Explainable AI for Curie Temperature Prediction in Magnetic Materials

M. Adeel Ajaib[1], Fariha Nasir[2], and Abdul Rehman[3]

[1]*Department of Data Sciences, Penn State Abington, Abington, PA 19001, USA*
[2]*Department of Physics, Penn State Abington, Abington, PA 19001, USA*
[3]*Department of Physics,University of Delaware, Newark, DE 19716, USA*

Abstract

We explore machine learning techniques for predicting Curie temperatures of magnetic materials using the NEMAD database. By augmenting the dataset with composition-based and domain-aware descriptors, we evaluate the performance of several machine learning models. We find that the Extra Trees Regressor delivers the best performance reaching an R^2 score of up to 0.85 ± 0.01 (cross-validated) for a balanced dataset. We employ the k-means clustering algorithm to gain insights into the performance of chemically distinct material groups. Furthermore, we perform the SHAP analysis to identify key physicochemical drivers of Curie behavior, such as average atomic number and magnetic moment. By employing explainable AI techniques, this analysis offers insights into the model's predictive behavior, thereby advancing scientific interpretability.

1 Introduction

Curie temperature (T_C) is a critical property that defines the transition of a material from a ferromagnetic to a paramagnetic state. Predicting T_C with high accuracy is essential for the discovery of magnetic materials for data storage, spintronics, and energy applications. Traditional approaches based on quantum mechanical computations or empirical models are often limited in scalability and accuracy. In recent years, machine learning (ML) has emerged as a promising alternative for property prediction across materials science domains [1–5].

Building on this momentum, several works have proposed the use of ML models trained on curated magnetic datasets. In particular, the recent study [6] introduced the NEMAD database, which aggregates experimentally measured magnetic transition temperatures and compositions. Similarly, the study by Belot et al. [7] utilized two of the largest available datasets of experimental Curie temperatures—comprising over 2,500 materials for training and more than 3,000 entries for validation—to compare machine learning strategies for predicting Curie temperature solely from chemical composition.

Our work is inspired by these prior efforts and aims to improve the predictive accuracy and gain insights into model interpretability. We develop a pipeline that starts from the NEMAD dataset, augments it with compositional and elemental features, and evaluates several ML models. A key contribution of our work is the integration of explainable AI (XAI) through SHAP (SHapley Additive exPlanations) analysis, which allows us to quantify how each input feature contributes to the model's prediction. Moreover, we benchmark our models on external datasets from literature to demonstrate generalization.

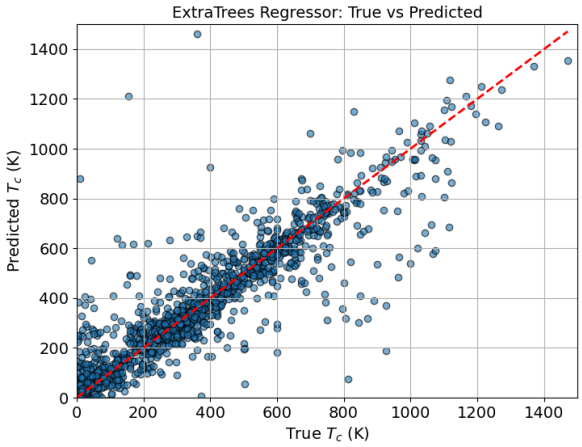
2 Dataset Preparation and Model Evaluation

We begin our study with the NEMAD database [6], a publicly available dataset compiled using large language models, comprising 26,706 magnetic materials with records for chemical composition, phase transition temperatures (e.g., Curie or Néel temperatures) and related magnetic properties. We preprocessed the database to remove ambiguous entries, duplicates, and entries with missing or non-numeric temperatures. We also standardized chemical formulas using the pymatgen library.

