stable performance across different training subsets.

Table 1: 5-fold cross-validation results on the training set (23,667 compounds).

Model	CV R^2 (mean)	CV R^2 (std)	CV MAE (μ_B/f.u.)	CV RMSE (μ_B/f.u.)
Ridge (linear baseline)	0.448	0.013	4.00	7.50
Random Forest	0.730	0.014	1.99	5.26
Gradient Boosting	0.753	0.004	2.05	5.02

Table 2 reports performance on the held-out test set (5,917 compounds), evaluated once after model selection. The ranking of models is consistent with the cross-validation results, with Gradient Boosting achieving the best test R^2 of 0.778, closely followed by Random Forest at 0.768; both are dramatically better than the linear baseline ($R^2 = 0.466$). Gradient Boosting is therefore taken as the primary model for the feature importance analysis and physical discussion that follow.

Table 2: Performance on the held-out test set (5,917 compounds), evaluated once.

Model	Test R^2	Test MAE (μ_B/f.u.)	Test RMSE (μ_B/f.u.)
Ridge (linear baseline)	0.466	3.88	7.22
Random Forest	0.768	1.86	4.75
Gradient Boosting	**0.778**	1.99	4.65

Figure 2 shows the parity plot for the best-performing model. Agreement with the ideal $y = x$ line is close for the densely populated low-moment region (0–50 μ_B/f.u.), which contains the overwhelming majority of compounds, confirming that the model reliably distinguishes weakly magnetic and compensated compounds from strongly magnetic ones. Scatter increases visibly for compounds with large actual moments (above roughly 75–100 μ_B/f.u.), where the model tends to underpredict; this is the expected signature of the right-skewed target distribution in Figure 1 – the training set simply contains far fewer examples in this regime, limiting the model's ability to learn their behavior precisely.

3.3 Feature Importance

Figure 3 shows the top 15 features by permutation importance for the Gradient Boosting model. The number of crystallographically distinct magnetic sites per formula unit (`num_magnetic_sites`) is overwhelmingly dominant, accounting for a mean R^2 drop of 0.459 when shuffled – roughly *four times* the contribution of any other single feature. The second most important descriptor is the Magpie-derived mean ground-state magnetic moment of the constituent elements (`MagpieData mean GSmagmom`, importance 0.124), followed by the number of symmetry-unique magnetic sites (0.052). The remaining top-15 features contribute more modestly (importance $\lesssim 0.03$ each) and span electronic descriptors related to d- and f-shell filling (average deviation in unfilled d-electron count, maximum unfilled f-electron count, mean unfilled d-electron count), crystal symmetry (space group number), thermodynamic stability (energy above the convex hull, formation energy per atom), and basic structural properties (unit-cell volume, density).

