



AAUP JOURNAL OF STEM AND
HEALTH SCIENCES (JSH-AAUP)

AAUP Journal of STEM and Health Sciences (JSH-AAUP)

Explainable Multi-Target Machine Learning Framework for Predicting Thermophysical Properties of Chemical Elements Using Periodic Table Descriptors

Feras Abujaber^{1,*}, Faady Siouri²

^{1,2} Department of Industrial Chemistry, Faculty of Sciences, Arab American University (AAUP), P.O. Box 240, Jenin, Palestine

ORCID: <https://orcid.org/0000-0001-5596-1093>

*Corresponding Author Email: Feras.Abujaber@aaup.edu

Date Received: 12/08/2026

Date Revised: 20/08/2026

Date Accepted: 02/09/2026

Date Published: 30/09/2026

DOI: <https://doi.org/10.35517/JSH-AAUP.v1.02.04>

Abstract

The thermophysical properties of the chemical elements underpin materials selection, process design, and fundamental chemistry, yet their experimental determination remains difficult for rare, radioactive, or synthetic species. This study developed an explainable, multi-target machine learning framework that predicted four key properties, namely melting point, boiling point, density, and atomic radius, directly from readily available periodic-table descriptors. Twelve descriptors, including atomic number, atomic mass, period, group, electronegativity, ionization energy, electron affinity, and valence-shell electron count derived from electron configurations, were compiled for the elements. Eight regression algorithms spanning linear, kernel, instance-based, and tree-ensemble families were systematically benchmarked using repeated five-fold cross-validation, with all targets modeled on a natural-logarithmic scale and missing descriptors handled by leakage-free iterative imputation. Tree-based ensembles considerably outperformed the linear models, which did not account for the nonlinear nature of the periodic relationships. The best mean performance was provided by the Extra Trees algorithm, which served as the framework base learner. After optimization, the framework reached coefficients of determination of 0.72, 0.76, 0.88, and 0.74 for the melting point, boiling point, density, and atomic radius, respectively, alongside low values of root-mean-square error. Model behavior was interpreted using Shapley additive explanations and permutation importance. Ionization energy emerged as the dominant descriptor for melting point, boiling point, and density owing to its correlation with bonding strength and electron localization, whereas the period governed atomic radius in line with the shell structure. The largest prediction errors were associated with elements exhibiting anomalous bonding (such as carbon, boron, and phosphorus) and strong relativistic effects (such as gold and mercury).

Keywords: machine learning, explainable artificial intelligence, thermophysical properties, periodic descriptors, ensemble learning, Shapley additive explanations.

1. Introduction

The physical and chemical behavior of every material ultimately traces back to the properties of its constituent elements. Among these, thermophysical properties such as the melting point, boiling point, density, and atomic radius are particularly significant because they govern phase stability, heat-transfer considerations, alloying behavior, separation techniques, and many other engineering choices [1]. Melting and boiling points give the temperatures at which solid–liquid and liquid–gas phase transitions occur and reflect the degree of interatomic bonding; density expresses the mass per unit volume and is critical in structural and nuclear applications; and the atomic radius influences solid-state atomic packing, coordination, and substitution. Knowledge of all these properties is necessary before designing novel materials. Traditionally, elemental properties have been measured directly or estimated from first-principles approaches. Direct measurement is more accurate but challenging, hazardous, or practically infeasible for rare, highly radioactive, or synthetic elements, of which only a few atoms may ever be produced. First-principles methods based on density functional theory [2] can predict several physical properties from the electronic configuration but are computationally expensive and difficult to apply, especially for finite-temperature properties such as the melting point, which is governed by complex collective behavior [9]. This has motivated the development of data-driven alternatives for estimating elemental properties from simple descriptors.

Despite considerable progress, most studies focus on a single property at a time and on complex, composition-based descriptors that require specialized software to compute. There remains a need for a unified, transparent framework that predicts several fundamental elemental properties simultaneously from a small set of intuitive periodic-table descriptors, and that explicitly explains its predictions in chemical terms. Addressing this gap, the present work develops an explainable multi-target machine learning framework for the joint prediction of melting point, boiling point, density, and atomic radius of the chemical elements.

The main contributions of this study are fourfold. First, a compact set of twelve periodic-table descriptors, several of which are engineered from electron configurations, is assembled and shown to be sufficient for accurate property prediction. Second, eight regression algorithms are systematically benchmarked under a rigorous repeated cross-validation protocol, establishing extremely randomized trees as the most effective base learner across all four targets. Third, the optimized framework is shown to achieve high predictive accuracy despite the small size of the periodic table and the presence of missing descriptor values, which are handled by a leakage-free imputation strategy. Fourth, and most importantly, Shapley additive explanations (SHAP) and permutation-importance analyses are used to interpret the framework, revealing chemically meaningful descriptor–property relationships and rationalizing the residual errors in terms of anomalous bonding and relativistic effects.

2. Related Work

Over the past decade, machine learning has emerged as a powerful third paradigm in the chemical and materials sciences, complementing experiment and simulation [3]. Data-driven models learn statistical relationships between the composition or structure of a material and its properties, enabling rapid screening of vast chemical spaces [4]. Comprehensive reviews have documented the rapid growth of materials informatics and its successful application to formation energies, band gaps, mechanical properties, and many other targets [5,6]. Representative studies include the prediction of molecular atomization energies with sub-kilocalorie accuracy [7], the acceleration of property predictions for polymers and inorganic solids [8], the estimation of melting temperatures of single- and binary-component solids from density functional theory data [9], and the prediction of band gaps of inorganic compounds [10]. Collectively, these works

demonstrate that carefully chosen descriptors combined with flexible regression models can reproduce quantum-mechanical and experimental data at a small fraction of the cost.