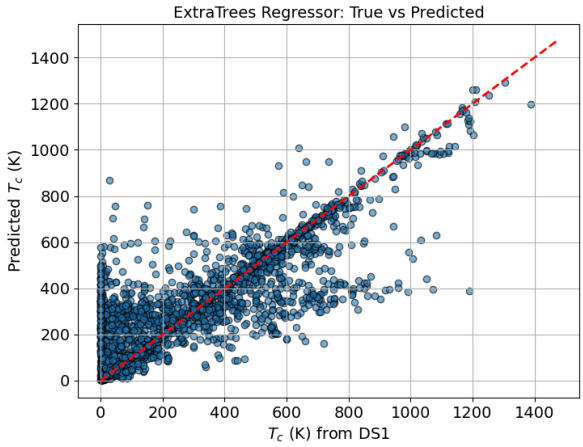
To generate descriptors, we first computed composition-based features such as average atomic number, atomic mass, group number, period, and electronegativity using pymatgen. We then added elemental property descriptors using the Magpie feature set via matminer, which includes statistics such as mean, range, and standard deviation of valence electron counts, ionic radii, and oxidation states. These features capture the diversity and bonding characteristics of constituent elements and have been shown to correlate with magnetic properties.

We evaluated several machine learning models for predicting Curie temperatures using the enriched dataset. Table 1 presents the comparative performance metrics in terms of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2).

The Extra Trees algorithm achieved the best overall performance, with an R^2 value of 0.87, MAE of 54 K, and RMSE of 106 K. This performance surpasses that of the Random Forest algorithm and other models considered. With a balanced dataset we were able to achieve $R^2 \sim 0.85$ with a standard



(a) Main dataset



(b) DS1 validation dataset

Figure 1: Predicted vs. true Curie temperatures using the ExtraTrees model for (a) the training/test dataset and (b) the DS1 validation dataset. A fairly strong alignment along the diagonal indicates the predictive performance of the model.

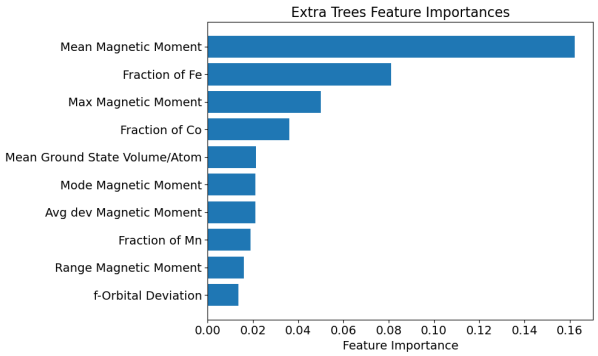


Figure 2: Feature importance graph for the Extra Trees Algorithm.

deviation of 0.01. The balanced dataset was obtained by randomly removing around 4% of samples with Curie temperatures less than 300 K [6].

The Extra Trees (Extremely Randomized Trees) algorithm employs an ensemble learning method similar to the Random Forests model, but with two notable distinctions:

- It selects cut-points at random rather than optimizing splits.
- It uses the entire training dataset without bootstrapping.

These differences introduce greater variance among the individual trees in the ensemble, which can improve generalization and reduce overfitting on noisy datasets. This may explain its superior performance in our analysis, where the feature space is high-dimensional and the temperature distribution is non-uniform.

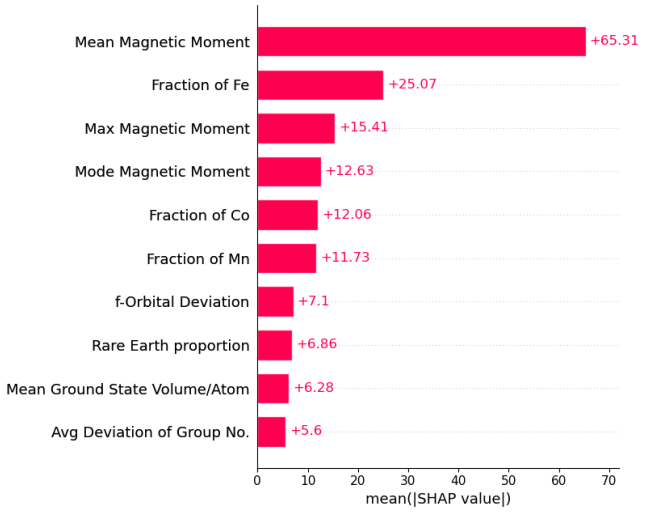
Table 1: Performance comparison of machine learning models for Curie temperature prediction after feature normalization where necessary.