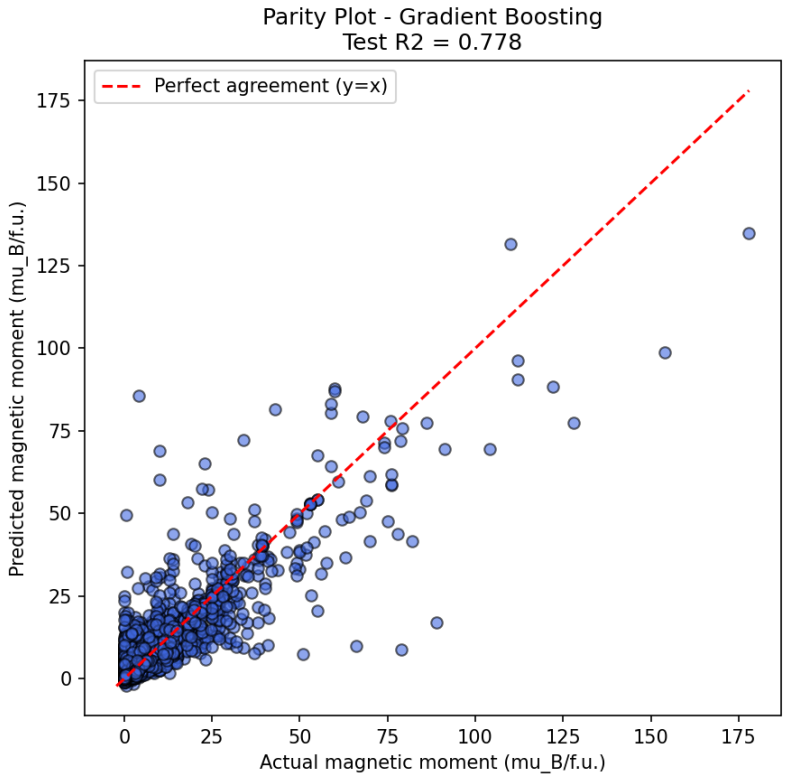


Figure 2: Parity plot of actual versus Gradient-Boosting-predicted magnetic moment per formula unit on the held-out test set (test $R^2 = 0.778$). The dashed red line denotes perfect agreement. Agreement is close for the densely populated low-moment region; scatter and systematic underprediction increase for the sparsely populated high-moment tail.

4 Discussion

4.1 Physical Interpretation of the Dominant Descriptors

The dominance of `num_magnetic_sites` calls for a careful, honest reading rather than an over-eager physical narrative. Because the regression target is normalized *per formula unit*, a compound with more magnetically active sites in its formula unit will, all else equal, tend toward a larger-magnitude net moment simply as a matter of counting – an essentially extensive relationship rather than a subtle electronic effect. We therefore regard this feature's dominance as a partly definitional artifact of how the target and this descriptor are both constructed from the same underlying site-counting information, rather than as a novel physical insight. This does not make the model useless – correctly weighting site count is necessary for good predictions – but it does mean the top-ranked feature should not be over-interpreted as revealing new magnetic physics.

The second-ranked feature, the mean Magpie ground-state magnetic moment of the constituent elements, is more genuinely informative: it encodes, at the level of the free-atom or elemental reference state, how strongly Hund's-rule-favoring the constituent $3d$/$4d$/$5d$ elements are before they are placed in a compound. That this elemental-level quantity survives as an important predictor after compound formation is physically sensible and consistent with Slater-Pauling-type behavior in transition-metal magnetism, where the alloy or compound moment correlates with the valence electron count and elemental magnetic character

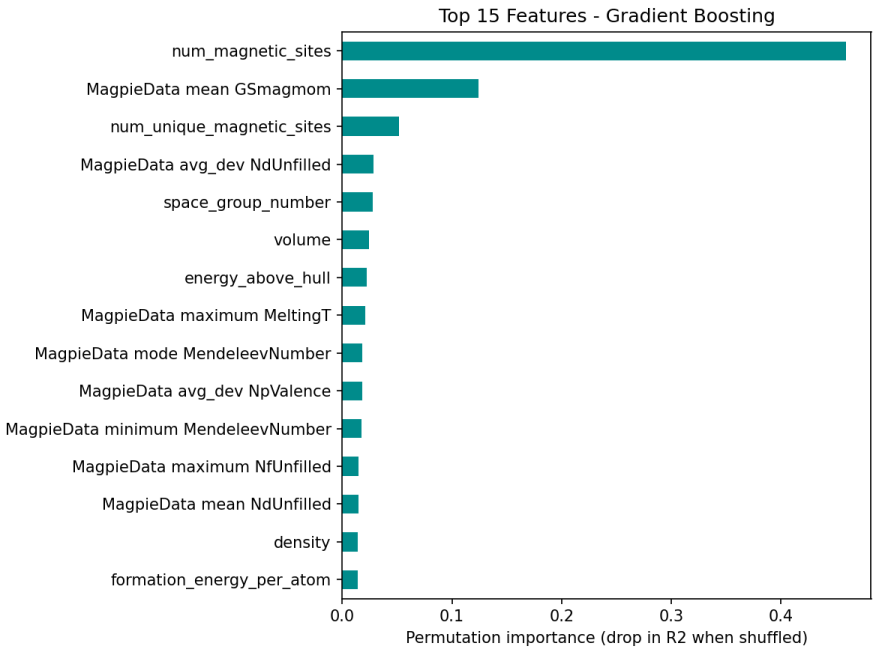


Figure 3: Top 15 features by permutation importance for the Gradient Boosting model, measured as the mean drop in test-set R^2 when each feature is independently shuffled (10 repeats).

