JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

MACHINE LEARNING-ASSISTED SCREENING OF THERMODYNAMICALLY STABLE LITHIUM-BASED BATTERY MATERIALS FOR SUSTAINABLE ENERGY STORAGE AND SOCIETAL BENEFITS

Jahanzaib[*1], Youshaib Masih Gil[2], Malaika Shahbaz[3], Usama Farhan Safdar[4], Fazal Ur Rehman Sheikh[5]

[*1,2,3,4,5]*Department of Physics and Mathematics Superior University, Lahore, Pakistan*

[*1]Jahanzaibfayyaz527@gmail.com

DOI: https://doi.org/10.5281/zenodo.23391288

Keywords

Article History
Received: *25 June 2026*
Accepted: *16 September 2026*
Published: *13 October 2026*

Copyright @Author
Corresponding Author: * Jahanzaib

Abstract

Efficient, sustainable energy storage hinges on our ability to identify stable lithium-based materials. Standard first-principles calculations take an immense amount of time and computing power, though, when applied across broad materials databases. In this study, we built a machine-learning-assisted workflow to screen for viable candidates, using predicted per-atom formation energies across lithium-bearing and related inorganic compounds. Our initial dataset held 86,988 raw records across 11 descriptors, but filtering out duplicates brought that down to 21,747 unique materials. From chemical, structural, energetic, and electronic descriptors, we built an input matrix totaling 92 features. We evaluated five machine-learning models Random Forest, XGBoost, LightGBM, CatBoost, and an Artificial Neural Network (ANN) using mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2). On the initial test set, Random Forest achieved the lowest MAE (0.1065), while the Artificial Neural Network achieved the lowest RMSE (0.2041) and the highest R^2 (0.9577). According to the SHAP analysis, energy_per_atom influenced the predictions dominantly, while dropping it shifted the main weight onto element_O, band_gap, element_F, density, and volume. When evaluated under a stricter five-fold Group K-Fold split, R^2 held at 0.9547 ± 0.0065 with energy_per_atom included, but fell to 0.8366 ± 0.0074 without it. The out-of-fold predictions yielded 16 distinct candidate compounds. Cross-referencing these candidates against Materials Project energy-above-hull figures showed that 8 of them (50%) are thermodynamically stable, four lie above the convex hull, and four have no E_hull records at all. While DFT calculations and convex-hull analysis remain necessary to confirm thermodynamic stability across all 16 candidates, these findings point to

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

the real utility of interpretable machine learning for screening energetics quickly.

1.1 INTRODUCTION

As renewable energy technologies, electric vehicles, and portable electronics have expanded quickly, the need for efficient, reliable, and sustainable energy-storage systems has grown significantly. Because they pack high energy density, last through many cycles, and fit into so many different uses, rechargeable lithium-ion batteries are now among the most widely adopted energy-storage options[1]. Pushing battery performance, safety, operating life, and sustainability any further, however, calls for advanced electrode and electrolyte materials that possess the right structural, electronic, and thermodynamic traits. Finding those materials is still difficult simply because the pool of candidate chemical compositions and crystal structures for energy-storage use is so vast[2].

Assessing whether a material is likely to form a stable compound relies heavily on its energetic profile. Formation energy shows how energetically favorable a compound is compared to its constituent elements, whereas the energy above the convex hull gives a more direct gauge of thermodynamic stability against competing phases [3]. Researchers now rely heavily on Density Functional Theory (DFT) to calculate these values, a method that has paved the way for massive computational materials databases. First-principles calculations demand substantial compute time and power, though, when applied across vast pools of candidate materials. Finding new candidates quickly across wide material spaces consequently calls for far more efficient computational techniques [4].

Machine learning (ML) has stepped in as an effective data-driven solution to this bottleneck [5]. Once trained on property data and descriptors drawn from experiments or calculations, ML models can estimate values for vast sets of materials in little time. Work across multiple studies shows that machine learning accurately estimates formation energies from elemental, compositional, and structural descriptors. Across large-scale materials datasets, these studies show that ML models achieve useful accuracy and can effectively replace costly first-principles calculations as surrogate tools [6].

Researchers have also broadened the use of machine learning beyond predicting formation energies, applying it to crystal-structure exploration and materials discovery[7]. To forecast how materials behave energetically and predict their properties, researchers have turned to a variety of algorithms, among them decision trees, gradient boosting, and artificial neural networks. Before committing to heavy first-principles calculations on chosen candidates, researchers can use ML models to quickly assess vast numbers of materials and flag those whose predicted properties look promising. This strategy lowers the computational burden of screening materials at high throughput without sacrificing valuable data on how they behave energetically[8].

Battery research stands to benefit from machine learning in particular, given the massive chemical search space required to find compounds that hit the right electrochemical, structural, and energetic targets. Researchers in earlier work have turned to neural networks and related machine-learning methods to predict battery electrode traits, including average voltage and the volume shifts that occur during charge and discharge cycles. What those studies showed was that ML can filter through candidate materials quickly, pointing researchers toward the most viable compounds for deeper quantum-mechanical analysis. Pairing computational materials databases with machine learning shows real promise for speeding up how quickly we find advanced materials for energy storage [9].

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Machine learning has moved past single-property forecasts as well, with researchers testing its capacity to gauge thermodynamic stability [10]. Using descriptors tied to elements and chemical composition, researchers have built ML models that estimate energy above the convex hull across perovskite oxides and other classes of materials. Comparable methods have been brought to bear on two-dimensional materials, grouping candidate compounds by how stable they are expected to be thermodynamically. Work like this shows how useful machine learning can be when screening material properties tied to stability at scale [11].

Pairing explainable artificial intelligence (XAI) with machine learning gives materials discovery an extra edge. With approaches like Shapley Additive Explanations (SHAP), researchers can test how much each material descriptor actually steers the ML output. Running a SHAP analysis helps pinpoint key chemical, electronic, and physical descriptors, which in turn clarifies the relationships the ML model has picked up. Earlier work has already shown that explainable machine learning helps make sense of the factors tied to predicted phase stability and other material properties [12].

All the same, calculating formation energy is not the same thing as directly establishing thermodynamic stability, and that boundary matters. A compound is not proven thermodynamically stable against all competing phases simply because its formation energy is known, even if ML models can predict that value with high accuracy[13]. Evaluating stability directly usually takes data on the energy above the convex hull, or comparable phase-equilibrium metrics. Predicting formation energies with ML should thus be treated as an energetic screening filter, not a standalone substitute for direct thermodynamic stability calculations[14].

With that distinction in mind, this work uses machine learning to screen lithium-based battery materials by evaluating their predicted formation energies alongside related energetic features tied to stability [15]. We draw on a large materials dataset, treating formation energy per atom as our main target. We test Random Forest (RF), XGBoost, LightGBM, CatBoost, and Artificial Neural Network (ANN) architectures against physical, electronic, and compositional features, including formula-derived elemental composition, element counts, transition-metal presence, energy per atom, band gap, semiconductor classification, volume, and density. We evaluate how well the models perform through MAE, RMSE, and R^2 [16].

SHAP analysis is then applied to pinpoint which features drive the predicted formation energies most heavily. To test how much the energy-per-atom descriptor matters, we compare how the models perform with that feature included versus left out[17]. The trained models are then set to screening lithium-bearing compounds, picking out those with promising predicted formation energies. These ML-screened candidates should be viewed as solid leads for further study rather than compounds confirmed to be thermodynamically stable. Taken as a whole, this work sets out to show how an interpretable machine-learning pipeline can speed up the discovery and analysis of promising lithium compounds for future battery work[18].

Meeting the rising demand for renewable power, electric vehicles, and portable electronics hinges more and more on developing energy-storage systems that are both efficient and sustainable. Lithium-based batteries carry much of the weight in these sectors, yet finding better materials for them means hunting down chemically and thermodynamically stable candidates within an enormous compositional space[19]. Screening those candidates through standard laboratory experiments can run up high costs in both hours and supplies. Here, machine learning offers a way to speed up discovery by parsing large datasets fast and surfacing the most promising candidates for follow-up testing. By flagging thermodynamically stable lithium compounds early, data-driven workflows like the one shown

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

here help streamline materials research and ultimately aid the push toward sustainable energy storage.

Literature Review

Predicting material formation energy through machine learning has proven effective for speeding up the discovery of new materials. Traditional descriptors (the tolerance factor for ABX_3 perovskites being one example) do give useful structural stability data, but they fail to capture the full complexity of atomic interactions [20]. Trained on formation-energy datasets, machine learning models pick up the intricate links between material descriptors and formation energy; earlier work on ABX_3 perovskites, for instance, showed strong predictive accuracy with an R^2 of 0.928 and an RMSE of 0.301 eV/atom. As broad computational materials databases become more widely available, using machine learning to predict formation energy has grown more practical, lowering the computational expense of screening candidates with large-scale DFT[21]. Specialized tools, including MatLearn, were created to navigate chemical composition spaces and flag compounds that might be thermodynamically accessible [22].

Earlier work indicates that pairing structural details with compositional descriptors leads to better formation-energy predictions. When trained on roughly 435,000 formation energies drawn from the Open Quantum Materials Database (OQMD), models that paired elemental properties with Voronoi-tessellation structural descriptors reached mean absolute errors of about 80 meV/atom [23]. Neural networks have likewise shown high accuracy in these predictions. Tested across roughly 21,800 OQMD compounds, the Structure-Informed Prediction of Formation Energy using Neural Networks (SIPFENN) framework produced mean absolute errors between 28 and 42 meV/atom, varying by the specific prediction task. Together, this work shows how machine learning can either replace or complement DFT to screen materials efficiently at scale [24].

Lithium compounds and related battery chemistries have also been targeted directly with machine learning [25]. For binary Li–Ge systems, kernel ridge regression and support vector regression using partial radial distribution functions kept prediction errors near 20 meV/atom while predicting relaxed formation energies directly from unrelaxed structures, pointing to ML's promise for cutting back on costly relaxation runs [26]. In other work, researchers turned to deep neural networks to estimate qualities like volume changes and average voltage as battery electrodes cycle, using ML to pre-screen prospective battery materials rapidly[27]. Together, these uses reflect how machine learning is increasingly relied upon to speed up how candidate materials for energy storage are spotted and tested [28].