Model	**MAE (K)**	**RMSE (K)**	**R^2**
Extra Trees	51	98	0.87
Random Forest	57	103	0.85
XGBoost	63	106	0.85
Neural Network	69	123	0.79
KNN	74	129	0.77

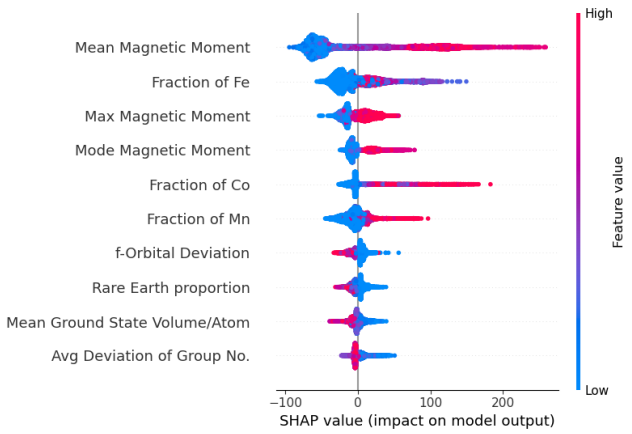
The Random Forest model also performs well but slightly underperforms Extra Trees, likely due to its more deterministic node-splitting strategy, which may be less adaptive to irregularities in the dataset.

XGBoost, KNN, and neural networks were also tested. XGBoost shows competitive results but did not outperform tree-based ensembles. KNN and neural networks showed relatively weaker performance, potentially due to the complexity of the feature relationships and sensitivity to scaling.

We validate the performance of the Extra Trees model on the DS1 dataset [19]. Figure 1 presents the predicted versus actual Curie temperatures using the ExtraTrees model, both for the main dataset and the DS1 validation test dataset. In both subfigures, the data points cluster along the $y = x$ diagonal, indicating that the predicted and actual values agree fairly well ($R^2 \sim 0.71$, MAE $\sim$ 91 K) which demonstrates the model's generalizability.



(a) SHAP bar plot showing the average impact of each feature on model output.



(b) SHAP beeswarm plot showing feature influence per sample. Color indicates feature value.

Figure 3: SHAP analysis of the ExtraTrees model. (a) shows the average contribution of the top 10 features, while (b) reveals the variation of feature impacts for individual predictions.

2.1 SHAP Analysis

To better understand the model's behavior and identify the most influential features contributing to the prediction of Curie temperature, we performed SHAP (SHapley Additive exPlanations) analysis. SHAP values offer a model-agnostic, consistent way to explain predictions by attributing importance scores to individual features.

SHAP Bar Plot: Figure 3 shows the SHAP summary bar plot of the top 10 features based on their mean absolute SHAP values. Among the dominant predictors, the mean magnetic moment stands out, reinforcing the well-established role of atomic-scale magnetism in determining T_C. The fraction of Fe also shows high importance, which aligns with its prevalence in high-T_C compounds and well-known ferromagnets. Features such as the mode magnetic moment, max magnetic moment, and fraction of Co all rank highly, further confirming that the magnetic nature of the constituent atoms is a primary driver of the model's predictions.

Interestingly, features such as rare earth proportion, mean ground state volume per atom, and group number deviation appear in the top contributors. These descriptors capture the complexity of atomic arrangements and compositional heterogeneity, both of which are known to modulate magnetic interactions in complex systems.

SHAP Beeswarm Plot: The SHAP beeswarm plot Figure (3), right panel, illustrates the distribution of SHAP values for each feature across all samples. Each point represents a single material, with color indicating the corresponding feature value. This visualization reveals not only which features are most influential, but also how they affect the model's predictions. For example, we can see from the plot that the Mean Magnetic Moment, which represents the average magnetic moment of the ground-state structure, exhibits a strong positive influence when its value is high (shown in red), shifting SHAP values to the right. This suggests that materials with higher magnetic moments are strongly associated with higher Curie temperatures, which is physically intuitive.

In contrast, the f-Orbital Deviation shows a somewhat different pattern. High values (in red) are generally associated with negative SHAP values, indicating a suppressive effect on Curie temperature. This deviation metric captures the compositional spread in elements possessing f-electrons, such as rare-earth or actinide elements.

