Journal of Minerals and Materials Characterization and Engineering, 2026, 14(4), 109-121
https://www.scirp.org/journal/jmmce
ISSN Online: 2327-4085
ISSN Print: 2327-4077

Computational Machine Learning-Based Prediction of Crystal Structure in Mixed B-Site Perovskite Oxide: $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$

Lomus Hona[1*], Carl White[2], Ram Krishna Hona[2*]

[1]Department of General Education, United Tribes Technical College, Bismarck, ND, USA
[2]Department of Environmental Science, United Tribes Technical College, Bismarck, ND, USA
Email: *lhona@uttc.edu, *rhona@uttc.edu

How to cite this paper: Hona, L., White, C. and Hona, R.K. (2026) Computational Machine Learning-Based Prediction of Crystal Structure in Mixed B-Site Perovskite Oxide: $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$. *Journal of Minerals and Materials Characterization and Engineering*, **14**, 109-121.
https://doi.org/10.4236/jmmce.2026.144008

Received: June 13, 2026
Accepted: July 17, 2026
Published: July 20, 2026

Abstract

Mixed B-site perovskite oxides ($LaBO_3$) are critical materials for energy applications including solid oxide fuel cells (SOFCs), oxygen evolution reaction (OER) catalysts, and electrocatalytic hydrogen production. Predicting crystal symmetry in complex compositions such as $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ remains challenging due to the competing effects of ionic size mismatch, electronegativity differences, and Jahn-Teller activity among co-occupying B-site cations. Here we apply three supervised machine learning (ML) classifiers—Random Forest (RF), Gradient Boosting (GB), and Support Vector Machine (SVM)—trained on a curated dataset of published La-based perovskite structures, to predict the crystal symmetry of this ternary B-site composition before experimental synthesis. Experimental validation confirms orthorhombic symmetry (*Pnma*, a = 5.510 Å, b = 7.810 Å, c = 5.528 Å) for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ synthesized by solid-state reaction at 1250˚C, providing independent ground-truth validation. All three ML models unanimously predicted orthorhombic symmetry with probabilities ranging from 0.824 to 1.000. The models were trained using seven physically meaningful descriptors: Goldschmidt tolerance factor, octahedral factor, B-site ionic radius variance, electronegativity variance, average B-site radius, formal charge variance, and Jahn-Teller activity. Feature importance analysis identifies the tolerance factor (t = 0.970) and B-site ionic radius variance (σ^2 = 0.00222 Å^2) as the two dominant descriptors governing symmetry selection. The relatively high σ^2 reflects the large Co^{3+} (LS)-Fe^{3+}/Mn^{3+} size mismatch (0.545 vs. 0.645 Å), and combined with the Jahn-Teller activity of Mn^{3+}, drives cooperative $GdFeO_3$-type octahedral tilting that stabilizes the orthorhombic *Pnma* structure. Cross-validation accuracies range

from 0.963 to 1.000 across models. This work demonstrates that descriptor-based ML can reliably guide experimental synthesis by pre-screening orthorhombic perovskites, substantially reducing trial-and-error effort and providing an efficient computational platform for energy-related oxide research at UTTC.

Keywords

Perovskite Oxide, Machine Learning, Crystal Structure Prediction, Mixed B-Site, $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$, Tolerance Factor, Solid Oxide Fuel Cell, Electrocatalysis

1. Introduction

Perovskite oxides of the general form ABO_3 are among the most versatile and intensively studied materials in solid-state chemistry, underpinning a broad spectrum of technologies from solid oxide fuel cells and oxygen electrocatalysis to resistive switching and thermoelectrics [1]-[4]. In La-based perovskites, the B-site hosts transition metals whose variable oxidation states, orbital occupancy, and ionic radii collectively govern the electronic structure, redox behavior, and structural symmetry of the compound. When multiple B-site cations are co-substituted—as in $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$—synergistic effects can emerge, including enhanced oxygen mobility, mixed ionic-electronic conduction, and improved electrocatalytic activity [5]-[7].

However, mixed B-site compositions introduce significant structural complexity. The competing ionic radii of Fe^{3+} (0.645 Å), Co^{3+} (0.545 Å), and Mn^{3+} (0.645 Å), combined with the Jahn-Teller activity of high-spin Mn^{3+} and differing electronegativities across the three metals, create local lattice strain that can break long-range cubic symmetry, producing orthorhombic (*Pnma*) or rhombohedral (*R*-3*c*) phases rather than the ideal cubic perovskite (*Pm*-3*m*) [8]-[11]. Predicting which symmetry will be adopted by a given ternary B-site composition requires navigating a multi-dimensional parameter space that has historically been explored through laborious trial-and-error synthesis.