Ensemble methods based on decision trees are among the most effective and widely used algorithms for tabular property prediction. Random forests [11], gradient boosting machines [12], extreme gradient boosting [13], and extremely randomized trees [14] combine many weak learners to capture nonlinear relationships and feature interactions while remaining robust to noise and requiring relatively little tuning. Their high performance, however, is accompanied by a loss of interpretability, since a large number of trees makes the model behave like a black box, which is problematic in scientific contexts where the interpretability of predictions is important.

The field of explainable artificial intelligence has developed rigorous tools to address this problem. Explanations based on Shapley values (SHAP), which rely on cooperative game theory, provide a locally faithful and consistent attribution of a model output to its input features [15], and an exact and efficient method exists for SHAP computation in tree ensembles [16]. Local interpretable model-agnostic explanations [17] and the broader suite of interpretable machine learning techniques [18] now make it feasible to look inside the black box and relate model predictions to underlying domain knowledge. This offers an opportunity not only to generate accurate predictions but also to validate that the discovered relationships are consistent with established periodic trends. To the best of our knowledge, no prior study has combined a compact periodic-descriptor representation with multi-target ensemble learning and explainability analysis for the joint prediction of these four elemental properties.

3. Methodology

3.1 Dataset

The study is based on a publicly available periodic-table dataset containing the 118 known chemical elements and their tabulated physical and chemical characteristics (PubChem Periodic Table Dataset). For each element the dataset provides the atomic number, symbol, name, atomic mass, ground-state electron configuration, electronegativity (Pauling scale), atomic radius, first ionization energy, electron affinity, common oxidation states, standard state, and the four target thermophysical properties considered here: melting point (K), boiling point (K), density (g cm^{-3}), and atomic radius (pm), together with a qualitative element category (for example alkali metal, transition metal, or noble gas). The melting and boiling points are expressed as absolute temperatures, and densities correspond to standard conditions.

As is characteristic of the periodic table, the data are incomplete for the heaviest elements: the superheavy synthetic elements beyond curium have few or no measured thermophysical properties, and several descriptors such as electron affinity are undefined or unmeasured for a number of elements. To construct a consistent multi-target learning problem, elements possessing measured values for all four target properties were retained, yielding a modeling set of 93 elements spanning atomic numbers 1 to 96. This complete-target requirement was necessary because multi-output evaluation requires every target to be defined for each element.

3.2 Feature Engineering

Twelve periodic-table descriptors were used as model inputs. Eight are tabulated directly: atomic number, atomic mass, electronegativity, first ionization energy, and electron affinity, together with the period and group, which were assigned from the atomic number using the standard structure of the periodic table (lanthanides and actinides were placed in group 3). The remaining four descriptors were engineered from the ground-state electron configuration by parsing the occupancy of the valence subshells, giving the numbers of valence s, p, d, and f electrons; their sum defines the total number of valence electrons. These configuration-derived features compactly encode the position of an element within its block and

its bonding capacity. The complete descriptor set and its definitions are summarized in Table 1. The four target properties were deliberately excluded from the descriptor list, since they were to be predicted by the framework from fundamental atomistic descriptors alone and not from any other thermodynamic property.

Table 1. Periodic-table descriptors used as inputs to the machine learning framework.

Descriptor	Symbol / unit	Description
Atomic number	Z	Number of protons in the nucleus
Atomic mass	u	Relative atomic mass
Period	—	Horizontal row of the periodic table (1–7)
Group	—	Vertical column (1–18); f-block assigned to group 3
Electronegativity	χ	Pauling electronegativity
Ionization energy	eV	First ionization energy
Electron affinity	eV	Electron affinity
Valence electrons	—	Total number of valence-shell electrons
Valence s-electrons	—	Electrons in the valence s subshell
Valence p-electrons	—	Electrons in the valence p subshell
Valence d-electrons	—	Electrons in the valence d subshell
Valence f-electrons	—	Electrons in the valence f subshell

3.3 Data Preprocessing

Two preprocessing steps were essential. First, the target properties span several orders of magnitude; the density, for example, ranges from about $9 \times 10^{-5} \text{ g cm}^{-3}$ for hydrogen to more than 22 g cm^{-3} for osmium, and the melting and boiling points range from a few kelvin for the light noble gases to several thousand kelvin for refractory solids. Modeling such heavy-tailed targets directly biases the models toward the largest values. Each target was therefore transformed to its natural logarithm before training, and predictions were back-transformed to the original scale for evaluation. A preliminary comparison confirmed that the logarithmic transformation improved the cross-validated accuracy of every target relative to modeling on the original scale.

Second, a minority of descriptor values were missing within the modeling set, most notably electron affinity (about 41% missing) and electronegativity (about 9% missing). Rather than discarding these informative descriptors or the affected elements, missing values were reconstructed by multivariate iterative imputation, in which each descriptor with missing entries is modeled in turn as a function of the others [19]. Crucially, the imputer was fitted only on the training partition within each cross-validation fold and then applied to the held-out partition, so that no information from the test elements leaked into training. The target values were never used during imputation. The strong mutual correlations among periodic descriptors make such imputation reliable, as an element's position in the periodic table largely determines its atomic properties.