Taken together, these plots indicate that the model captures meaningful physical trends: magnetic moment-related features dominate, while elemental diversity exhibit subtle patterns. This supports the model's reliability and its alignment with known physics.

3 K-means Clustering Analysis

In this section, we present our implementation of the k-means clustering algorithm to identify subgroups within the dataset that may exhibit distinct magnetic behavior. To determine the optimal number of clusters, we employed the elbow method, as shown in Figure 4. The curve shows a sharp drop in inertia from $k = 1$ to $k = 3$, with diminishing returns beyond $k = 4$. This led us to select $k = 4$ for clustering.

We then applied the Extra Trees algorithm independently to each cluster and observed that Clusters 1 and 3 exhib-

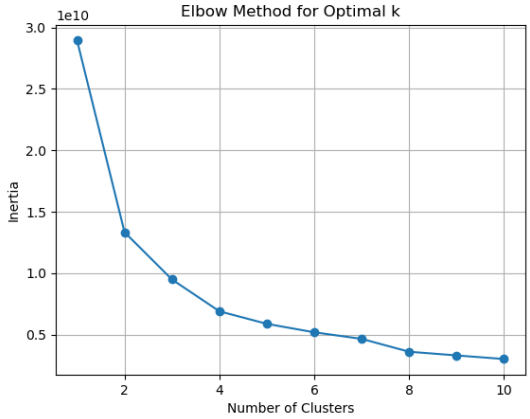


Figure 4: Elbow plot showing the inertia values for different numbers of clusters. The curve begins to flatten around $k = 4$, suggesting a reasonable balance between variance capture and model complexity.

ited significantly lower predictive performance, with R^2 values of approximately 0.58 and 0.38, respectively. Cluster 1 was found to be dominated by oxygen-rich compounds, particularly complex oxides, as illustrated in Figure 5. Examples of compounds in this cluster include, $LaMnO_3$, $SrFeO_3$, and $CaMnO_3$. From a physical standpoint, such oxide systems differ substantially from the metallic systems that dominate other clusters. Their magnetism is governed by mechanisms like superexchange, strong electron correlation, and structural distortions (e.g., Jahn–Teller effects), which are not well captured by simple compositional descriptors. This explains the degradation in model performance within this cluster.

Cluster 3 yields the lowest R^2 compared to other clusters, with the highest MAE and RMSE values among all clusters. This underperformance likely arises from both its small size and its unique chemical composition. Unlike other clusters dominated by transition metal oxides or alloys, Cluster 3 contains a high proportion of carbon, and phosphorus-rich compounds, including intermetallics and carbides. These materials often exhibit complex bonding environments and non-collinear magnetic structures, which are challenging to model using composition-only descriptors. The limited number of training samples in this chemically diverse space further hinders generalization.

Therefore, the under performance of the model in Clusters 1 and 3 provides a strong motivation for incorporating structure-aware or quantum-level descriptors in future studies. Their distinct magnetic behaviors and modeling challenges highlight the limitations of current descriptor sets and the need for more physically grounded representations.

Consequently, excluding these clusters from our analysis improves overall model performance and this improves the value of R^2 to approximately 0.85. This decision is supported both statistically and physically and highlights the importance of domain knowledge when applying machine learning to datasets for magnetic materials. After removing Clusters 1 and 3 from the data, we obtain the following results from 5-fold cross validation:

Table 2: Cross-validation (5-fold) results for the best-performing model.

Metric	**Value**	**Mean $\pm$ Std**
MAE (K)	54	54 ± 2.5
RMSE (K)	105	105 ± 4.0
R^2	0.85	0.85 ± 0.012

t-SNE Visualization of Clusters

To further understand the separation and characteristics of clusters formed in the dataset, we employed t-SNE (t-distributed Stochastic Neighbor Embedding) to reduce the high-dimensional feature space to two dimensions. The resulting 2D visualization, shown in Figure 5, illustrates clear separation between chemically distinct groups.