of its constituents. The subsequent cluster of d- and f-shell filling descriptors (average deviation in unfilled d-electrons, unfilled f-electron counts) reinforces this picture: they capture how close the constituent elements sit to a half-filled or fully-filled shell configuration, which historically correlates with the size of the local magnetic moment in transition-metal and rare-earth-containing compounds. The presence of space group number among the top features is a useful, if modest, reminder that crystal symmetry – the same ingredient that ultimately distinguishes altermagnetic from conventional antiferromagnetic order – also carries statistical information about magnetic behavior even at the level of a scalar moment regression.

4.2 Relevance to Altermagnetic Candidate Screening

We emphasize what this model does and does not accomplish. It does not predict spin-splitting, magnetic space group, or altermagnetic order directly; those require spin-resolved band structure and symmetry classification that a scalar magnetic-moment regressor cannot provide. What it does provide is a computationally free way to flag, among tens of thousands of catalogued compounds, those whose predicted (and database-reported) net moment is close to zero despite non-trivial magnetic site counts and elemental magnetic character – i.e., compounds that are locally magnetic but globally compensated. Within the Materials Project's own magnetic ordering labels, such compounds are typically classified as antiferromagnetic; cross-referencing our moment predictions against that ordering label provides a computationally inexpensive first filter for exactly the kind of compensated, collinear magnetic candidates from which altermagnets are drawn [22, 1]. Confirming genuine altermagnetic order in any shortlisted compound would still require the symmetry-resolved (spin space group) analysis and, ideally, spin-resolved DFT or ARPES verification used in the confirmed cases of RuO_2, CrSb, and MnTe [12, 13, 17].

4.3 Limitations

Several limitations constrain how far these results should be extrapolated. First, and most importantly, the magnetic labels and total magnetization values used both as our target and as contextual information are drawn from Materials Project's automated, collinear spin-polarized DFT workflow, which Fahmy *et al.* recently showed carries a systematic bias toward ferromagnetic convergence relative to experimentally validated magnetic structures in the MAGNDATA database [35]. Because our model is trained to reproduce these DFT-derived labels, any such bias is inherited by our predictions; a compound flagged by our model as having near-zero moment reflects the DFT relaxation's outcome, not necessarily the true experimental ground state, and this caveat should be attached to any screening list derived from this work.

Second, the strongly right-skewed target distribution (Figure 1) means that the reported R^2 and error metrics are dominated by performance in the low-moment regime; the parity plot (Figure 2) shows that high-moment compounds are systematically underpredicted, a pattern consistent with the class-imbalance challenges documented elsewhere for magnetic moment regression [39]. Future work should explore log-transformed targets, quantile-based resampling, or explicitly weighted loss functions to improve tail performance.

Third, as discussed above, the dominant feature (`num_magnetic_sites`) is partly definitional given the per-formula-unit normalization of the target; its importance should not be read as a physical discovery, and models intended for genuine physical insight into compensation mechanisms would benefit from also reporting results with this feature excluded or from working with an intensive (per-atom or per-site) target formulation.

Fourth, this study only evaluates one train/test split, albeit cross-validated on the training portion; a fully nested cross-validation scheme, along with hyperparameter tuning (which was not performed here beyond fixed, reasonable defaults), could refine the reported performance figures.

Finally, no compound identified by our screening approach as a compensated candidate has been independently validated here against spin space group symmetry analysis or spin-resolved calculations; the present work is a triage tool for prioritizing candidates, not a confirmation method, and should be used accordingly.