3.4 Regression Algorithms

Eight regression algorithms representing four distinct methodological families were evaluated. Two linear baselines were included: ordinary multiple linear regression and ridge regression, the latter adding an L2 penalty to stabilize coefficient estimation [20]. An instance-based k-nearest-neighbors regressor [21] and a kernelized support vector regressor with a radial basis function kernel [22] captured local and nonlinear structure, respectively. Four tree-ensemble methods were

then assessed: the random forest [11], which averages many decorrelated bagged trees [23]; the gradient boosting regressor, which fits trees sequentially to the residuals of the current model [12]; extreme gradient boosting, a regularized and highly optimized boosting implementation [13]; and the extremely randomized trees (Extra Trees) ensemble, which introduces additional randomization by selecting split thresholds at random [14]. For the linear, distance-, and kernel-based models, descriptors were standardized to zero mean and unit variance within the pipeline; the tree ensembles are scale-invariant and required no such scaling.

3.5 Model Evaluation

Predictive performance was estimated by repeated five-fold cross-validation with ten repetitions (fifty train–test splits in total), which provides a robust estimate of generalization on a small dataset and quantifies the variability of the estimate. Within every fold the complete pipeline, comprising imputation, optional scaling, target transformation, and model fitting, was refitted on the training elements only. Three complementary metrics were computed on the original (back-transformed) scale: the coefficient of determination (R^2), the root-mean-square error (RMSE), and the mean absolute error (MAE). The R^2 measures the fraction of variance explained, whereas RMSE and MAE express the typical prediction error in the physical units of each property. Cross-validation as a principle of predictive model assessment follows established statistical practice [24].

3.6 Hyperparameter Optimization

The algorithm with the best mean cross-validated performance across the four targets was selected as the framework base learner and subsequently optimized. The most influential Extra Trees hyperparameters were tuned by grid search, considering the number of features evaluated at each split, the minimum number of samples at a leaf node, and the minimum number of samples required to split a node. Grid search was carried out using an inner five-fold cross-validation loop while keeping the number of trees fixed at 500. The best configuration for each target was then re-evaluated using the same cross-validation protocol employed for model comparison.

3.7 Model Interpretation

The optimized framework was interpreted using two complementary techniques. Shapley additive explanations assign to each descriptor, for each element, a contribution to the predicted value such that the contributions are locally accurate and satisfy consistency axioms derived from cooperative game theory [15,25]. For tree ensembles these values can be computed exactly and efficiently using the TreeSHAP algorithm [16]. The mean absolute SHAP value of a descriptor, averaged over all elements, quantifies its global importance, while the distribution of signed SHAP values reveals the direction and nonlinearity of its effect. As an independent cross-check, permutation importance was also computed by measuring the decrease in model performance when each descriptor was randomly shuffled. Because the models operate on log-transformed targets, SHAP contributions are expressed on the logarithmic scale.

3.8 Implementation

All computations were performed in Python. The regression algorithms, cross-validation, iterative imputation, and permutation importance were implemented with the scikit-learn library [26]; extreme gradient boosting used the XGBoost library [13]; and Shapley explanations used the SHAP library [16]. Numerical computation and data handling relied on NumPy [27] and pandas [28], and figures were produced with Matplotlib [29] and seaborn [30]. Fixed random seeds were used throughout to ensure reproducibility. The overall workflow, from data acquisition and feature engineering through model training, evaluation, selection, and interpretation, is summarized in Figure 1.

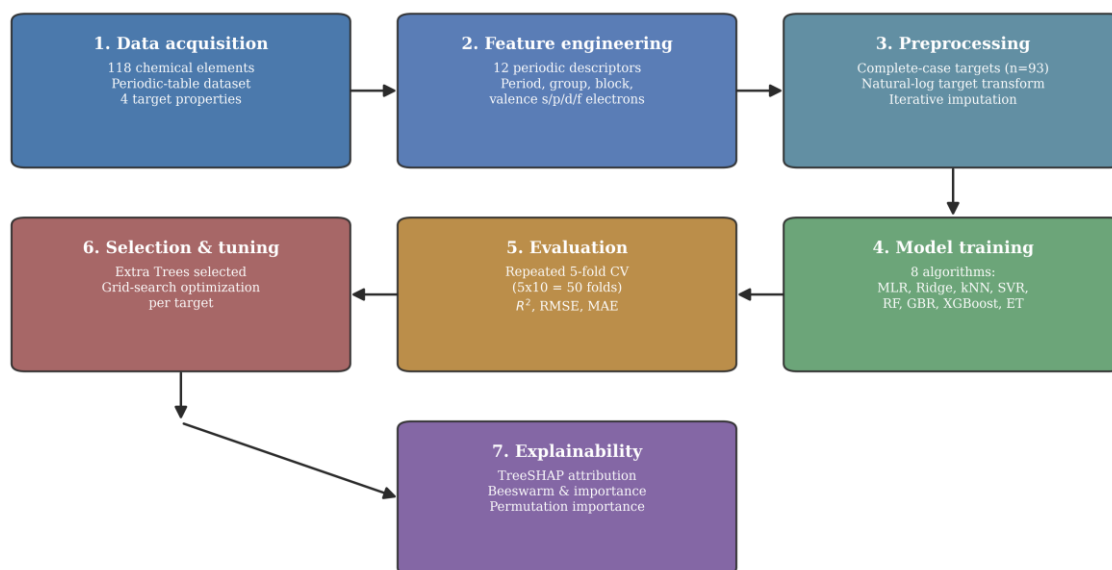


Figure 1. Overview of the explainable multi-target machine learning workflow, from data acquisition and feature engineering through model training, evaluation, selection, and interpretation.

4. Experiments and Results

4.1 Exploratory Analysis of Descriptor–Property Relationships

Table 2 summarizes descriptive statistics for the four target properties across the modeling set of 93 elements. All four target variables display the broad right-skewed distributions expected for the chemical elements (Figure 2, top row). Melting temperatures vary from 0.95 K in helium to 3823 K in carbon; boiling temperatures from 4.22 K to 5869 K; densities vary by more than five orders of magnitude; and atomic radii vary from 120 pm to 343 pm. The logarithmic transform (Figure 2, bottom row) narrows these ranges and produces much more symmetric distributions, which validates its appropriateness for regression analysis.