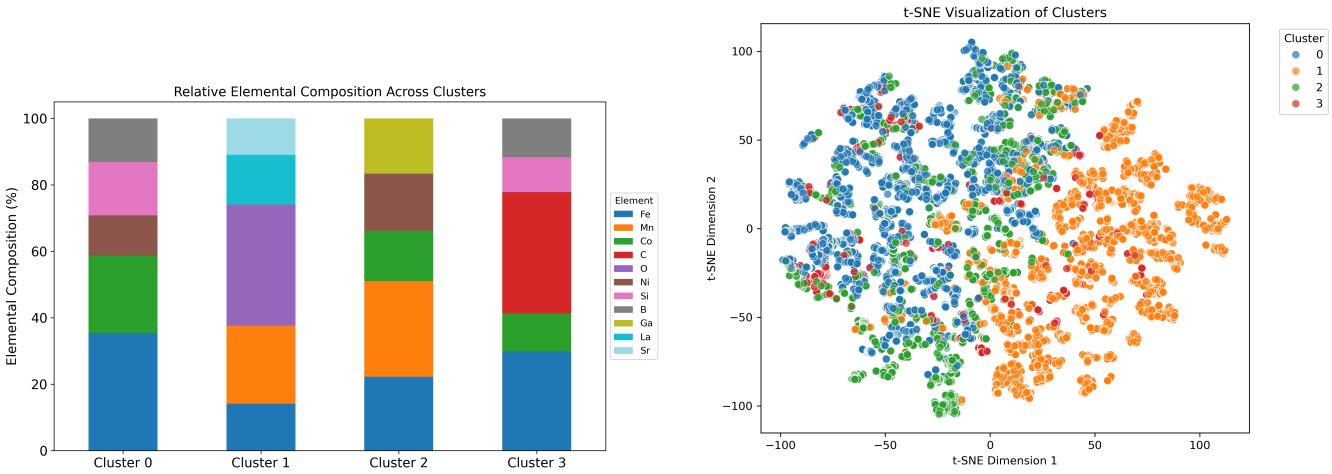
Cluster 0 contains high concentrations of metallic elements such as iron (Fe), cobalt (Co), silicon (Si), and boron (B), reflecting typical magnetic intermetallics. Cluster 1 is dominated by complex oxides, rich in oxygen (O), manganese (Mn), lanthanum (La), and strontium (Sr)—a chemical signature common in perovskite and layered oxide systems. Cluster 2 is enriched in manganese (Mn), iron (Fe), gallium (Ga), and nickel (Ni), resembling Heusler-type alloys. Cluster 3 includes mainly carbon (C), iron (Fe), and Cobalt (Co), and represents simpler binary and ternary compounds.

These clusters not only form distinct regions in t-SNE space but also exhibit coherent chemical themes, validating the clustering strategy and emphasizing the influence of composition on Curie temperature behavior.

4 Cluster-wise SHAP Analysis

To further enhance the interpretability of our machine learning models for Curie temperature (T_c), we applied the SHAP analysis for each of the four clusters identified via KMeans. By isolating materials into chemically similar groups, we can uncover subtle relationships and also better understand which descriptors govern T_c in different regimes.

Cluster 0, which is the largest cluster (4,744 materials), contains a significant number of Fe-, Co-, and rare-earth-based compounds. The model achieves strong performance with MAE = 86.83 K and $R^2 = 0.80$. As shown in top left panel of Fig. 6, the most influential features include the mean magnetic moment, mode magnetic moment, and fraction of Co, all positively correlated with T_c. These results reflect known physical



(a) Cluster-wise elemental composition percentages

(b) t-SNE projection with cluster labels

Figure 5: Figure 3(a) visualizes the dominant elements in each cluster, reinforcing chemically meaningful groupings. Figure 3(b) shows the t-SNE visualization for the dataset, where each point represents a material colored by its assigned cluster. The separation among clusters, particularly the distinct distribution of Cluster 1, suggests notable differences in compositional properties that impact Curie temperature prediction.

principles — high magnetic moments enhance exchange interactions and stabilize ferromagnetic ordering. Additional features such as minimum Mendeleev number and mode column suggest that periodic trends play a secondary yet non-trivial role.