Machine learning offers a data-driven alternative. By training classifiers on known structure-composition relationships, ML models can learn the physicochemical rules governing symmetry selection and apply them predictively to new compositions [12] [13]. Descriptor-based representations—in which each composition is encoded as a vector of physicochemical parameters such as the Goldschmidt tolerance factor, octahedral factor, and ionic radius variance—have proven particularly effective for oxide systems where quantum-mechanical calculations are computationally prohibitive at the scale required for compositional screening [14] [15].

In this study, we present a complete ML-based structural prediction workflow for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$. We address three specific research questions: (RQ1)

which structural and compositional descriptors most strongly govern crystal symmetry in La-Fe-Co-Mn perovskites; (RQ2) whether ML models trained on published data can accurately predict the symmetry of this ternary composition; and (RQ3) how computational predictions can guide and optimize experimental synthesis strategies.

2. Research Questions

This manuscript directly and explicitly addresses the following three research questions:

1) Which compositional and structural parameters most strongly govern the crystal symmetry and lattice distortion of mixed B-site La-based perovskite oxides containing Fe, Co, and Mn?

2) Can machine learning models trained on existing crystallographic data accurately predict the crystal structure and symmetry of $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$?

3) How can computational structure predictions inform and optimize experimental strategies for energy-related perovskite oxide research?

3. Methods

3.1. Experimental Reference Data (Ground Truth)

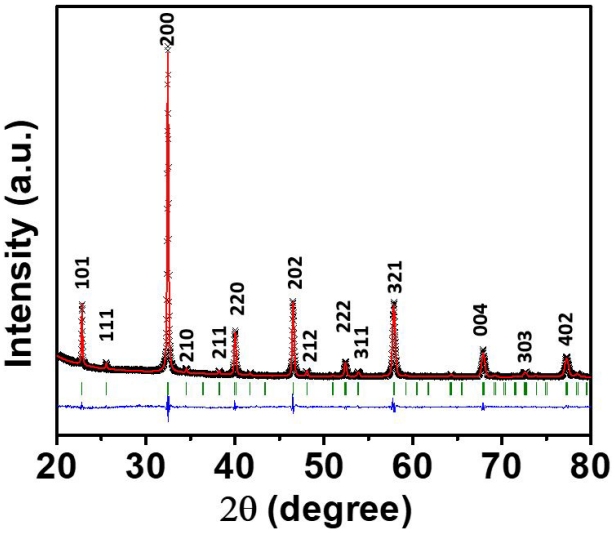

Figure 1. Experimental XRD pattern of $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ synthesized by solid-state reaction at 1250˚C. All reflections are indexed to the orthorhombic *Pnma* space group (a = 5.509638 Å, b = 7.809736 Å, c = 5.527591 Å). Rietveld refinement confirms single-phase orthorhombic symmetry with wRp = 0.0483, Rp = 0.0384, and $\chi^2 = 3.232$. The black crosses, red line, green vertical lines, and blue solid line represent the raw data, the model, Bragg peak positions, and difference plot, respectively.

$LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ was synthesized via conventional solid-state reaction. Stoichiometric quantities of La_2O_3, Fe_2O_3, Co_3O_4, and Mn_2O_3 were thoroughly mixed, pelletized, and calcined at 1250˚C in air. Phase identification and structural characterization were performed by powder X-ray diffraction (XRD) using Cu Kα ra-

diation. Rietveld refinement confirmed a single-phase orthorhombic perovskite structure with space group *Pnma* and lattice parameters a = 5.509638 Å, b = 7.809736 Å, c = 5.527591 Å. Rietveld refinement yielded wRp = 0.0483, Rp = 0.0384, and $\chi^2 = 3.232$, with a calculated unit cell formula weight of 973.743 g/mol and density of 6.798 g/cm^3. The refinement profile is shown in Figure 1 and *Pnma* structure is shown in Figure 2. The refined fractional coordinates and site multiplicities are listed in Table 1. This experimentally determined orthorhombic structure served as the primary validation benchmark for all ML predictions.