Table 2. Descriptive statistics of the four target thermophysical properties for the modeling set of 93 elements.

Property	Unit	n	Minimum	Maximum	Mean	SD
Melting point	K	93	0.95	3823	1279.0	921.1
Boiling point	K	93	4.22	5869	2536.2	1588.4
Density	g cm ⁻³	93	9 × 10 ⁻⁵	22.57	7.46	5.88
Atomic radius	pm	93	120	343	206.6	36.3

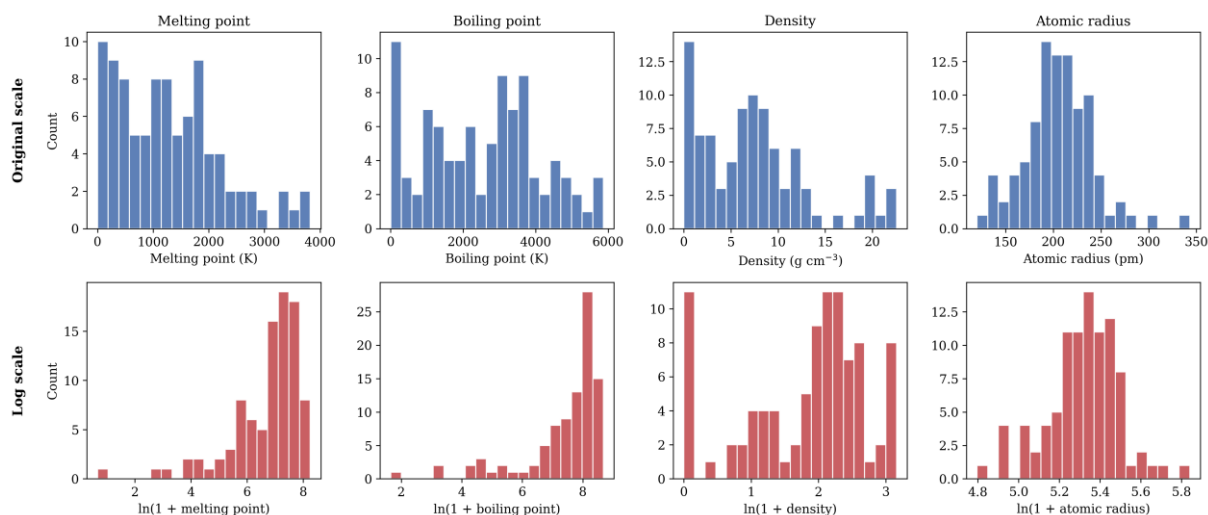


Figure 2. Distributions of the four target properties on their original scale (upper row) and after natural-logarithmic transformation (lower row) for the 93 modeled elements.

The linear correlation structure between descriptors and targets (Figure 3) is informative but limited. Atomic radius correlates most strongly and negatively with ionization energy and electronegativity (Pearson coefficients of about -0.64), consistent with the fact that tightly bound, electronegative atoms are compact. The density of elements is directly related to atomic number, atomic mass, and period, increasing as the mass of the element increases. Melting and boiling points show a very weak linear relationship with any single descriptor; the most significant correlation is observed with the number of p electrons in the valence shell and with ionization energy. The weakness of these linear correlations indicates that the descriptor–property relationships are nonlinear and interactive, and that linear modeling alone is inadequate.

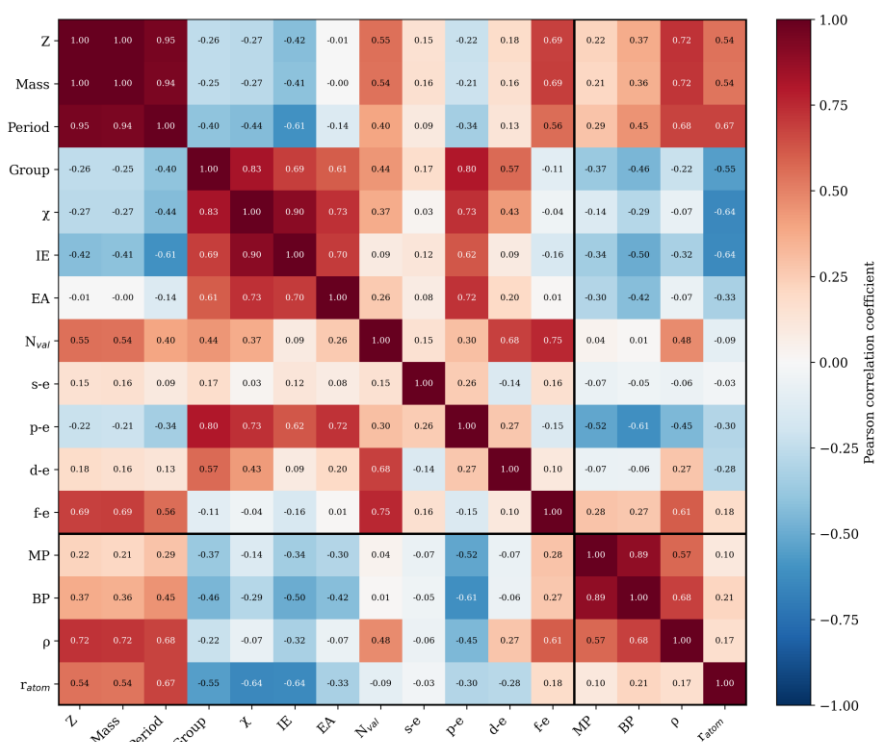


Figure 3. Pearson correlation matrix among the periodic descriptors and the four target properties. The target block is outlined. Symbols: Z, atomic number; χ , electronegativity; IE, ionization energy; EA, electron affinity; MP, melting point; BP, boiling point; ρ , density; r_{atom} , atomic radius.