Cluster 1 (4,826 materials) includes many oxide-based perovskites rich in Mn, La, and O. The model performance is lower (MAE = 100.21 K, $R^2 = 0.58$), likely due to the structural and chemical complexity of these materials. Fig. 6 (top right plot) reveals that fraction of Fe, average deviation of magnetic moment, and f-orbital deviation are key predictive features. The prominence of orbital and magnetic variance descriptors suggests the presence of competing magnetic mechanisms such as double exchange and Jahn-Teller distortions. These subtleties may explain the slightly reduced predictive power compared to Cluster 0.

Cluster 2 (2,730 materials) consists of Mn- and Ni-based intermetallics, including Heusler-like alloys. Despite compositional richness, the model achieves good performance (MAE = 88.68 K, $R^2 = 0.73$). In Fig. 6 (bottom left plot), features such as mean magnetic moment, fraction of Mn, and maximum covalent radius dominate. This aligns with the tunability of intermetallic compounds, where structural and electronic changes can significantly influence magnetic ordering. The importance of spatial descriptors further highlights the role of atomic packing and size effects in modulating T_c.

Cluster 3 is the smallest cluster (760 materials) and exhibits the poorest performance (MAE = 138.88 K, $R^2 = 0.38$). This may be due to its small size and its unique chemical composition. As seen in Fig. 6 (bottom right plot), while magnetic descriptors still lead, there is greater importance given to features such as La fraction, Sr fraction, and f-orbital deviation. These may hint at more subtle magnetic phenomena that our current descriptor set cannot fully capture. The high variance in SHAP values also suggests that no single feature consistently governs T_c across the cluster, pointing to the need for more tailored descriptors or domain-specific models.

SHAP analysis across clusters confirms that magnetic moment descriptors are universally important predictors of Curie temperature. However, the prominence of different supporting features (e.g., orbital variances, group numbers, atomic radii) shifts based on the chemical and structural nature of the materials within each cluster. This highlights that even with a shared target, distinct materials demand distinct explanatory regimes.

We conclude that cluster-specific modeling not only improves performance but also enriches physical insight. Future work may benefit from hybrid clustering methods (e.g., hierarchical or domain-guided) and the integration of additional physics-based descriptors tailored to each family. It is also worth noting that while SHAP helps explain model behavior, it doesn't always align with physical intuition, especially for complex oxides. This is expected, as SHAP reflects correlations based on available descriptors, which may miss deeper physics. Its insights are valuable, but approximate.

5 Conclusion

This study demonstrates that explainable machine learning can serve as a powerful tool for materials property prediction, par-