5 Conclusion

We have developed a machine learning framework that predicts the net magnetic moment per formula unit for a broad set of transition-metal-bearing compounds retrieved from the Materials Project database, using structural, thermodynamic, and Magpie composition-based descriptors. A histogram-based Gradient Boosting model achieved a held-out test R^2 of 0.778, substantially outperforming a linear baseline and comparable to prior magnetic-moment machine learning studies on more restricted chemistries. Permutation importance analysis shows that the number of magnetic sites and the elemental-level ground-state magnetic moment are the dominant predictors, with d-shell filling, crystal symmetry, and thermodynamic descriptors contributing secondary but physically interpretable signal. We have argued, with explicit attention to the model's limitations – inherited database bias, target skew, and a partly definitional dominant feature – that such a regression model is best understood as a first-pass screening tool for identifying magnetically compensated candidates among catalogued transition-metal compounds, narrowing the space that more expensive spin- and symmetry-resolved analysis would need to examine in the search for new altermagnetic materials. Future work should validate candidate compounds flagged by this approach against experimentally grounded magnetic databases, incorporate spin space group classification directly into the screening pipeline, and address the identified class-imbalance and definitional-feature concerns to sharpen both the accuracy and the physical interpretability of the resulting picture.

Data Availability

Data were retrieved from the Materials Project database via the mp-api Python interface [23] and are publicly available at `https://materialsproject.org`. Composition-based features were generated using the open-source `matminer` and `pymatgen` packages [25, 24].

References

[1] L. Šmejkal, J. Sinova, and T. Jungwirth, "Emerging research landscape of altermagnetism," *Phys. Rev. X* **12**, 040501 (2022).

[2] L. Šmejkal, J. Sinova, and T. Jungwirth, "Beyond conventional ferromagnetism and antiferromagnetism: A phase with nonrelativistic spin and crystal rotation symmetry," *Phys. Rev. X* **12**, 031042 (2022).

[3] L. Šmejkal, R. González-Hernández, T. Jungwirth, and J. Sinova, "Crystal time-reversal symmetry breaking and spontaneous Hall effect in collinear antiferromagnets," *Sci. Adv.* **6**, eaaz8809 (2020).

[4] I. Mazin (The PRX Editors), "Editorial: Altermagnetism—a new punch line of fundamental magnetism," *Phys. Rev. X* **12**, 040002 (2022).

[5] L. Bai, W. Feng, S. Liu, L. Šmejkal, Y. Mokrousov, and Y. Yao, "Altermagnetism: Exploring new frontiers in magnetism and spintronics," *Adv. Funct. Mater.* **34**, 2409327 (2024).

[6] S. S. Fender, O. Gonzalez, and D. K. Bediako, "Altermagnetism: A chemical perspective," *ChemRxiv* preprint (2024).

[7] C. Song *et al.*, "Altermagnets as a new class of functional materials," *Nat. Rev. Mater.* **10**, 473–485 (2025).

[8] L. Šmejkal, A. H. MacDonald, J. Sinova, S. Nakatsuji, and T. Jungwirth, "Anomalous Hall antiferromagnets," *Nat. Rev. Mater.* **7**, 482–496 (2022).

[9] R. González-Hernández *et al.*, "Efficient electrical spin splitter based on nonrelativistic collinear antiferromagnetism," *Phys. Rev. Lett.* **126**, 127701 (2021).

[10] V. Baltz, A. Manchon, M. Tsoi, T. Moriyama, T. Ono, and Y. Tserkovnyak, "Antiferromagnetic spintronics," *Rev. Mod. Phys.* **90**, 015005 (2018).

[11] T. Jungwirth, X. Marti, P. Wadley, and J. Wunderlich, "Antiferromagnetic spintronics," *Nat. Nanotechnol.* **11**, 231–241 (2016).

[12] Z. Feng *et al.*, "An anomalous Hall effect in altermagnetic ruthenium dioxide," *Nat. Electron.* **5**, 735–743 (2022).

[13] M. Zeng *et al.*, "Observation of spin splitting in room-temperature metallic antiferromagnet CrSb," *Adv. Sci.* **11**, 2406529 (2024).

[14] J. Ding *et al.*, "Large band-splitting in g-wave type altermagnet CrSb," *Phys. Rev. Lett.* **133**, 206401 (2024).

[15] S. Reimers, L. Odenbreit, L. Šmejkal *et al.*, "Direct observation of altermagnetic band splitting in CrSb thin films," *Nat. Commun.* **15**, 2116 (2024).