Researchers looking for new materials via machine learning have also turned their attention to thermodynamic stability. In work on perovskite oxides, researchers evaluated stability directly through energy above the convex hull (E_{hull}); their kernel ridge regression model showed high predictive accuracy, yielding cross-validation errors of roughly 28.5 meV/atom for RMSE and about 16.7 meV/atom for MAE [29]. In a comparable manner, researchers have sorted 2D materials by their thermodynamic stability through machine learning fed with structural and compositional descriptors, turning up over a thousand candidate compounds worth examining further [30]. Following that, several of these candidates were tested for photoelectrocatalytic water splitting, which illustrates how screening for stability with ML can steer materials science toward concrete energy applications [31].

Computational materials screening has grown even more useful thanks to interpretable machine learning[32]. Researchers working on lithium-metal and all-solid-state batteries paired machine learning with high-throughput screening across more than 15,000 Li-bearing materials, looking closely at migration barriers alongside electrochemical stability windows[33].Gradient-

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

boosting models predicted oxidation potential, reduction potential, and migration barriers with high accuracy, and researchers used SHAP alongside permutation feature-importance analyses to pinpoint which material properties drove those predictions most [34]. Work of this kind shows that pairing accurate predictions with interpretable models sheds extra light on what drives material behavior, helping guide materials discovery with greater focus.

Even with this progress, real limitations persist. Predicting formation energy with high accuracy does not mean a model will reliably predict thermodynamic stability[35].In an analysis of 85,014 unique chemical compositions from the Materials Project, compositional models predicted formation energy reasonably well, yet they struggled to predict stability directly in chemically sparse regions where structural data provided extra predictive value . Because of this, predicting formation energy cannot be treated as the same task as predicting thermodynamic stability directly[36].

Research Gap

Earlier studies have shown that machine learning works well for estimating formation energy, evaluating thermodynamic stability, and screening battery materials. Most existing work, though, targets either sweeping inorganic datasets or very narrow material families, leaving lithium-bearing compounds rarely studied within a single, interpretable framework. Direct checks on thermodynamic stability are also constrained because certain datasets leave out energy-above-hull and complete convex-hull data.

To fill these gaps, we apply machine learning to a large dataset of lithium-bearing and related inorganic compounds, predicting per-atom formation energy with Random Forest, XGBoost, LightGBM, CatBoost, and Artificial Neural Network models. To pinpoint which descriptors drive the predictions most heavily, we run feature-importance and SHAP analyses, while testing models with and without energy_per_atom to trace that specific descriptor's impact. Grouping the validation folds by material identifier limits the influence of duplicate entries, yielding a more dependable test of how well the models generalize. Lastly, out-of-fold predictions help flag lithium-bearing candidates worth pursuing in subsequent studies. Without energy-above-hull figures or full convex hulls in the dataset, these candidates should be treated as promising computational leads instead of proven stable materials.

3. Methodology

3.1 Dataset and processing

To build our machine-learning-assisted screening framework for lithium-containing materials, we drew on a large dataset of chemical, structural, electronic, and energetic descriptors. Initially, the dataset contained 86,988 entries spanning 11 variables: material ID, chemical formula, number of elements, transition-metal presence, formation energy per atom, energy per atom, band gap, semiconductor classification, density, volume, and elemental composition. We sourced these records from the Materials Project.

Our chosen target variable was formation energy per atom (formation_energy_per_atom). An initial check showed that the data contained no missing values at all. Multiple entries had identical material identifiers, so we deduplicated on `material_id` and ended up with 21,747 unique materials.

We identified which materials contained lithium straight from their chemical formulas. Using one-hot encoding, we converted the composition data in the `elements` field into binary elemental descriptors. In total, 85 different elements showed up across the dataset. On top of the elemental descriptors, we incorporated seven numerical and categorical features: element count, presence of transition metals, energy per atom, band gap, semiconductor classification, density, and volume. Together, these additions brought the final feature matrix to 92 input variables.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Feature Engineering alongside the Roles of Material ID and Chemical Formula

We relied on `material_id` to identify materials, drop duplicates, and split groups for validation, but we did not feed it to the model as a predictor. In the GroupKFold validation, using `material_id` as the grouping factor guaranteed that no material group ended up partly in training and partly in validation.

The machine-learning models were never fed the chemical formula as raw text; it served only as the source for the elemental composition data. Instead, elemental composition was translated into numerical terms through 85 binary one-hot encoded descriptors. This setup gave the models access to composition data without forcing them to work straight from raw formula text.

3.2 Selection of Machine Learning Models

Five regression algorithms were evaluated for predicting formation energy per atom: Random Forest (RF), XGBoost, LightGBM, CatBoost, and Artificial Neural Network (ANN).

For the initial random train-test evaluation, setting "random_state = 42", we split the data into an 80% training set and a 20% test set. This gave 17,397 samples in the training split and 4,350 in the test split.

The Random Forest model was configured with 200 estimators. For XGBoost, LightGBM, and CatBoost, training parameters were set to 300 estimators/boosting iterations, a learning rate of 0.05, and a maximum tree depth of 6. The Artificial Neural Network architecture consisted of two fully connected hidden layers: the first held 128 neurons, and the second held 64. We used the ReLU activation function together with a dropout rate of 0.2. Using a learning rate of 0.001, the Adam optimizer updated the network across 50 epochs with batches of 32. `StandardScaler` was used to standardize the input features fed into the ANN.

3.3 Model Evaluation

We evaluated how well the models predicted the target using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2).

MAE, MSE, and RMSE quantified error magnitude, whereas R^2 tracked the share of variance in formation energy accounted for by each model. Plotting actual values directly against predicted ones gave us a clear visual check on how closely the two sets of formation energies lined up.

3.4 Feature Importance and SHAP Analysis

To determine which descriptors drove predictions most heavily, we began by extracting feature importances directly from the trained Random Forest model. The top 20 features were then ranked according to their Random Forest importance scores and plotted.

Next, to interpret individual predictions without relying on model architecture, we ran a SHAP (SHapley Additive exPlanations) analysis on the trained Random Forest model. SHAP values were calculated for every one of the 4,350 test-set samples. Each feature's overall contribution was captured by its mean absolute SHAP value, while SHAP summary plots illustrated the magnitude and directional influence of each descriptor on the predictions.

3.5 Impact of Effect of Energy per Atom

Because `energy_per_atom` relates directly to the energetic profile of a material, we performed an ablation study to track what happens to prediction accuracy when it is left out. Removing `energy_ per_atom` reduced the feature matrix from 92 input descriptors to 91.

A Random Forest model was then trained using this reduced feature set. We then compared its MAE, RMSE, and R^2 directly with the baseline from the full-feature model. This was done to find out whether the model, stripped of that one input, could still predict formation energy

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

accurately from the remaining chemical, structural, electronic, and physical descriptors alone.

3.6 Group-Based Validation

We grouped a five-fold GroupKFold cross-validation by `material_id` so we could prevent leakage between related records and test generalization under harder terms. This kept every material group strictly inside a single fold, so no group was ever split across training and validation at the same time.

Across all five folds, we evaluated the Random Forest models and recorded their MAE, RMSE, and R^2. We calculated and reported the mean alongside the standard deviation across the five folds for each metric. This group-based validation was performed twice: once using the full 92-feature matrix, and a second time on the 91 features left after removing `energy_per_atom.

3.7 Using Machine learning to Screen Candidates

To screen the final candidates, we performed a five-fold GroupKFold cross-validation using the Random Forest model and gathered out-of-fold (OOF) predictions for every sample. Because training runs never included samples from a model's own material group, relying on these OOF predictions reduced bias in our screening assessment.

We then evaluated the lithium-bearing materials by pairing their OOF predictions with their corresponding prediction errors. From this filtered pool we selected candidate materials, making sure no chemical composition appeared more than once. This selection left us with 16 unique lithium-containing candidate materials to evaluate in the next phase.

For each candidate, we recorded the material ID, chemical formula, actual formation energy, OOF-predicted formation energy, and absolute prediction error. Neither experiment nor computation had yet confirmed these materials as thermodynamically stable; at this stage, they were only ML-screened candidates.

3.8 Evaluating Thermodynamic Stability

We evaluated whether the ML-screened candidates were thermodynamically stable by retrieving their energy-above-hull (E_hull) values recorded in the Materials Project. Energy above hull simply reflects how far a material sits above the thermodynamic convex hull in energy terms. Candidates with E_hull = 0 eV/atom` were treated as lying right on the convex hull and marked stable according to current Materials Project records, whereas anything with `E_hull` over zero counted as above hull.

If matching stability entries could not be found in the Materials Project, we left the candidate without a firm thermodynamic label. They remained listed as unverified instead, set aside for later evaluation with standalone density functional theory calculations and convex-hull analysis.

Results and Discussion

Distribution of Formation Energy per Atom

Figure 1 shows the distribution of energy formations per atom among the unique 21,747 materials that are present in the dataset. Most materials are clustered around the main peak between -1 and 0 eV/atom, with only a thin scatter to higher positive values. The curve is partially wrapped and folded into the loose right tail rather than spreading uniformly. While examples of high energy can be very rare, most data sets form energy-like compounds. This type of skew has real consequences for machine learning as the training set makes the model much more midrange examples than marginal. As a result, densest regions forecasts are expected to stay fairly reliable. But less common entries are more susceptible to significant errors.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

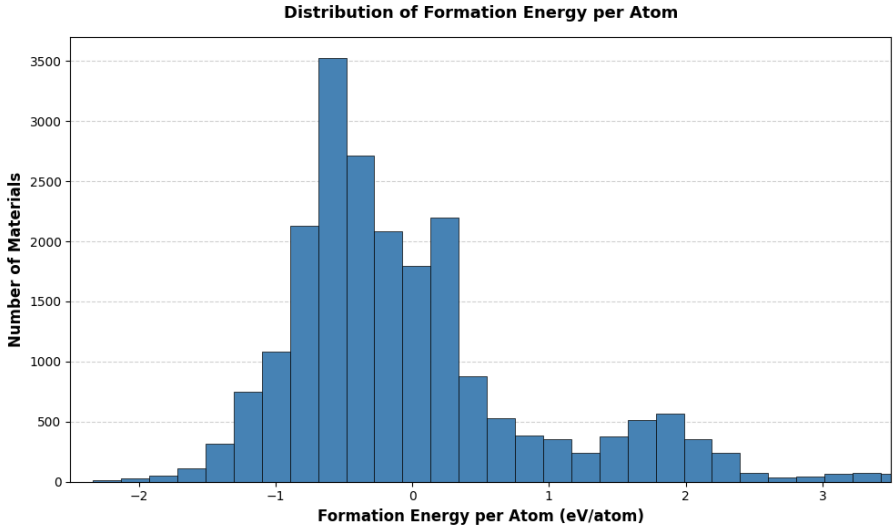


Figure 1. Distribution of Formation Energy per Atom

Model Performance Comparison

Table 1 Test-set performance of five machine learning models trained on the Lithium Battery Materials dataset (with energy_per_atom included as a feature).