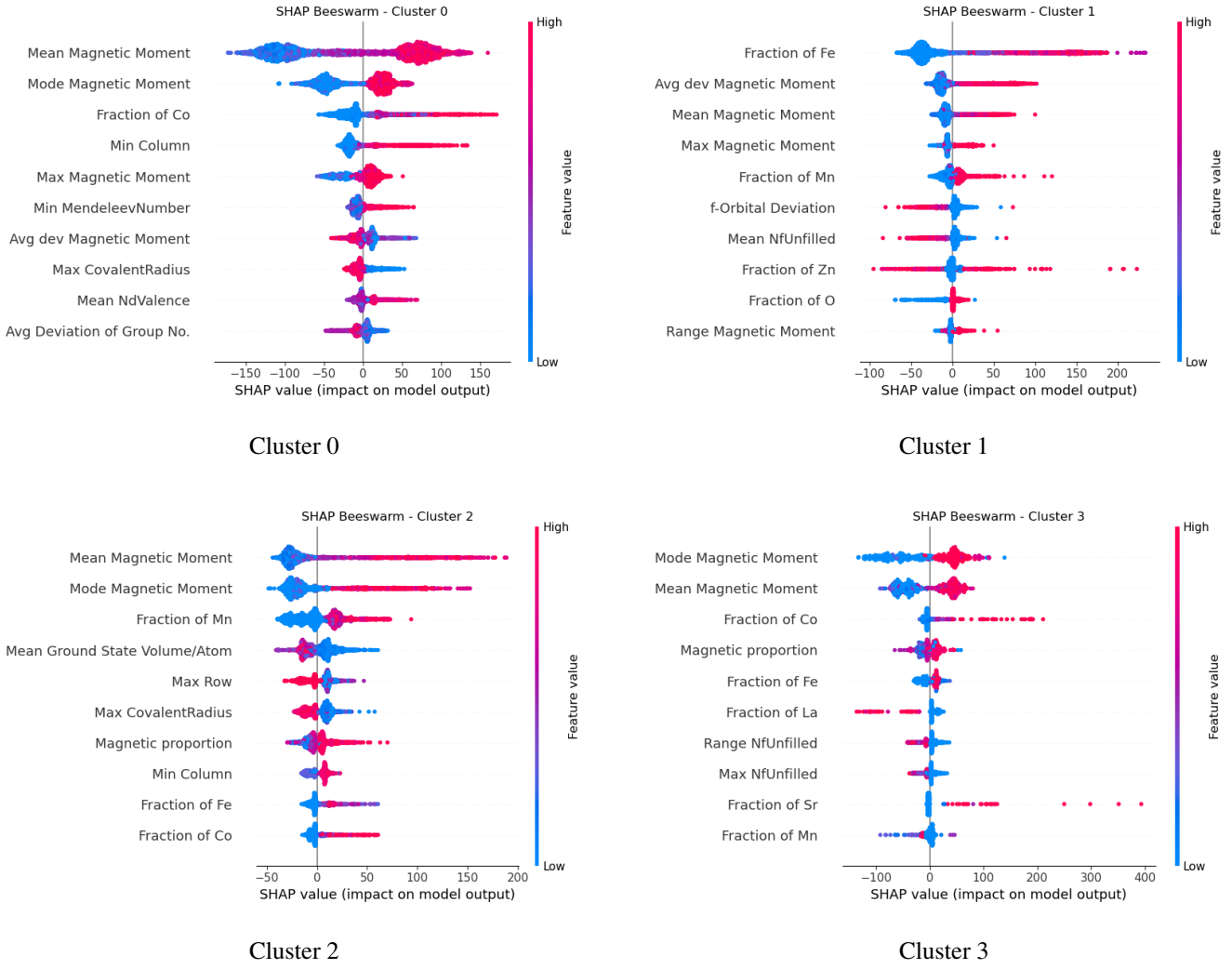


Figure 6: SHAP beeswarm summary plots for each cluster (0 to 3). These plots visualize the most influential features determining T_c predictions in each cluster.

ticularly for estimating Curie temperatures using composition-based descriptors. We employed various machine learning models and found that, compared to other models, the optimized Extra Trees Regressor demonstrated robust predictive performance for Curie temperature estimation.

Furthermore, by leveraging explainable machine learning, we uncover feature-driven insights into the compositional factors that influence ferromagnetism in inorganic materials. The use of SHAP analysis allowed us to interpret model outputs in a physically meaningful way. Additionally, the application of clustering and t-SNE visualization revealed distinct patterns and subgroups within the dataset.

Our framework demonstrates that incorporating explainable AI not only improves transparency but also further strengthens the scientific value of machine learning models in materials science. This approach can be readily extended to other property prediction tasks across the domain.

6 Acknowledgments

We would like to thank Babar Shabbir for useful discussions.

The code and data for this project is available on github: https://github.com/dradeelajaib/Curie-Prediction

References

[1] Rajan, K., 2015. Materials informatics: The materials "gene" and big data. *Annual Review of Materials Research*, 45, pp.153–169.

[2] Jung, S. G., et al. "Machine-learning prediction of Curie temperature from chemical compositions of ferromagnetic materials." *Journal of Chemical Information and Modeling* 64.16 (2024): 6388–6409. https://doi.org/10.1021/acs.jcim.4c00947

[3] Lu, S., Zhou, Q., Guo, Y., and Wang, J. "On-the-fly interpretable machine learning for rapid discovery of two-dimensional ferromagnets with high Curie temperature." *Chem* 8.3 (2022): 769–784. https://doi.org/10.1016/j.chempr.2021.12.004

[4] Bränvall, M. A., Persson, G., Casillas-Trujillo, L., Armiento, R., and Alling, B. "Predicting the Curie temperature of magnetic materials with automated calculations across chemistries and structures." *Physical Review Materials* 8 (2024): 114417. https://doi.org/10.1103/PhysRevMaterials.8.114417

[5] Ward, L., et al. "A general-purpose machine learning framework for predicting properties of inorganic materials." npj Computational Materials 2 (2016): 16028.