[16] G. Yang *et al.*, "Three-dimensional mapping of the altermagnetic spin splitting in CrSb," *Nat. Commun.* **16**, 1442 (2025).

[17] J. Krempaský *et al.*, "Altermagnetic lifting of Kramers spin degeneracy," *Nature* **626**, 517–522 (2024).

[18] S. Lee *et al.*, "Broken Kramers degeneracy in altermagnetic MnTe," *Phys. Rev. Lett.* **132**, 036702 (2024).

[19] A. Bentien, S. Johnsen, G. K. H. Madsen, B. B. Iversen, and F. Steglich, "Colossal Seebeck coefficient in strongly correlated semiconductor $FeSb_2$," *Europhys. Lett.* **80**, 17008 (2007).

[20] M. Battiato, J. M. Tomczak, Z. Zhong, and K. Held, "Unified picture for the colossal thermopower compound $FeSb_2$," *Phys. Rev. Lett.* **114**, 236603 (2015).

[21] H. Takahashi *et al.*, "Colossal Seebeck effect enhanced by quasi-ballistic phonons dragging massive electrons in $FeSb_2$," *Nat. Commun.* **7**, 12732 (2016).

[22] I. I. Mazin, K. Koepernik, M. D. Johannes, R. González-Hernández, and L. Šmejkal, "Prediction of unconventional magnetism in doped $FeSb_2$," *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2108924118 (2021).

[23] A. Jain *et al.*, "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation," *APL Mater.* **1**, 011002 (2013).

[24] S. P. Ong *et al.*, "Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis," *Comput. Mater. Sci.* **68**, 314–319 (2013).

[25] L. Ward *et al.*, "Matminer: An open source toolkit for materials data mining," *Comput. Mater. Sci.* **152**, 60–69 (2018).

[26] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, "A general-purpose machine learning framework for predicting properties of inorganic materials," *npj Comput. Mater.* **2**, 16028 (2016).

[27] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature* **559**, 547–555 (2018).

[28] L. Breiman, "Random forests," *Mach. Learn.* **45**, 5–32 (2001).

[29] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

[30] G. Ke *et al.*, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*.

[31] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.* **29**, 1189–1232 (2001).

[32] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," *J. Mach. Learn. Res.* **20**, 1–81 (2019).

[33] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nat. Mach. Intell.* **1**, 206–215 (2019).

[34] S.-O. Kaba, B. Groleau-Paré, M.-A. Gauthier, A.-M. S. Tremblay, S. Verret, and C. Gauvin-Ndiaye, "Prediction of large magnetic moment materials with graph neural networks and random forests," *Phys. Rev. Mater.* **7**, 044407 (2023).

[35] A. E. Fahmy *et al.*, "Learning magnetic order classification from large-scale materials databases," *arXiv:2509.05909* (2025).

[36] N. C. Frey, M. K. Horton *et al.*, "High-throughput search for magnetic and topological order in transition metal oxides," *Sci. Adv.* **6**, eabd1076 (2020).

[37] A. Verma, J. Jami, and A. Bhattacharya, "Advancing magnetic materials discovery: A structure-based machine learning approach for magnetic ordering and magnetic moment prediction," *arXiv:2507.01913* (2025).

[38] T. L. Pham, H. Kino, K. Terakura, T. Miyake, K. Tsuda, I. Takigawa, and H. C. Dam, "Machine learning reveals orbital interaction in materials," *Sci. Technol. Adv. Mater.* **18**, 756–765 (2017).

[39] R. Y. Sanspeur and J. R. Kitchin, "Circumventing data imbalance in magnetic ground state data for magnetic moment predictions," *Mach. Learn.: Sci. Technol.* **5**, 015023 (2024).

[40] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.* **136**, B864–B871 (1964).

[41] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.* **140**, A1133–A1138 (1965).

[42] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

[43] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Phys. Rev. B* **54**, 11169–11186 (1996).

[44] T. Xie and J. C. Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties," *Phys. Rev. Lett.* **120**, 145301 (2018).

[45] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, "Graph networks as a universal machine learning framework for molecules and crystals," *Chem. Mater.* **31**, 3564–3572 (2019).