Model	MAE	RMSE	R^2
Random Forest	0.1065	0.2093	0.9555
XGBoost	0.1455	0.2298	0.9464
LightGBM	0.1498	0.2383	0.9424
CatBoost	0.1671	0.2593	0.9317
ANN	0.1193	0.2041	0.9577

Both MAE and RMSE are reported in standardized (z-score) units rather than eV/atom. Random Forest yields the lowest error alongside the highest R^2, with ANN trailing just behind

MAE and RMSE Comparison of Machine-Learning Models

With mean absolute error (MAE) and root mean square error (RMSE) Figure 2 compares five models of machine learning side by side. While ANN reaches lowest RMSE (0.2041), Random Forest reaches lowest MAE (0.1065). CatBoost finishes last on the two metrics (MAE = 0.1671, RMSE = 0.2593). XGBoost and LightGBM generate error rates landed in-between these extremes. Since RMSE penalizes major errors heavier than MAE in order to see that it goes higher for each model, it suggests that there are relatively significant residuals available. The whole chart shows that Random Forest and ANN had the fewest errors in this set of tests. In Z-score standard units instead of eV/atom, the error metrics are reported.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

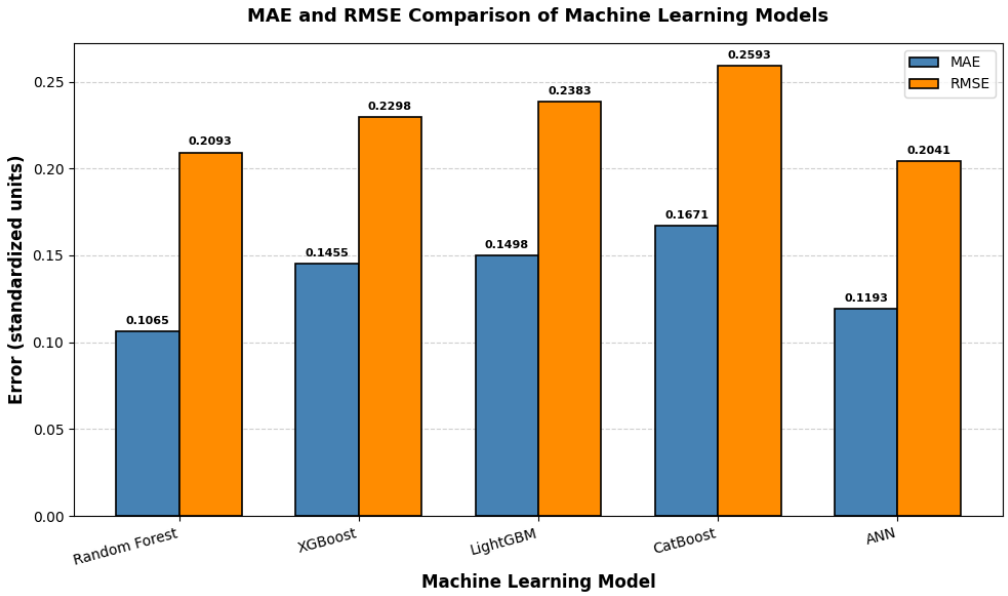


Figure 2. MAE and RMSE Comparison of Machine-Learning Models

Comparing Machine-Learning Models by R^2

A comparison of the coefficient of determination (R^2) for each of the five machine-learning models appears in Figure 3. Leading the group, ANN reached an R^2 of 0.9577, closely trailed by Random Forest at 0.9555. XGBoost, LightGBM, and CatBoost yielded R^2 values of 0.9464, 0.9424, and 0.9317, respectively. Every one of the five models exceeds an R^2 of 0.93, meaning all of them capture the bulk of the formation energy variance across the test set. The narrow gap between ANN and Random Forest indicates that both methods capture the dataset's underlying nonlinearities effectively. CatBoost's lower R^2 aligns directly with its higher MAE and RMSE shown in Figure 2.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

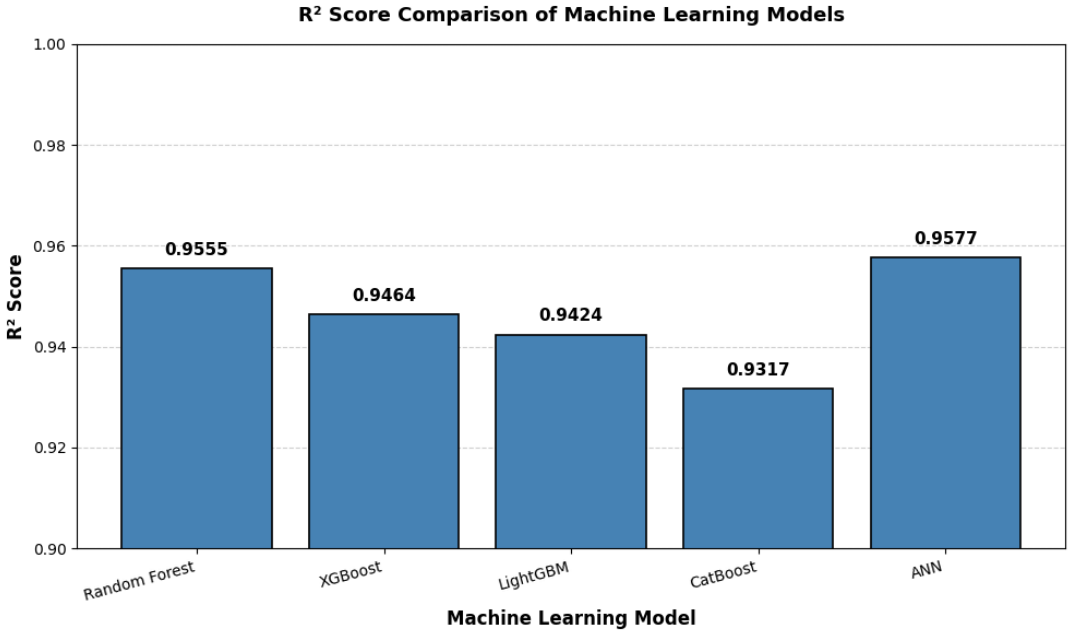


Figure 3. R^2 Comparison of Machine-Learning Model

Figure 4 plots measured formation energy against the values predicted by each of the five machine-learning models. Predictions hug the 1:1 line of perfect agreement across all models, spanning R^2 values from 0.9317 for CatBoost up to 0.9577 for ANN. Random Forest and ANN keep closest to the 1:1 line, fitting their higher R^2 values of 0.9555 and 0.9577, respectively. XGBoost and LightGBM sit in the middle, posting R^2 values of 0.9464 and 0.9424, respectively. CatBoost disperses the most, especially past 4 eV/atom on the positive side, which matches its lower R^2 and higher MAE and RMSE. Predictions from every model grow more dispersed at higher formation energies because data points are thinner in that range. The broad trends in formation energy come through across all five models, though Random Forest and ANN make the most accurate predictions.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

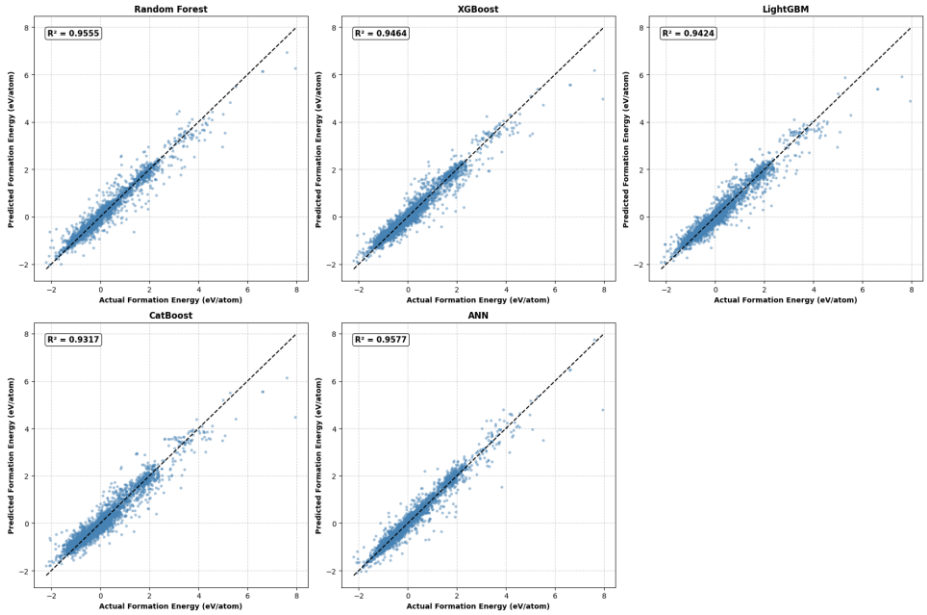


Figure 4. Actual vs Predicted Formation Energy for All Five Machine-Learning Models

Random Forest Feature Importance (Top 20, with energy_per_atom)

Table 2. Feature importances from a Random Forest model trained with energy_per_atom included, predicting formation energy_per_atom.

Rank	Feature	Importance
1	**energy_per_atom**	0.544634
2	**band_gap**	0.094659
3	**element_F**	0.083917
4	**density**	0.073585
5	**element_O**	0.069736
6	element_N	0.020598
7	contains_transition_metal	0.016554
8	volume	0.013656
9	n_elements	0.007897
10	element_S	0.005885
11	element_Cl	0.005350
12	element_Mg	0.005298
13	element_C	0.005178
14	element_Ti	0.004507
15	element_V	0.004184
16	element_P	0.004089
17	element_Cr	0.003144
18	element_Si	0.002884
19	element_H	0.002261

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Rank	Feature	Importance
20	element_Br	0.001971

The 20 most important features for the Random Forest model trained with energy_per_atom are listed in Table 2, led decisively by energy_per_atom at 0.545, more than five times the 0.095 value of band_gap in second place. Further down the list are element_F at 0.084, density at 0.074, and element_O at 0.070. None of the other features reaches an importance of 0.021. That spread makes clear how much the model depends on energy_per_atom whenever it can use it. Still, that reliance requires some caution, because energetic arguments link energy_per_atom straight to the target property. Testing generalization through the group-based cross-validation below puts the models under a much harder test than a random split ever would.