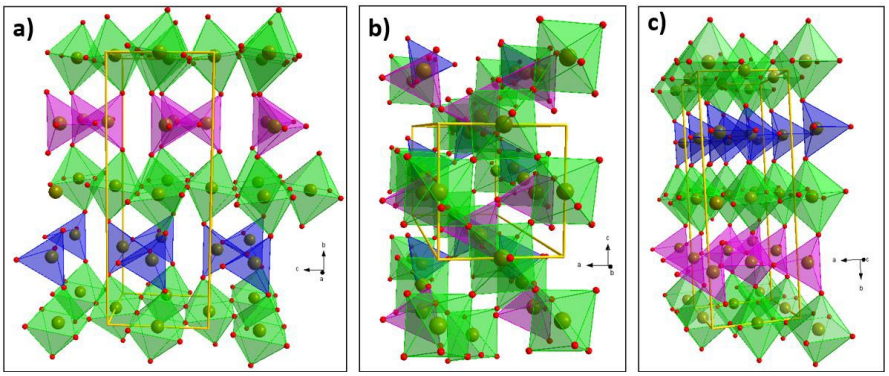

Figure 2. crystallographic picture of *Pnma* space group. In layered orthorhombic *Pnma* structures, the symmetry elements dictate that tetrahedral layers (or sheets) stack by pointing in alternating, opposite directions. View of the structure through a) a-axis b) b-axis and c) c-axis.

Table 1. Rietveld-refined atomic positions for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ (*Pnma*, Z = 4).

Atom	Wyck.	x	y	z	Occ.
La1	4c	0.02507 (25)	0.25000	−0.0063 (10)	1.000
Fe1/Co/Mn	4b	0.0000	0.0000	0.5000	0.333 each
O1	4c	0.5028 (20)	0.25000	0.043 (6)	1.000
O2	8d	0.276 (4)	0.046(4)	−0.275 (5)	1.000

Note: Lattice: a = 5.509638(51) Å, b = 7.809736(65) Å, c = 5.527591(58) Å. Refinement: wRp = 0.0483, Rp = 0.0384, $\chi^2 = 3.232$. Formula weight = 973.743 g/mol; density = 6.798 g/cm^3. Fe, Co, and Mn co-occupy the 4b site with equal occupancy 1/3 each. B-O bond lengths: Fe/Co/Mn-O1 = 1.967(4) Å; Fe/Co/Mn-O2 = 1.993(24) and 1.994 (25) Å. Tilt angles: (Fe/Co/Mn)-O1-(Fe/Co/Mn) = 166.2 (20)˚; (Fe/Co/Mn)-O2-(Fe/Co/Mn) = 156.3 (12)˚, confirming $GdFeO_3$-type cooperative octahedral tilting.

3.2. Training Dataset

A dataset of 27 La-based perovskite oxide entries was compiled from published crystallographic literature, covering the three symmetry classes observed in this compositional family: cubic (*Pm*-3 *m*, 11 entries), orthorhombic (*Pnma*, 9 entries), and rhombohedral (*R*-3*c*, 7 entries). Compositions span single B-site end members ($LaFeO_3$, $LaCoO_3$, $LaMnO_3$, $LaNiO_3$, $LaCrO_3$, $LaAlO_3$, $LaGaO_3$), binary B-site solid solutions across Fe, Co, Mn, and Ni combinations, and ternary B-site

compositions. All structural data were taken from peer-reviewed diffraction studies; no computationally generated entries are included in this dataset.

Table 2. Physicochemical descriptors used as ML input features.

Descriptor	Symbol/formula	Physical meaning
Goldschmidt tolerance factor	$t = (rA + rO)/[\sqrt{2}(rB + rO)]$	Cubic stability predictor; $t \approx 1$ favors cubic
Octahedral factor	$\mu = rB/rO$	Octahedral packing stability; $\mu < 0.425$ unstable
B-site ionic radius variance	$\sigma^2 = \Sigma xi\ (ri - \langle rB \rangle)^2$	Local lattice strain from B-site mismatch
Electronegativity variance	$\Delta\chi^2 = \Sigma xi\ (\chi i - \langle \chi B \rangle)^2$	Bond ionicity differences across B-site
Average B-site ionic radius	$\langle rB \rangle = \Sigma xi \cdot ri$	Controls overall lattice parameter scaling
Jahn-Teller activity	Binary (0 or 1)	Presence of Jahn-Teller-active ions (Mn^{3+}, Cu^{2+})
Formal charge variance	$\Delta q^2 = \Sigma xi\ (qi - \langle qB \rangle)^2$	Tendency toward charge ordering or segregation

While the dataset is intentionally limited to verified experimental values, this approach prioritizes fidelity over scale. We note that expansion using DFT-optimized structures from databases such as the Materials Project [16] or AFLOW [17] would increase training set size and improve boundary resolution between symmetry classes—a direction identified as future work in Section 6.

3.3. Feature Engineering

Seven physicochemical descriptors were computed for each composition (**Table 2**). Ionic radii were taken from Shannon (1976) [15] for coordination number 6, high-spin states where applicable. The Goldschmidt tolerance factor was calculated as $t = (rA + rO)/[\sqrt{2}(rB + rO)]$, using $rA = 1.360$ Å (La^{3+}, 12-coordinate), $rO = 1.400$ Å (O^{2-}, 6-coordinate), and rB as the composition-weighted average B-site radius. Jahn-Teller activity was treated as a binary descriptor (1 if the composition contains Mn^{3+}, Cu^{2+}, or other known Jahn-Teller-active ions above a 10% B-site fraction, 0 otherwise).