4.2 Comparison of Regression Algorithms

The eight algorithms were benchmarked against one another using repeated cross-validation under identical conditions (Table 3 and Figure 4). The two linear algorithms performed poorly and even yielded negative coefficients of determination for melting point, boiling point, and density, indicating that they made predictions worse than simply predicting the mean of the data; multiple linear regression on the logarithm of density gave $R^2 = -3.28$. This drastic drop in predictive ability confirms that the relationship between periodic descriptors and thermophysical properties is highly nonlinear and cannot be captured by linear terms. Incorporating nonlinearity through the k-NN and SVR algorithms improved the coefficients of determination to the range 0.42–0.74.

The tree-ensemble models clearly outperformed the other model types for every target. The random forest, gradient boosting, extreme gradient boosting, and Extra Trees models produced coefficients of determination ranging from 0.70 to 0.90. Extreme gradient boosting gave the best performance for density and atomic radius, with R^2 values of 0.90 and 0.75, respectively, while Extra Trees performed best for melting point and boiling point, with R^2 values of 0.72 and 0.76, respectively. Across the four properties, Extra Trees achieved the highest average coefficient of determination (0.78), followed closely by extreme gradient boosting and random forest with averages of 0.77 and 0.75, respectively. The Extra Trees model was therefore chosen as the base learner because of its ability to minimize variance on small datasets.

Table 3. Cross-validated coefficient of determination (R^2) for the eight regression algorithms, reported as mean $\pm$ standard deviation over repeated five-fold cross-validation. The best value for each property is shown in bold.

Algorithm	Melting point	Boiling point	Density	Atomic radius
Multiple linear regression	-0.07 ± 0.54	-0.08 ± 0.72	-3.28 ± 5.30	0.65 ± 0.14
Ridge regression	-0.07 ± 0.48	-0.03 ± 0.63	-2.35 ± 3.71	0.62 ± 0.14
k-nearest neighbors	0.48 ± 0.25	0.42 ± 0.23	0.74 ± 0.15	0.56 ± 0.13
Support vector regression	0.54 ± 0.31	0.55 ± 0.32	0.63 ± 0.33	0.66 ± 0.12
Random forest	0.70 ± 0.27	0.72 ± 0.22	0.87 ± 0.08	0.71 ± 0.11
Gradient boosting	0.68 ± 0.28	0.70 ± 0.21	0.89 ± 0.10	0.72 ± 0.12
Extreme gradient boosting	0.70 ± 0.29	0.71 ± 0.20	0.90 ± 0.07	0.75 ± 0.10
Extra Trees	0.72 ± 0.29	0.76 ± 0.22	0.90 ± 0.09	0.75 ± 0.10

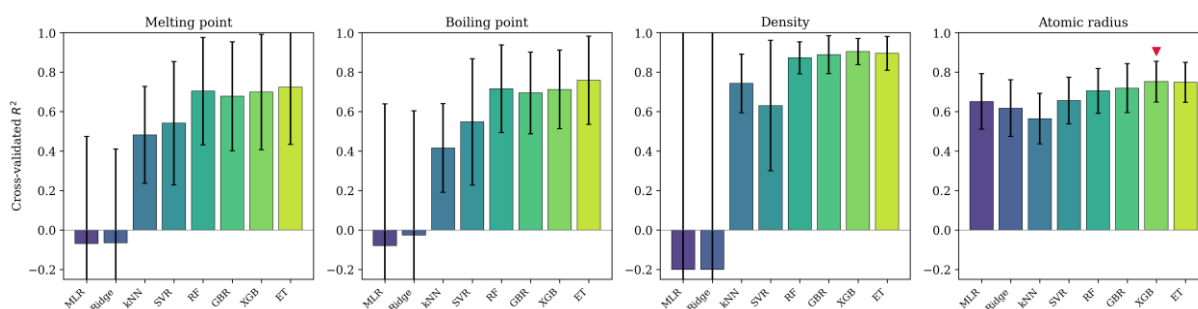


Figure 4. Cross-validated coefficient of determination for the eight regression algorithms across the four properties. Bars show the mean and error bars the standard deviation over repeated five-fold cross-validation; the best model for each property is marked. Abbreviations: MLR, multiple linear regression; kNN, k-nearest neighbors; SVR, support vector regression; RF, random forest; GBR, gradient boosting; XGB, extreme gradient boosting; ET, Extra Trees.

4.3 Performance of the Optimized Framework

Following hyperparameter tuning, the framework was assessed again using repeated cross-validation (Table 4). It accounted for 72% of the variance in melting point ($RMSE \approx 434$ K), 76% of the variance in boiling point ($RMSE \approx 683$ K), 88% of the variance in density ($RMSE \approx 1.84$ g cm $^{-3}$), and 74% of the variance in atomic radius ($RMSE \approx 16.7$ pm). Given the chemical complexity of the periodic table and the relatively small number of training samples, these figures represent a good fit, especially for density, which shows near-quantitative behavior. The associated errors are also small: the mean absolute errors correspond to 12–19% of the mean values of the respective properties. In the observed-versus-

predicted diagrams generated from out-of-fold predictions (Figure 5), all points lie close to and evenly around the ideal line, with no curvature or systematic bias.