Top 20 Random Forest Features with energy_per_atom

Figure 5 plots feature importance scores for the top 20 variables from a Random Forest run that included energy_per_atom. That feature clearly heads the list with a score near 0.545, over five times higher than the second-ranked descriptor, band_gap, which sits at just around 0.095. Several additional features carry discernible weight, namely element_F, density, and element_O. A disparity this pronounced shows that the model leans heavily on energy_per_atom whenever the feature is present. Still, this calls for careful interpretation, since energetic considerations mean energy_per_atom follows the target property directly. Evaluating generalization via the upcoming group-based cross-validation sets a much higher bar than a basic random split, giving a sounder read on how the model actually performs.

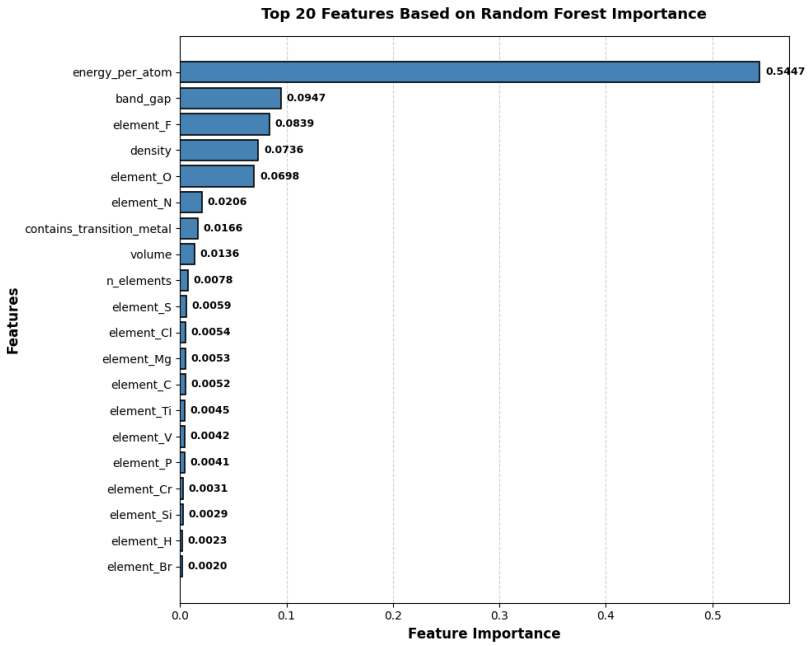


Figure 5. Top 20 Random Forest Features with energy_per_atom

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

SHAP Summary Plot for Random Forest with energy_per_atom

In Figure 6, the SHAP summary distribution as defined by the Random Forest trains with energy_per_atom is shown. The distribution shows the dominant role of Energy_per_atom, whose SHAP values extends far beyond zero and indicates an important effect on individual predictions, while band_gap, element_O, element_F. Density exert visible but considerably more confined effects. This pattern is well aligned with feature importances that are reported on Figure 5. SHAP values includes both scope and the impact of introductions; consequently the width of distribution is the effect on model predictions and is not the unsteady positive or negative rule of physics.

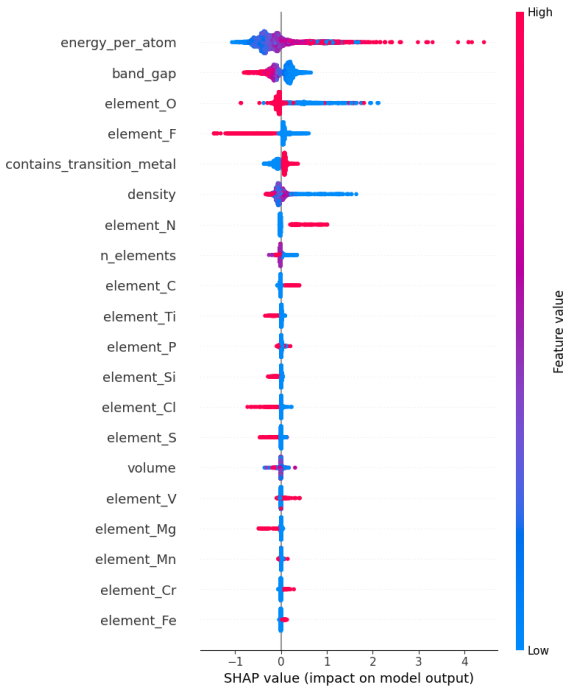


Figure 6. SHAP Summary Plot for Random Forest with energy_per_atom

Top 20 Random Forest Feature Importance without energy_per_atom

Figure 7 displays the 20 most important features for the Random Forest model when energy_per_atom is excluded from training. Taking energy_per_atom out of the mix drastically alters how the other features rank against one another. Instead, chemical and structural variables take on most of the weight, paced by element_F (˜0.28), element_O (˜0.24), density (˜0.12), band_gap (˜0.09), and volume (˜0.08). Predictive power is clearly not confined to energy_per_atom: this reordering proves the model can also read formation-energy trends out of structural features, composition, and electronic structure. Relying less on an energy-based feature clarifies the interpretation and shows plainly which descriptors fill the gap once energy_per_atom is excluded.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

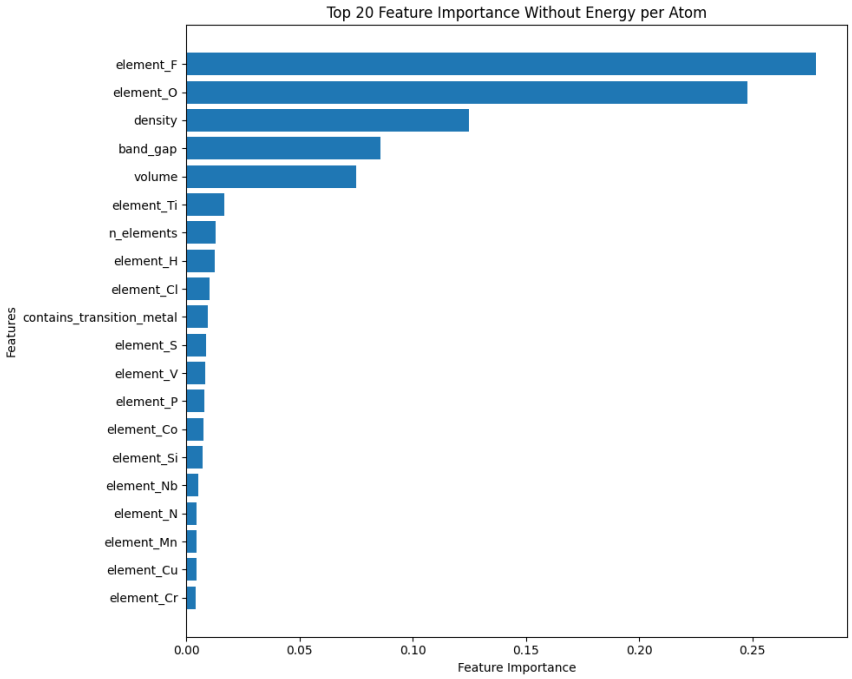


Figure 7. Top 20 Random Forest Feature Importance without energy_per_atom

Table 3 Top 20 SHAP Features for Random Forest (Without energy_per_atom)
Mean absolute SHAP value per feature, Random Forest model trained without energy_per_atom.

Rank	Feature	Mean Absolute SHAP Value
1	**element_O**	0.407876
2	**band_gap**	0.241063
3	**element_F**	0.236668
4	**density**	0.167180
5	**volume**	0.105153
6	element_Ti	0.053247
7	contains_transition_metal	0.052904
8	n_elements	0.044822
9	element_Co	0.042220
10	element_P	0.033923
11	element_V	0.029775
12	element_Mn	0.026991
13	element_H	0.026039
14	element_S	0.025500
15	element_Si	0.024563
16	element_Cl	0.018413
17	element_Ni	0.018343
18	element_Nb	0.016299

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Rank	Feature	Mean Absolute SHAP Value
19	element_Cu	0.015332
20	element_C	0.014637

Without energy_per_atom leaves element_O as the most influential feature at 0.408, with band_gap at 0.241 and element_F at 0.237 coming next. Together, the top five features—element_O, band_gap, element_F, density, and volume—make up around 72% of the total importance among the top 20.

SHAP Summary Plot for Random Forest without energy_per_atom

The SHAP summary plot in Figure 8 corresponds to the Random Forest model trained after leaving out energy_per_atom. Under these conditions, predictions are driven most heavily by element_O, band_gap, and element_F, with density and volume trailing just behind. This matches what Table 3 shows for mean absolute SHAP values: 0.408, 0.241, 0.237, 0.167, and 0.105, in that order. Because model influence is concentrated in these few top descriptors, a small set of chemical, electronic, and structural properties clearly supplies the bulk of the predictive signal. Because SHAP values capture contribution sizes, the plot reveals how strongly each feature shifts predictions rather than establishing any universal causal relationship with formation energy.

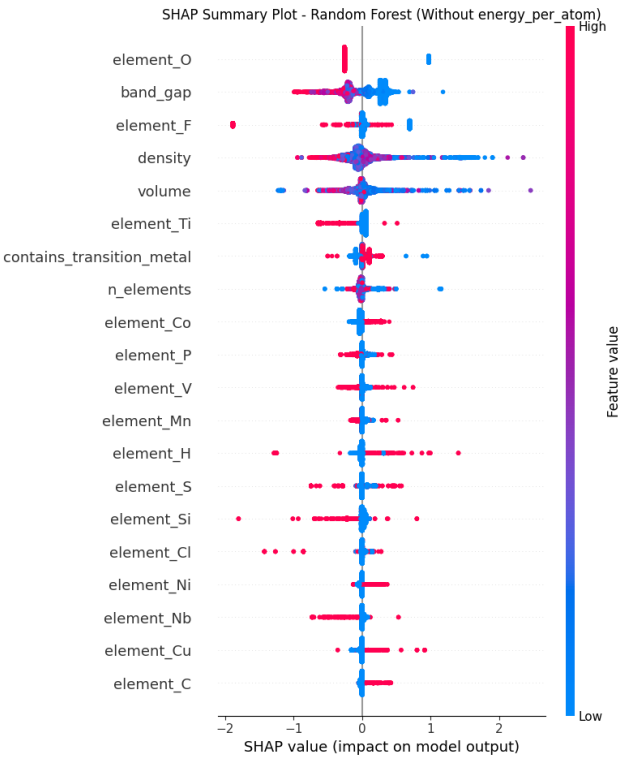


Figure 8. SHAP Summary Plot for Random Forest without energy_per_atom

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Top 20 SHAP Feature Importance without energy_per_atom

In Figure 9, mean absolute SHAP values are ranked directly for the Random Forest model trained without energy_per_atom: element_O stands out well ahead at 0.408, followed by band_gap at 0.241 and element_F at 0.237, with density at 0.167 and volume at 0.105 adding notable weight as well. Combined, these first five descriptors make up a heavy portion of the run's total SHAP contribution. Past those top five, the remaining features taper into a low tail, which shows that none of them individually shifts the model's predictions by much. That pattern backs up the idea that when energy_per_atom is omitted, baseline structural metrics, chemical makeup, and electronic properties still deliver genuine predictive utility.