[6] Itani, S., Zhang, Y., Zang, J. (2024). Northeast Materials Database (NEMAD): Enabling Discovery of High Transition Temperature Magnetic Compounds. arXiv preprint arXiv:2409.15675.

[7] Belot, J.F., Taufour, V., Sanvito, S., Hart, G.L.W. (2023). Machine learning predictions of high-Curie-temperature materials. *arXiv preprint* arXiv:2307.06879.

[8] Lundberg, S.M. and Lee, S.-I., 2017. A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (pp. 4765-4774).

[9] Tshitoyan, V., et al., 2019. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763), pp.95–98.

[10] Blundell, S., 2001. *Magnetism in Condensed Matter*. Oxford University Press.

[11] Klie, R.F., et al., 2005. Direct evidence for charge inhomogeneities at the Curie temperature of $La_{1-x}Ca_xMnO_3$. *Phys. Rev. B*, 71(14), p.144508.

[12] Schmidt, J., et al., 2019. Recent advances and applications of machine learning in solid-state materials science. *Chem. Rev.*, 119(12), pp.7359–7413.

[13] Balachandran, P.V., et al., 2018. Adaptive strategies for materials design using uncertainties. *npj Comput. Mater.*, 4(1), p.43.

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

[15] Goodenough, J.B. and Zhou, J.S., 2006. Localized to itinerant electronic transitions in perovskite oxides. *Chem. Mater.*, 18(12), pp.2787–2794.

[16] Xie, T. and Grossman, J.C., 2018. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.*, 120(14), p.145301.

[17] Curtarolo, S., Hart, G.L.W., Nardelli, M.B., Mingo, N., Sanvito, S., Levy, O. (2013). The high-throughput highway to computational materials design. *Nature Materials*, 12(3), 191–201.

[18] Sanvito, S., Oses, C., Xue, J., Tiwari, A., Zic, M., Archer, T., Tozman, P., Venkatesan, M., Coey, M., Curtarolo, S. (2017). Accelerated discovery of new magnets in the Heusler alloy family. *Science Advances*, 3(4), e1602241.

[19] Nelson, J. Sanvito, S. (2019). Predicting the Curie temperature of ferromagnets using machine learning. *Phys. Rev. Materials*, 3(10), 104405.

[20] Byland, J.K., Shi, Y., Parker, D.S., Zhao, J., Ding, S., Mata, R., Magliari, H.E., Palasyuk, A., Bud'ko, S.L., Canfield, P.C. (2022). Statistics on magnetic properties of Co compounds: A database-driven method for discovering Co-based ferromagnets. *Phys. Rev. Materials*, 6(6), 063803.

[21] Merker, H.A., Heiberger, H., Nguyen, L., Liu, T., Chen, Z., Andrejevic, N., Drucker, N.C., Okabe, R., Kim, S.E., Wang, Y. (2022). Machine learning magnetism classifiers from atomic coordinates. *iScience*, 25(7), 105192.

[22] Kabiraj, A., Kumar, M., Mahapatra, S. (2020). High-throughput discovery of high Curie point two-dimensional ferromagnetic materials. *npj Computational Materials*, 6(1), 35.

[23] Vedmedenko, E.Y., Kawakami, R.K., Sheka, D.D., Gambardella, P., Kirilyuk, A., Hirohata, A., Binek, C., Chubykalo-Fesenko, O., Sanvito, S., Kirby, B.J. (2020). The 2020 magnetism roadmap. *J. Phys. D: Appl. Phys.*, 53(45), 453001.

[24] Coey, J.M.D. (2010). *Magnetism and Magnetic Materials*. Cambridge University Press.

[25] Graf, T., Felser, C., Parkin, S.S.P. (2011). Simple rules for the understanding of Heusler compounds. *Progress in Solid State Chemistry*, 39(1), 1–50.