Table 4. Performance of the optimized Extra Trees framework, reported as mean $\pm$ standard deviation over repeated five-fold cross-validation (RMSE and MAE in the physical units of each property).

Property	Unit	R ² (mean $\pm$ SD)	RMSE (mean $\pm$ SD)	MAE
Melting point	K	0.72 $\pm$ 0.28	434 $\pm$ 232	271
Boiling point	K	0.76 $\pm$ 0.19	683 $\pm$ 225	474
Density	g cm ⁻³	0.88 $\pm$ 0.08	1.84 $\pm$ 0.73	1.13
Atomic radius	pm	0.74 $\pm$ 0.10	16.7 $\pm$ 3.7	12.7

The optimized model also demonstrated the stability of the base estimator: the best-performing parameter sets for all four targets used all descriptors at each node with minimal leaf regularization, a configuration very close to the default Extra Trees settings. Tuning therefore brought only marginal improvements over the default, indicating that the strong performance arose from the ensemble learning method and the descriptor set rather than from the tuned parameters.

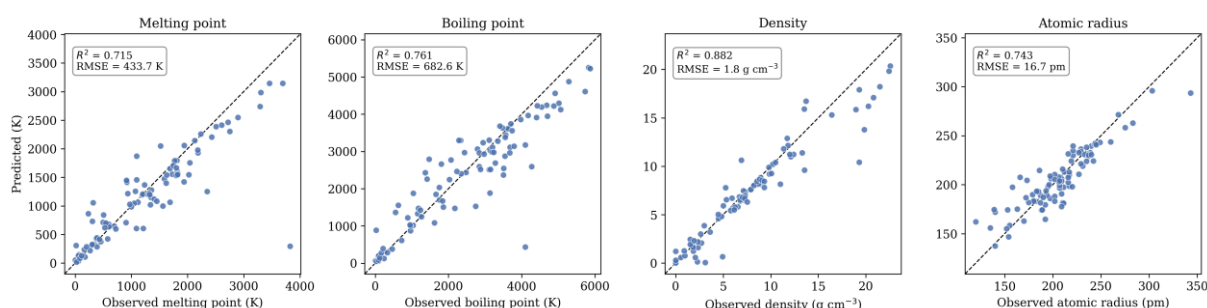


Figure 5. Predicted versus observed values for the optimized framework. Predictions are out-of-fold estimates averaged over ten resampling repeats; the annotated statistics are the repeated five-fold cross-validation averages of Table 4. The dashed line denotes perfect prediction.

5. Discussion

5.1 Model Interpretation and Chemical Insight

The central aim of an explainable framework was to reveal why the model makes its predictions. The global importance of descriptors derived from SHAP are shown in Figure 6, and the leading descriptors for each property are listed in Table 5. A clear and chemically coherent picture emerges. Ionization energy is the single most important descriptor for melting point, boiling point, and density, with the largest mean absolute SHAP value in each case. This is physically reasonable: the first ionization energy reflects how tightly the valence electrons are bound to the nucleus, which correlates with the strength of metallic and covalent bonding, and hence with the cohesive energy that must be overcome to melt or vaporize an element, and with the compactness that governs density. The beeswarm plots confirm the expected directions: high ionization energy pushes the predicted density upward, whereas for atomic radius high ionization energy pushes the prediction downward, since strongly bound electrons yield smaller atoms.

The group number and electronegativity are consistently the next most influential descriptors for the melting and boiling points, together with the number of valence p electrons. This reflects the systematic variation of bonding character across a period, from strongly bonded metals and network solids on the left and center of the table to weakly bonded molecular and monatomic species on the right. For density, the period and electronegativity join ionization energy as the dominant descriptors, encoding the position of an element within the table and thus its mass-to-volume ratio. For atomic radius, the ordering changes: the period becomes the most important descriptor, in direct agreement with the fundamental periodic trend that atomic radius increases with the principal quantum number of the valence shell down a group, followed by ionization energy, group, and electronegativity. The overall SHAP magnitudes for atomic radius are smaller than for the other properties because atomic radius varies over a comparatively narrow range on the logarithmic scale.

It is noteworthy that descriptors expressing similar underlying information, such as atomic number, atomic mass, and period, tend to share importance among themselves, a well-understood consequence of feature correlation in additive attribution. The broad agreement between the SHAP rankings and the independently computed permutation importances, both identifying ionization energy, group, electronegativity, and period as the key descriptors, increases confidence that the framework has learned genuine periodic relationships rather than spurious correlations. In this sense the explainability analysis serves as a scientific validation of the model.

Table 5. The five most important descriptors for each property, ranked by mean absolute SHAP value (given in parentheses, on the natural-logarithmic scale).

Rank	Melting point	Boiling point	Density	Atomic radius
1	Ionization energy (0.34)	Ionization energy (0.42)	Ionization energy (0.69)	Period (0.043)
2	Group (0.23)	Group (0.17)	Period (0.27)	Ionization energy (0.024)
3	Electronegativity (0.13)	Valence p-electrons (0.13)	Electronegativity (0.26)	Group (0.023)
4	Valence d-electrons (0.11)	Electronegativity (0.13)	Valence p-electrons (0.19)	Electronegativity (0.019)
5	Valence p-electrons (0.10)	Valence d-electrons (0.10)	Group (0.19)	Atomic mass (0.019)