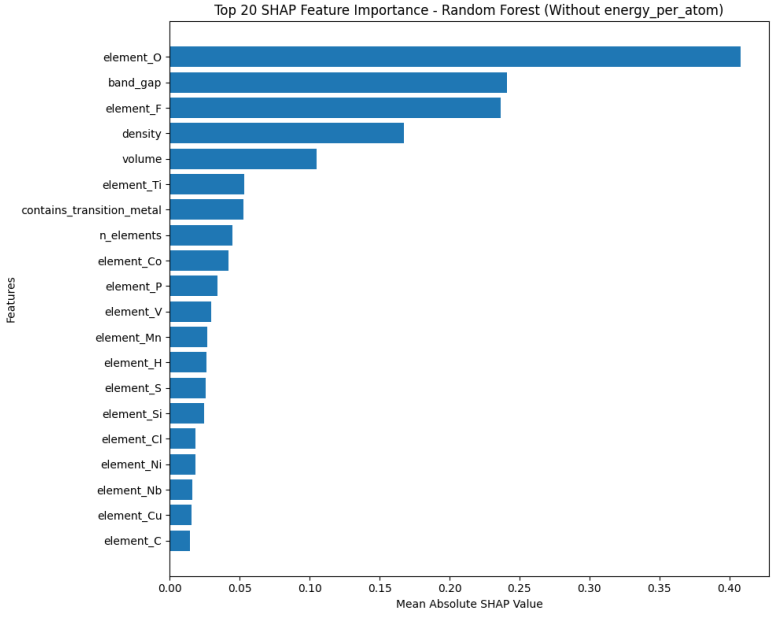


Figure 9. Top 20 SHAP Feature Importance without energy_per_atom

Table 4. Random Forest Performance With and Without energy_per_atom Using 5-Fold Group Cross-Validation

Feature Set	MAE (Mean ± SD)	RMSE (Mean ± SD)	R^2 (Mean ± SD)
RF with energy_per_atom	0.1071 ± 0.0046	0.2123 ± 0.0161	0.9547 ± 0.0065
RF without energy_per_atom	0.1719 ± 0.0040	0.4042 ± 0.0108	0.8366 ± 0.0074

The Random Forest results evaluated across 5-fold GroupKFold cross-validation, comparing setups with and without energy_per_atom, are listed in Table 4. A seemingly contradictory trend appears: taking out energy_per_atom drops the mean MAE from 0.1071 to 0.1719, yet RMSE climbs from 0.2123 to 0.4042 and R^2 falls from 0.9547 to 0.8366. This split suggests that although errors are smaller on average once energy_per_atom is gone, a handful of larger misses drive up the RMSE and drag down the R^2. This highlights why model evaluation needs to rely on more than a single metric.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

With group-based cross-validation, omitting energy_per_atom causes mean R^2 to decline by roughly 0.118, which reveals that data leakage across duplicate rows of the same material was behind part of its strong showing in the random 80/20 split. Even in this stripped-back form, the model reaches an R^2 of 0.8366, which confirms that what remains of the chemical, electronic, and structural descriptors still holds substantial predictive power.

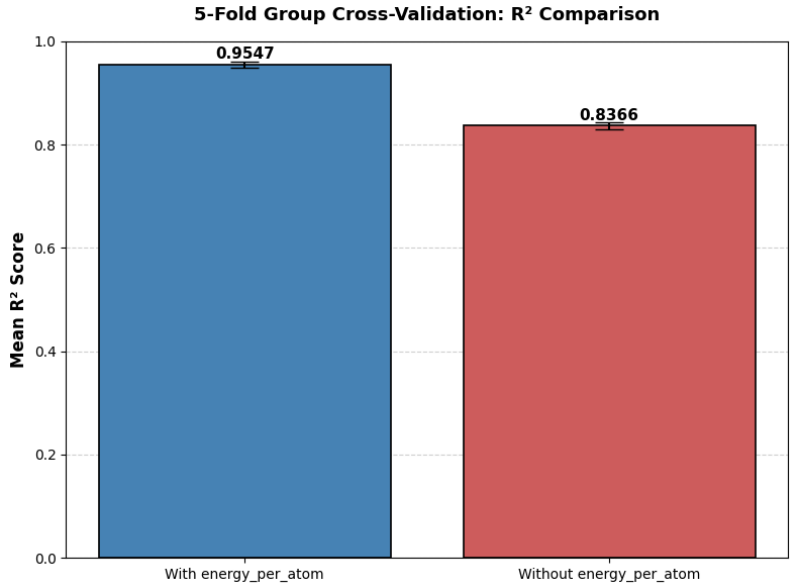


Figure 10 5-Fold Group Cross-Validation: R^2 Comparison

Evaluating the Random Forest through group-based cross-validation subjects it to a much stiffer test than a simple random train-test split provides. With energy_per_atom included, the average R^2 comes out to 0.9547 ± 0.0065, whereas excluding it brings that number down to 0.8366 ± 0.0074. Removing energy_per_atom reduces the average R^2 by about 0.118. Since that decline is fairly modest, the model evidently retains much of its predictive power without the feature. This divergence between group-level results and feature importance calculated on a random split makes it clear that importance values from a single random split cannot serve as direct indicators of true out-of-sample generalization.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

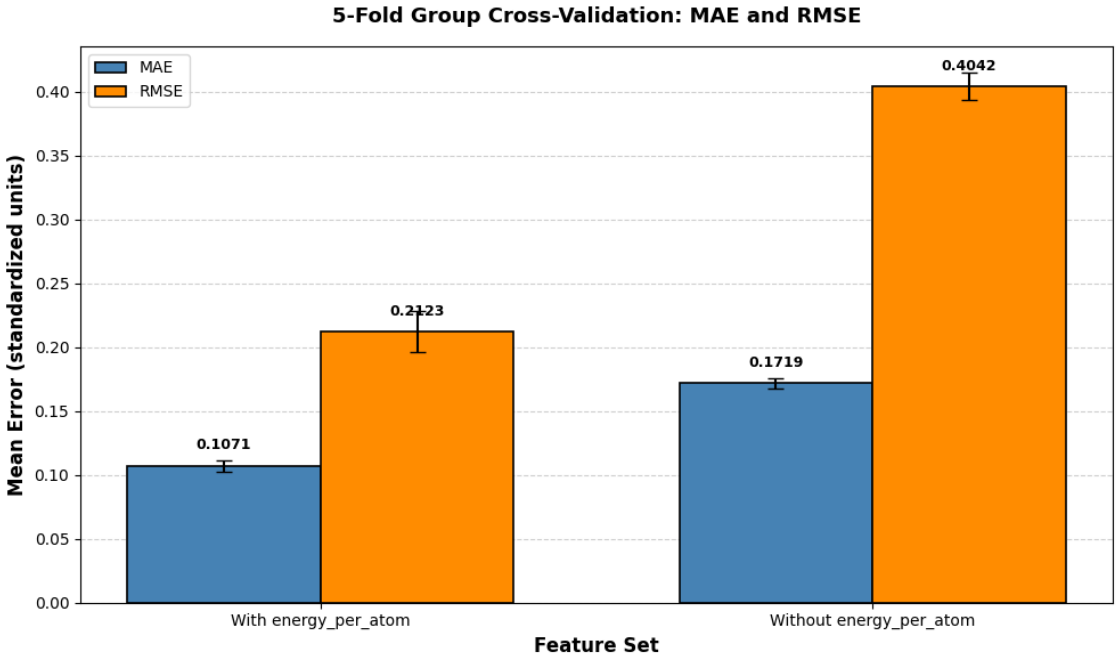


Figure 11 5-Fold Group Cross-Validation: MAE and RMSE

Both MAE and RMSE metrics from 5-fold group cross-validation appear together in Figure 11. When energy_per_atom remains in the feature set, the model posts an MAE of 0.1071 ± 0.0046 and an RMSE of 0.2123 ± 0.0161. Without energy_per_atom, the MAE rises to 0.1719 ± 0.0040, while the RMSE climbs even more steeply to 0.4042 ± 0.0108. Omitting that feature worsens the mean absolute error, and the inflated RMSE reflects a model that has become somewhat more vulnerable to occasional large errors. A difference like that shows why model evaluation requires multiple metrics instead of relying on any single summary number. With standard deviations remaining fairly narrow, results across the individual folds look reasonably stable too.

Table 5.Final ML-Screened Candidate Materials
Top candidate materials ranked by lowest out-of-fold (OOF) prediction error between actual and model-predicted formation energy per atom.

Rank	Material ID	Formula	Actual Formation Energy	OOF Predicted Formation Energy	Absolute Error
1	mp-1210003	$NaLiHo_2F_8$	-2.253269	-2.253893	0.000624
2	mp-1209986	$NaLiEr_2F_8$	-2.261011	-2.248127	0.012885
3	mp-1210006	$NaLiTm_2F_8$	-2.272452	-2.245109	0.027342
4	mp-1209991	$NaLiLu_2F_8$	-2.269808	-2.175352	0.094457
5	mp-3941	$LiYF_4$	-2.270734	-2.154550	0.116184
6	mp-1195827	$SrLiAlF_6$	-2.091252	-2.045602	0.045650
7	mp-1211459	$KLiErF_5$	-2.023767	-2.044624	0.020857
8	mp-1211480	$KLiLuF_5$	-2.028912	-2.035403	0.006491

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Rank	Material ID	Formula	Actual Formation Energy	OOF Predicted Formation Energy	Absolute Error
9	mp-1194722	$KLiTmF_5$	-2.032419	-2.026090	0.006329
10	mp-543044	$BaLiAlF_6$	-2.075023	-2.022238	0.052785
11	mp-1211435	$KLiTbF_5$	-2.016970	-2.018399	0.001429
12	mp-1211461	$KLiDyF_5$	-2.018769	-2.018109	0.000660
13	mp-1211494	$KLiHoF_5$	-2.020374	-2.016972	0.003402
14	mp-556237	$KLiYF_5$	-1.993220	-2.002469	0.009249
15	mp-542219	Li_2ZrF_6	-1.891901	-1.953976	0.062075
16	mp-1193222	$LiMgAlF_6$	-1.915061	-1.830382	0.084679

All 16 candidates are lithium fluoride compounds, consistent with the earlier finding that fluorine strongly lowers predicted formation energy. The top-ranked material ($NaLiHo_2F_8$) has an absolute error of only 0.000624 eV/atom, indicating an almost exact match between actual and predicted values.

Final ML-Screened Candidate Materials

Out-of-fold (OOF) predicted formation energy is written on Figure 12 for 16 applicants. Every candidate has a negative predicted formation energy between –2.254 eV/atom—for the Naliho2F8-1.830-eV/atom—for LiMgAlF6 indicating favorable thermodynamic stability. This confirms that the screening pipeline can pick materials if model predictions are closely related to formation energies. These compounds remain strictly ML to select candidates and are still not tested.

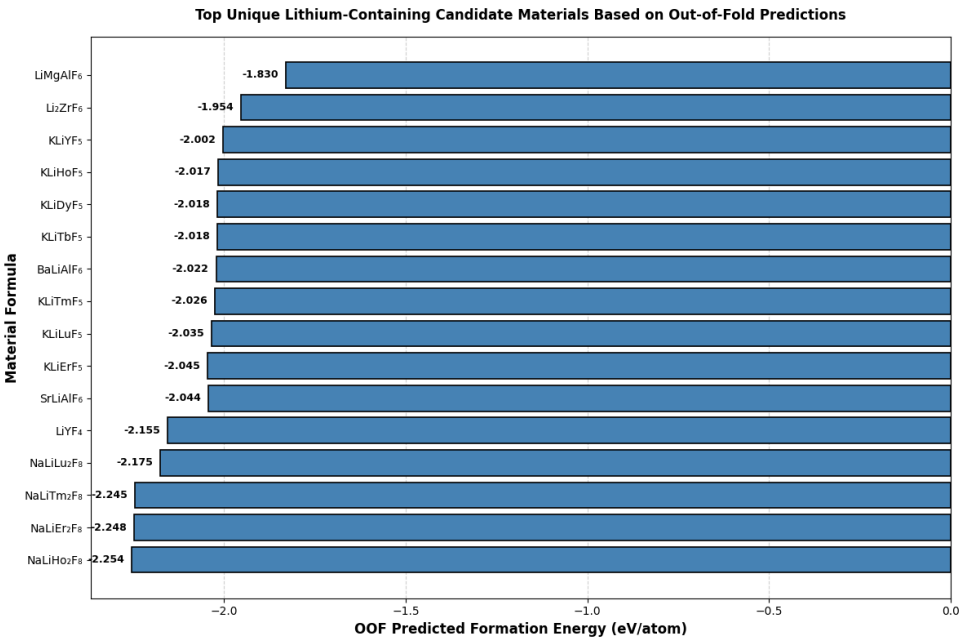


Figure 12. Final ML-Screened Candidate Materials

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Random Forest Actual vs. Predicted Formation Energy (energy_per_atom Excluded)

Figure 13 plots actual against predicted formation energy for the Random Forest model when energy_per_atom is omitted from training. A clear positive correlation remains between the predictions and actual values, showing that general formation-energy trends are still captured by the remaining structural, electronic, and chemical features. The points spread out much further from the diagonal line than they did with the full feature set (Figure 4), and a handful drift into noticeably higher errors. What the plot shows matches the GroupKFold metrics directly, where omitting energy_per_atom drops R^2 from 0.9547 to 0.8366 and raises RMSE from 0.2123 to 0.4042. What these results show is that energy_per_atom helps tighten the fit, yet the other descriptors still retain substantial predictive signal, even if they fail to track every data point quite as well.

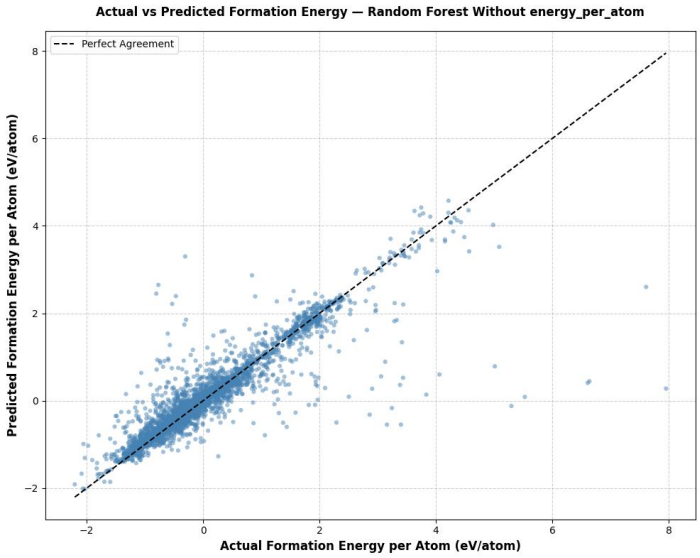


Figure 13. Actual vs Predicted Formation Energy Random Forest without energy_per_atom

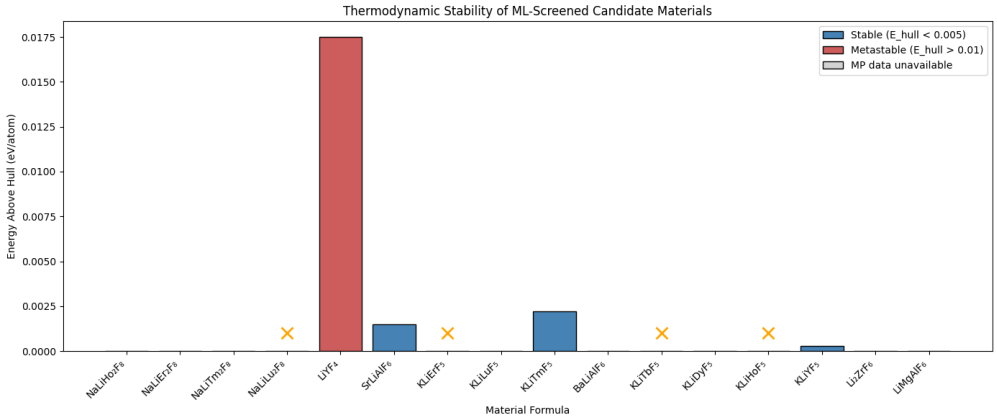


Figure 14. Thermodynamic Stability of ML-Screened Candidate Materials

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

The E_hull values calculated for all 16 candidates appear in Figure 14. Half of the candidates (eight, or 50.0%) register E_hull values of 0.0000 eV/atom, meaning they sit right on the convex hull and qualify as thermodynamically stable. Positive E_hull values between 0.0003 and 0.0174 eV/atom put four of the candidates—$LiYF_4$, $SrLiAlF_6$, $KLiTmF_5$, and $KLiYF_5$—just above the convex hull. Missing E_hull data meant that the last four candidates—$NaLiLu_2F_8$, $KLiErF_5$, $KLiTbF_5$, and $KLiHoF_5$—could not be evaluated.

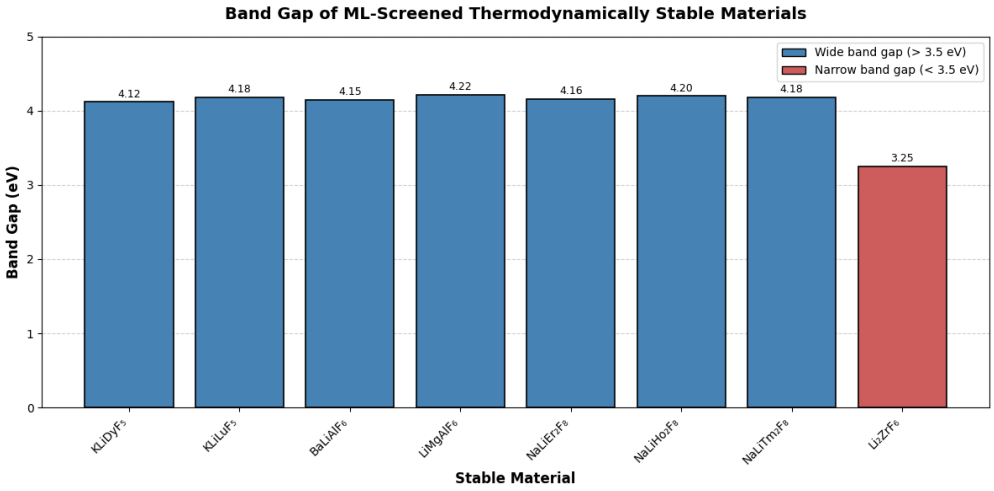


Figure 15. Band Gap of ML-Screened Thermodynamically Stable Materials

The band gaps for the 8 stable materials are plotted in Figure 15. Seven of the candidates have wide band gaps between 4.12 and 4.22 eV, but Li_2ZrF_6 drops to a narrower 3.25 eV. Because these band gaps stay broad across the board, the fluoride-based compounds appear to be electronic insulators.