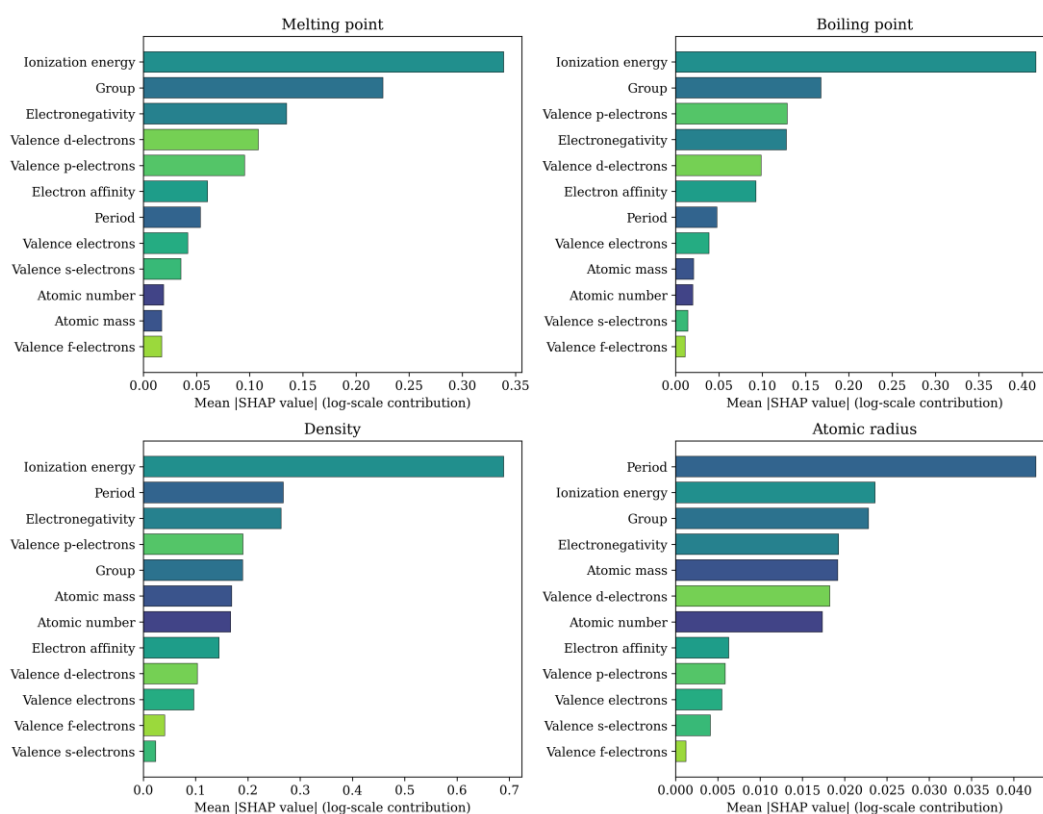


Figure 6. Global descriptor importance for each property, quantified by the mean absolute SHAP value (contribution on the natural-logarithmic scale).

5.2 Error Analysis

Examining the elements with the largest prediction errors provides further insight and, reassuringly, the outliers are precisely those that chemistry would flag as anomalous. For melting point, the most severely underestimated element is carbon, whose measured value of 3823 K vastly exceeds the framework's estimate because carbon, in its diamond and graphite forms, possesses a three-dimensional covalent network with exceptionally strong bonding that is not encoded by single-atom descriptors; boron and silicon, which also form network solids, are underestimated for the same reason.

Conversely, phosphorus is overestimated because its low melting point arises from a discrete molecular structure. For density, gold and mercury are underestimated, reflecting the pronounced relativistic contraction of the 6s orbital in the heaviest elements, an effect absent from the descriptor set; mercury is further exceptional as the only metal that is liquid at room temperature. For atomic radius, the alkali metal cesium, which has the largest radius in the dataset, lies at the extreme of the descriptor space and is underestimated.

These residual errors are not failures of the methodology so much as manifestations of chemistry that transcends the periodic descriptors, namely allotropy, molecular versus network bonding, and relativistic effects. The fact that the interpretable framework isolates exactly these physically meaningful exceptions underscores the value of coupling accurate prediction with transparent explanation, and it indicates where richer, structure-aware descriptors would be required to improve accuracy further.

5.3 Limitations and Outlook

Several limitations should be acknowledged. The periodic table provides only about one hundred fully characterized elements, so the training set is inherently small, which limits the achievable accuracy and widens the uncertainty of the performance estimates, as reflected in the standard deviations in Table 4. The descriptors are single-atom quantities and therefore cannot capture crystal structure, allotropy, or relativistic effects, which set a ceiling on the accuracy attainable for structure-sensitive properties such as melting point. The imputation of the substantial fraction of missing electron-affinity values, although performed without leakage, introduces some additional uncertainty. Finally, the framework is interpolative and should be applied with caution to the superheavy elements that lie outside the range of the training data.

Future work could extend the framework in several directions: incorporating structural or cohesive-energy descriptors to better model network solids; adopting genuinely joint multi-output learning that exploits correlations among the four properties; enlarging the dataset with compounds and alloys; and quantifying predictive uncertainty on a per-element basis. Nevertheless, within its intended scope the framework already provides accurate, transparent, and essentially instantaneous estimates of four fundamental elemental properties, and the explainability analysis confirms that its reasoning aligns with established chemical principles.

6. Conclusions and Future Work

An explainable multi-target machine learning framework was developed to predict the melting point, boiling point, density, and atomic radius of the chemical elements from a compact set of twelve periodic-table descriptors, several of which were engineered from electron configurations. A systematic comparison of eight algorithms showed that tree-based ensembles vastly outperform linear models on this nonlinear problem, and the extremely randomized trees method delivered the best mean performance across the four properties. After optimization and leakage-free handling of missing descriptors, the framework explained between 72% and 88% of the variance in the four properties with low absolute errors, using only tabulated atomic information and negligible computation. Shapley and permutation analyses revealed chemically coherent descriptor–property relationships, identifying ionization energy as the principal driver of melting point, boiling point, and density, and the period as the principal driver of atomic radius, while the largest residual errors were traced to elements with anomalous bonding or relativistic effects.