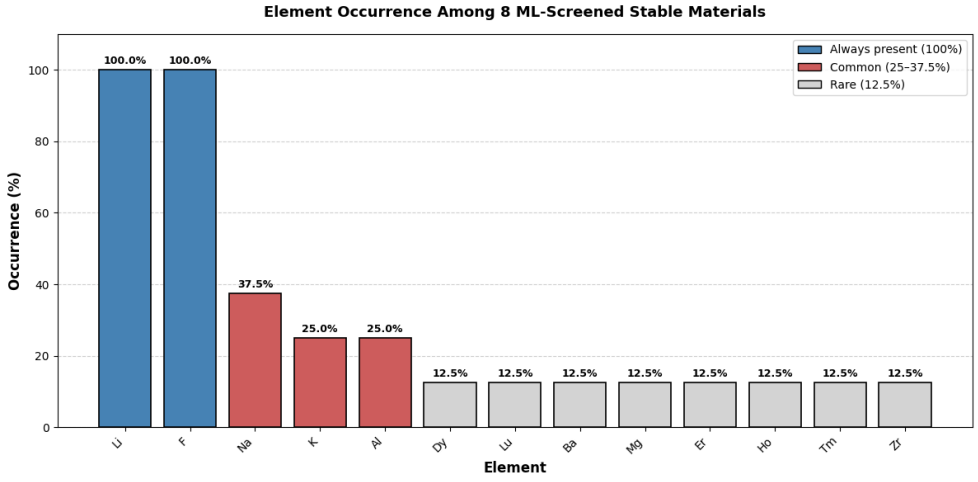


Figure 16. Element Occurrence Among 8 ML-Screened Stable Materials

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

The elemental makeup of the 8 stable materials identified through ML screening is given in Figure 16. Every one of the 8 candidates contains lithium (Li) and fluorine (F) (100%), which confirms their identity as lithium fluoride-based compounds. Sodium (Na) turns up in 3 candidates (37.5%), whereas potassium (K) and aluminum (Al) are each found in 2 (25.0%). Dysprosium (Dy), lutetium (Lu), barium (Ba), magnesium (Mg), erbium (Er), holmium (Ho), thulium (Tm), and zirconium (Zr) make up the rest, with each found in just 1 candidate (12.5%). Among them are several rare earths (Dy, Lu, Er, Ho, Tm) that typically help stabilize structure and open up wide band gaps in fluorides. Because Li and F turn up in every single candidate, the stabilizing effect of fluorine on lithium-based materials comes through clearly.

Table 6: Statistical Summary of Stable Materials

Property	Count	Mean	Std	Min	Max
n_elements	8	3.875	0.354	3.000	4.000
formation_energy_per_atom	8	-2.0895	0.1549	-2.2725	-1.8919
band_gap	8	4.0898	0.3402	3.2542	4.2556

A statistical overview of the 8 stable materials is given in Table 6. On average, the materials incorporate 3.875 elements in each formula unit. Formation energies run from -2.2725 to -1.8919 eV/atom, yielding a mean of -2.0895 eV/atom. Band gaps fall between 3.2542 and 4.2556 eV, with an average value of 4.0898 eV. These relatively tight standard deviations show that the ML-screened candidates sit within a well-defined pocket of chemical space.

Table 7: Final Stability Status Summary

Status	Count	Percentage	Materials
Stable	8	50.0%	$NaLiHo_2F_8$, $NaLiEr_2F_8$, $NaLiTm_2F_8$, $KLiLuF_5$, $BaLiAlF_6$, $KLiDyF_5$, Li_2ZrF_6, $LiMgAlF_6$
Above Hull	4	25.0%	$LiYF_4$, $SrLiAlF_6$, $KLiTmF_5$, $KLiYF_5$
MP Data Unavailable	4	25.0%	$NaLiLu_2F_8$, $KLiErF_5$, $KLiTbF_5$, $KLiHoF_5$

The final stability status for each of the 16 ML-screened candidates is outlined in Table 7. Half of the materials (eight, or 50.0%) proved to be thermodynamically stable, while four (25.0%) lie above the convex hull and another four (25.0%) lack E_hull data in the Materials Project database.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

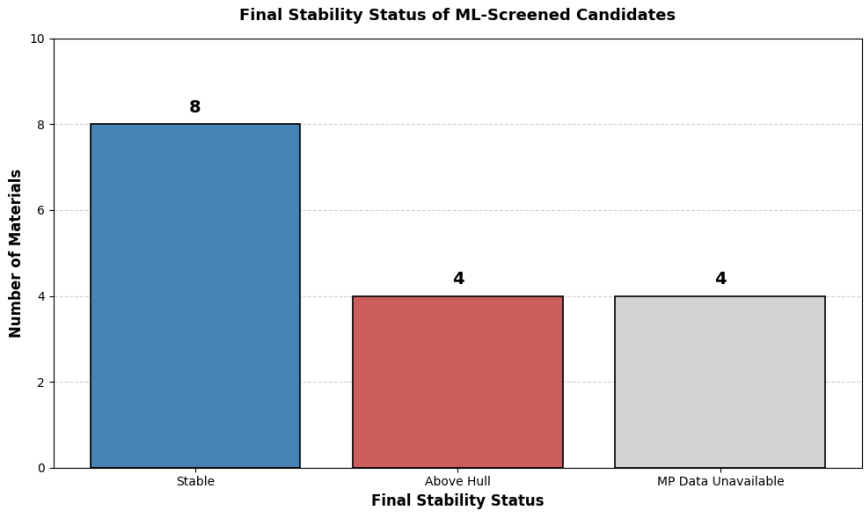


Figure 17. Final Stability Status of ML-Screened Candidates.

A bar chart of the final stability outcomes across the 16 ML-screened candidates is provided in Figure 17. Half of them (eight materials, or 50.0%) proved thermodynamically stable, while four (25.0%) sit above the convex hull and another four (25.0%) have no E_hull entries in the Materials Project database. More candidates ended up stable than either above-hull or unverified. Reaching a 50% hit rate confirms that the ML screening workflow did a good job identifying stable lithium-based compounds. The four materials without E_hull values could be resolved in future work using DFT convex-hull calculations.

Table 8: ML-Screened Candidate Materials with Stability Status

Rank	Material ID	Formula	Actual FE (eV/atom)	OOF Pred. (eV/atom)	Abs. Error (eV/atom)	E_hull (eV/atom)	Final Status
1	mp-1210003	$NaLiHo_2F_8$	-2.2533	-2.2539	0.0006	0.0000	Stable
2	mp-1209986	$NaLiEr_2F_8$	-2.2610	-2.2481	0.0129	0.0000	Stable
3	mp-1210006	$NaLiTm_2F_8$	-2.2725	-2.2451	0.0273	0.0000	Stable
4	mp-1209991	$NaLiLu_2F_8$	-2.2698	-2.1754	0.0945	—	MP data unavailable
5	mp-3941	$LiYF_4$	-2.2707	-2.1546	0.1162	0.0174	Above hull
6	mp-1195827	$SrLiAlF_6$	-2.0913	-2.0456	0.0457	0.0015	Above hull
7	mp-1211459	$KLiErF_5$	-2.0238	-2.0446	0.0209	—	MP data unavailable
8	mp-1211480	$KLiLuF_5$	-2.0289	-2.0354	0.0065	0.0000	Stable
9	mp-1194722	$KLiTmF_5$	-2.0324	-2.0261	0.0063	0.0022	Above hull
10	mp-543044	$BaLiAlF_6$	-2.0750	-2.0222	0.0528	0.0000	Stable
11	mp-1211435	$KLiTbF_5$	-2.0170	-2.0184	0.0014	—	MP data

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

Rank	Material ID	Formula	Actual FE (eV/atom)	OOF Pred. (eV/atom)	Abs. Error (eV/atom)	E_hull (eV/atom)	Final Status
							unavailable
12	mp-1211461	$KLiDyF_5$	-2.0188	-2.0181	0.0007	0.0000	Stable
13	mp-1211494	$KLiHoF_5$	-2.0204	-2.0170	0.0034	—	MP data unavailable
14	mp-556237	$KLiYF_5$	-1.9932	-2.0025	0.0092	0.0003	Above hull
15	mp-542219	Li_2ZrF_6	-1.8919	-1.9540	0.0621	0.0000	Stable
16	mp-1193222	$LiMgAlF_6$	-1.9151	-1.8304	0.0847	0.0000	Stable

The 16 ML-screened candidate materials and their stability classifications appear in Table 8. Each entry in the table gives the material ID, chemical formula, actual formation energy, OOF-predicted formation energy, absolute prediction error, energy above hull (E_hull), and final stability classification. Ranking the 16 candidates by absolute prediction error places $NaLiHo_2F_8$ at the lowest (0.0006 eV/atom) and $LiMgAlF_6$ at the highest (0.0847 eV/atom). According to Materials Project E_hull values, 8 candidates are stable at 0.0000 eV/atom, 4 lie above the hull ($LiYF_4$, $SrLiAlF_6$, $KLiTmF_5$, $KLiYF_5$), and 4 lack E_hull data entirely ($NaLiLu_2F_8$, $KLiErF_5$, $KLiTbF_5$, $KLiHoF_5$). Every one of the 16 candidates is a lithium fluoride compound, matching the observation that fluorine drives down predicted formation energy substantially.

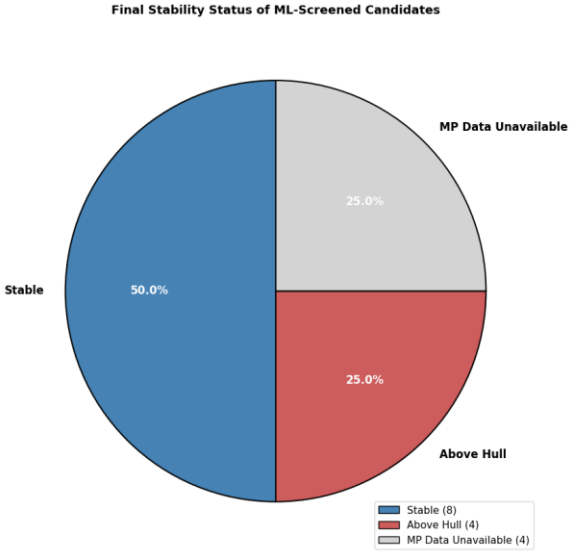


Figure 18. Final Stability Status of ML-Screened Candidates.

A breakdown of the final stability outcomes for all 16 ML-screened candidates appears as a pie chart in Figure 18. Half of the candidates, 8 materials (50.0%), proved thermodynamically stable, each registering an energy above hull of 0.0000 eV/atom. Another four materials (25.0%), specifically $LiYF_4$, $SrLiAlF_6$, $KLiTmF_5$, and $KLiYF_5$, fell above the hull, yielding positive

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

E_hull values between 0.0003 and 0.0174 eV/atom. Because the Materials Project database lacked E_hull values for them, the last four materials (25.0%)—$NaLiLu_2F_8$, $KLiErF_5$, $KLiTbF_5$, and $KLiHoF_5$—could not be assessed. With 50% of the candidates verified as stable, these findings confirm that the ML screening workflow effectively singled out thermodynamically favorable lithium-based materials.

Conclusion

We developed an interpretable machine-learning pipeline to estimate per-atom formation energies and screen candidates containing lithium. After preprocessing and removing duplicates, we were left with a final set of 21,747 unique materials described by chemical, structural, electronic, and energetic features. Across the five machine-learning models evaluated, Random Forest achieved the lowest MAE(0.1065) while ANN achieved the highest R^2(0.9577) .Group-based cross-validation gave us a much better read on true generalization. With energy_per_atom included, Random Forest reached an R^2 of 0.9547 ± 0.0065, which fell to 0.8366 ± 0.0074 when that feature was removed. That the decline remains modest indicates the surviving chemical, electronic, and structural descriptors retain substantial predictive power on their own. With energy_per_atom excluded, SHAP values identified element_O, band_gap, element_F, density, and volume as the strongest contributors. Applying out-of-fold predictions to lithium-containing compounds narrowed the field to 16 unique ML-screened candidates. When checked against Materials Project energy-above-hull figures, 8 (50%) proved thermodynamically stable, while four sit above the convex hull and four have no recorded E_hull data. $NaLiHo_2F_8$ (mp-1210003) gave the smallest absolute prediction error reported, at 0.000624. All told, these findings show that interpretable machine learning has a practical place in high-throughput screening for lithium-based materials. All the same, these remain ML-screened candidates rather than confirmed thermodynamically stable materials, and proving them out will require DFT calculations, energy-above-hull checks, and convex-hull analyses.

Broader Context and Social Impact

These results show that screening materials with machine learning could aid the larger effort to build sustainable energy-storage systems. The proposed framework predicts formation energy quickly and picks out the descriptors tied to thermodynamic stability, cutting down the compute time and lab work needed to filter large sets of lithium-based candidates. That could speed up how quickly promising candidates are flagged for lab testing while helping research teams use their resources more effectively. More broadly, getting better at finding stable lithium-based compounds feeds into building advanced storage systems, which matter directly for grid-scale renewables, electric vehicles, and other new energy applications. What this study shows, then, is the societal and technological benefit that comes from putting machine learning alongside computational materials science in the search for clean energy.

REFERENCES

[1] Un C. Comparative Analysis of Lithium-Ion Battery Chemistries and Cell Formats: Implications for Charging Performance, Discharge Behavior and Thermal Characteristics in Electric Vehicle Applications.

[2] Guo W, Dun C, Guo J, Urban JJ, Yu C, Zhang Q, Qiu J. Tailoring Materials Design for Aqueous Energy Storage and Conversion through Electrochemical Reconstruction. Chemical Reviews. 2026 Mar 17;126(7):4062-134.

[3] Dallmann B, Saha A, Rosen AS. Predicting the Thermodynamic Limits of Metal–Organic Framework Metastability. Journal of the American Chemical Society. 2026 May 11;148(19):19487.

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

[4] Guo Y, Wang Y, Zhang W, Shen Z, Zeng L. From algorithms to applications: a comprehensive review of machine learning in computational materials science. Archives of Computational Methods in Engineering. 2026 Jan;33(1):1445-66.

[5] Xu L, Zhou J, Bandarenka AS, Chen Z. Data-Driven Electrocatalyst Discovery: Recent Trends in Machine Learning Approaches and Descriptor-Based Design Principles. Accounts of Materials Research. 2026 Feb 4;7(3):294-308.

[6] Abraham BM, Gogotsi Y. Machine learning toolkits and frameworks for materials design. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2026 Mar;16(2):e70067.

[7] Xin Y, Liu P, Xie Z, Mi W, Gao P, Zhao HJ, Lv J, Wang Y, Ma Y. Bridging theory and experiment in materials discovery: Machine-learning-assisted prediction of synthesizable structures. Physical Review B. 2026 Feb 1;113(5):054102.

[8] Kumar Yadav A. Artificial intelligence in accelerating materials discovery: opportunities and challenges. ChemistrySelect. 2026 Feb;11(7):e07191.

[9] Kim KS. Machine learning for accelerating energy materials discovery: Bridging quantum accuracy with computational efficiency. Advanced Energy Materials. 2026 Jan;16(2):e03356.

[10] Wei G, Li M, Cui B, Xiu W, Sarman AM. Machine Learning for Structural Steels: Materials Design, Property Prediction, Durability, and Future Directions. Materials. 2026 Aug 25;19(17):3612.

[11] Kim J, Cho H, Jeon H, Jung J, Han S. Reactive machine learning interatomic potentials for chemistry and materials science. Chemical Reviews. 2026 Apr 22;126(8):4467-510.

[12] Oviedo F, Ferres JL, Buonassisi T, Butler KT. Interpretable and explainable machine learning for materials science and chemistry. Accounts of Materials Research. 2022 Jun 3;3(6):597-607.

[13] Park IK, Pourasad S, Mun J, Rajan AC, Park TH, Rohit A, Zafari M, Pham TL, Jung E, Koo B, Ha M. Machine Learning Interatomic Potentials for Energy Materials: Architectures, Training Strategies, and Applications. Advanced Energy Materials. 2026 May 12:e71046.

[14] Purohit A. Free energy calculations in molecular modeling: from classical methods to machine learning. Journal of Molecular Modeling. 2026 Apr;32(4):110.

[15] Wu J, Wang H, Wang C, Chen H, Han J, Zhang L. Machine Learning-Driven Design of Electrolytes and Interfaces for Solid-State Lithium Batteries: Progress and Outlook. Advanced Functional Materials. 2026:e78357.

[16] Mundu MM, Sempewo JI, Goparaju A, Uti DE. Comparative analysis of model evaluation metrics in energy systems, environmental modeling, and sustainability science. International Journal of Energy Research. 2026;2026(1):6170467.

[17] Vasconcelos HC, Meirelles M. Interpretable formation-energy structure from composition, density, and band-gap descriptors. Applied Physics A. 2026 Aug;132(8):748.

[18] Jain V, Wang Z, You F. Machine learning pipelines for the design of solid-state electrolytes. Materials Horizons. 2026 Jan 12;13(1):15-44.

[19] Jefferson TE, Rana S, Banerjee N. Lithium-Ion Batteries: A Review of Materials, Design Advances, Sustainability Challenges, and Future Prospects. Energy Storage. 2026 Feb;8(1):e70320.

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

[20] Alhashmi A, Kanoun MB, Goumri-Said S. Machine learning for halide perovskite materials ABX3 (B= Pb, X= I, Br, Cl) assessment of structural properties and band gap engineering for solar energy. Materials. 2023 Mar 27;16(7):2657.

[21] GP JJ. Integration of density functional theory and machine learning for materials discovery in energy applications. Discover Chemistry. 2026 Sep 22;3(1):531.

[22] Soares TA, Nunes-Alves A, Bereau T, Friederich P, Gryn'ova G, Müller C, Wade RC, Merz Jr K. Mapping the Molecular Universe: Exploring Chemical Compound Space by Multiscale High-Throughput Screening and Machine Learning. Journal of Chemical Information and Modeling. 2026 Jun 22;66(12):6809-13.

[23] Parackal AS. Crystal Symmetry and Machine Learning for Systematic Materials Discovery. Linköping University Electronic Press; 2026 Feb 13.

[24] Wu K, Sang Z, Zhang G, Pang H. The Applications of Machine Learning in Micro-Nano Materials Research: From High-Throughput Screening to Intelligent Design. Small. 2026 Aug 10:e75132.

[25] Liu S, Dong D, Xiao Z, Bedrov D, Xu K, Xing L, Li W. Machine learning in electrolyte design: Accelerating discovery and multidimensional optimization for next-generation batteries. Chem. 2026 Jul 9;12(7).

[26] Majewski M, Malo L, Montero-Blay A, Marengo M, Dos Santos RN, Gkeka P, Minoux H. Agent-driven Model Development for RNA 3D Structure Prediction. bioRxiv. 2026 Sep 8:2026-09.

[27] Wang S, Liu J, Song X, Xu H, Gu Y, Fan J, Sun B, Yu L. Artificial intelligence empowers solid-state batteries for material screening and performance evaluation. Nano-Micro Letters. 2025 Dec;17(1):287.

[28] Emani S, Velidi G. Machine Learning and Artificial Intelligence in Digital Twins. Digital Twin Technology for Solar Concentrators. 2026 Sep 10:239-80.

[29] Zhou Y, Hu S, Zhang X, Wang H, Tan G, Jia W. MatRIS: Toward reliable and efficient pretrained machine learning interatomic potentials. arXiv preprint arXiv:2603.02002. 2026 Mar 2.

[30] Barhoumi M. AI-Guided Two-Dimensional Materials: From Discovery and Property Prediction to Device Engineering. Available at SSRN 6314398. 2026 Feb 27.

[31] Yoon H, Cho JH, Ong WJ, Kim SY. Machine Learning Meets Photocatalytic Water Splitting for Sustainable Hydrogen Production. ChemPhotoChem. 2026 Aug;10(8):e70239.

[32] Zhao L, Zong H. AI-Driven Decoding of Material Dynamics: From Machine Learning Potentials and Interpretability to Generative Prediction. Advanced Materials. 2026 May;38(29):e14626.

[33] He T, Xu B, Choi JU, Zhu M, Zhang Y, Niu C, Zhou G, Wang W, Liu D, Lee PS, Chen J. Engineering a low lattice mismatch interface layer for durable and bi-stable dynamic smart windows based on reversible metal electrodeposition. Energy & Environmental Science. 2026 May 12;19(9):2949-62.

[34] Romiluyi O. Beyond the Black Box: An Interpretable Machine Learning Framework for Predicting Electronic Structure Microdescriptors and Structure-Performance Relationships in Fe-based Catalytic Systems. arXiv preprint arXiv:2605.08994. 2026 May 9.

[35] Choi J, Kim S, Park J, Choi J, An S, Jung Y. AI for Accelerated Materials Discovery: From Generative Design to Autonomous Realization. Chemical Reviews. 2026 Aug 12;126(15):8644-93.

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 6 • October 2026

[36] Farooq MU, Chen F. Predicting ground-state structures of refractory W-Mo-Ta-Ti nanoclusters using deep reinforcement learning and machine learning potentials. International Journal of Refractory Metals and Hard Materials. 2026 Jul 28:108014.