By uniting predictive accuracy with transparent, chemically meaningful explanations, the proposed framework offers a practical and trustworthy complement to experimental measurement and first-principles calculation for the rapid estimation of fundamental elemental properties. Future efforts should focus on incorporating structure-aware and cohesive-energy descriptors to capture network solids and relativistic effects, adopting fully joint multi-output learning, extending the dataset to compounds and alloys, and providing per-element uncertainty estimates to further broaden the applicability of the framework.

7. Acknowledgments

Financial support from the Arab American University (AAUP) is gratefully acknowledged. The authors thank the Faculty of Sciences for providing computational facilities and technical assistance.

8. Declarations

Funding

This work was supported by the Arab American University (AAUP).

Conflict of Interest / Competing Interests

The authors declare no conflict of interest.

Ethics Approval and Consent to Participate

Not applicable. This study involved only publicly available data and no human participants.

Consent for Publication

All authors have read and approved the final manuscript and consent to its publication.

Data Availability

The dataset analyzed in this study is derived from publicly available periodic-table data (<https://pubchem.ncbi.nlm.nih.gov/rest/pug/periodictable/CSV>). The engineered features and analysis code are available from the corresponding author upon reasonable request.

Materials Availability

Not applicable.

Code Availability

The analysis code is available from the corresponding author upon reasonable request.

Authors' Contributions

F. Abujaber: Conceptualization, Supervision, Methodology, Writing–Review and Editing, Funding Acquisition. F. Siouri: Methodology, Software, Formal Analysis, Visualization, Writing–Original Draft. All authors have read and agreed to the published version of the manuscript.

References

- [1] Greenwood NN, Earnshaw A. Chemistry of the Elements. 2nd ed. Oxford, UK: Butterworth-Heinemann; 1997. doi:10.1016/C2009-0-30414-6
- [2] Kohn W, Sham LJ. Self-consistent equations including exchange and correlation effects. Phys Rev. 1965;140(4A):A1133–A1138. doi:10.1103/PhysRev.140.A1133
- [3] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature. 2018;559(7715):547–555. doi:10.1038/s41586-018-0337-2
- [4] Ward L, Agrawal A, Choudhary A, Wolverton C. A general-purpose machine learning framework for predicting properties of inorganic materials. npj Comput Mater. 2016;2:16028. doi:10.1038/npjcompumats.2016.28
- [5] Ramprasad R, Batra R, Pilia G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Comput Mater. 2017;3:54. doi:10.1038/s41524-017-0056-5
- [6] Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater. 2019;5:83. doi:10.1038/s41524-019-0221-0
- [7] Rupp M, Tkatchenko A, Müller KR, von Lilienfeld OA. Fast and accurate modeling of molecular atomization energies with machine learning. Phys Rev Lett. 2012;108(5):058301. doi:10.1103/PhysRevLett.108.058301
- [8] Pilia G, Wang C, Jiang X, Rajasekaran S, Ramprasad R. Accelerating materials property predictions using machine learning. Sci Rep. 2013;3:2810. doi:10.1038/srep02810

- [9] Seko A, Maekawa T, Tsuda K, Tanaka I. Machine learning with systematic density-functional theory calculations: application to melting temperatures of single- and binary-component solids. *Phys Rev B*. 2014;89(5):054303. doi:10.1103/PhysRevB.89.054303
- [10] Zhuo Y, Mansouri Tehrani A, Brgoch J. Predicting the band gaps of inorganic solids by machine learning. *J Phys Chem Lett*. 2018;9(7):1668–1673. doi:10.1021/acs.jpcllett.8b00124
- [11] Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32. doi:10.1023/A:1010933404324
- [12] Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29(5):1189–1232. doi:10.1214/aos/1013203451
- [13] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA; 2016. p. 785–794. doi:10.1145/2939672.2939785
- [14] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63(1):3–42. doi:10.1007/s10994-006-6226-1
- [15] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems 30*. Long Beach, CA, USA; 2017. p. 4765–4774.
- [16] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9
- [17] Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA; 2016. p. 1135–1144. doi:10.1145/2939672.2939778
- [18] Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2nd ed. Munich, Germany; 2022.
- [19] van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1–67. doi:10.18637/jss.v045.i03
- [20] Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*. 1970;12(1):55–67. doi:10.1080/00401706.1970.10488634
- [21] Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13(1):21–27. doi:10.1109/TIT.1967.1053964
- [22] Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–297. doi:10.1007/BF00994018
- [23] Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123–140. doi:10.1007/BF00058655
- [24] Stone M. Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc Series B*. 1974;36(2):111–147. doi:10.1111/j.2517-6161.1974.tb00994.x
- [25] Shapley LS. A value for n-person games. In: Kuhn HW, Tucker AW, editors. *Contributions to the Theory of Games II*. Princeton, NJ, USA: Princeton University Press; 1953. p. 307–317.
- [26] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–2830.
- [27] Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, et al. Array programming with NumPy. *Nature*. 2020;585(7825):357–362. doi:10.1038/s41586-020-2649-2
- [28] McKinney W. Data structures for statistical computing in Python. In: *Proceedings of the 9th Python in Science Conference*. Austin, TX, USA; 2010. p. 56–61. doi:10.25080/Majors-92bf1922-00a
- [29] Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng*. 2007;9(3):90–95. doi:10.1109/MCSE.2007.55

-
- [30] Waskom ML. seaborn: statistical data visualization. J Open Source Softw. 2021;6(60):3021. doi:10.21105/joss.03021