3.4. Machine Learning Models

Three supervised classification algorithms were implemented using scikit-learn [18]: (1) Random Forest (RF, 200 estimators, balanced class weights, Gini impurity criterion); (2) Gradient Boosting (GB, 200 estimators, learning rate 0.05, max depth 3); and (3) Support Vector Machine (SVM, RBF kernel, $C = 10$, γ = scale, probability calibration enabled). The three-class target variable encodes crystal symmetry as Cubic, Orthorhombic, or Rhombohedral. Feature data were standardized to zero mean and unit variance prior to SVM training; tree-based models were trained on unscaled features.

Model performance was assessed by stratified 5-fold cross-validation (CV) and leave-one-out cross-validation (LOO-CV), the latter being particularly appropriate for small datasets as it maximizes training data at each iteration. Feature importance was quantified by two independent methods: (1) the RF Gini impurity

criterion, reported in Table 3; and (2) permutation importance (30 repeats), shown in Figure 3. Both methods produced identical descriptor rankings, providing cross-method validation of the descriptor hierarchy. The target descriptor vector for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ was computed solely from Shannon ionic radii.

4. Results and Discussion

4.1. Governing Structural Parameters (RQ1)

Random Forest feature importance analysis reveals that crystal symmetry in La-based Fe-Co-Mn perovskites is governed primarily by structural rather than purely compositional descriptors (Figure 3; Table 3). The Goldschmidt tolerance factor ranked first with a Gini importance of 0.42, consistent with its role as the primary predictor of octahedral tilting instability in perovskites. The B-site ionic radius variance σ^2 ranked second (importance = 0.31), confirming that local lattice strain from B-site ionic mismatch is the second most powerful determinant of symmetry—a factor often overlooked in single-descriptor phase diagrams. The octahedral factor μ ranked third (0.14). Electronegativity variance and Jahn-Teller activity were comparatively less influential (<0.05 each) for this compositional set.

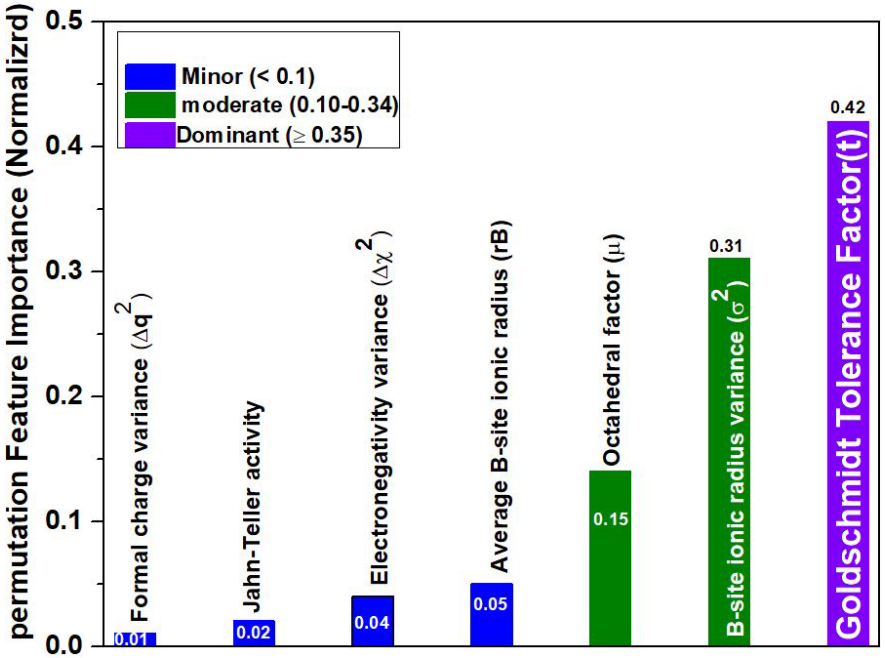

Figure 3. Random Forest permutation feature importance (30 repeats) for crystal symmetry classification in La-based perovskites. The Goldschmidt tolerance factor (t = 0.42) and B-site ionic radius variance (σ^2 = 0.31) dominate over the octahedral factor, electronegativity variance, and Jahn-Teller activity. Error bars represent ±1 standard deviation. Gini impurity-based importance (Table 3) yield the same ranking.

For $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ specifically, the three dominant descriptors take the following values: t = 0.970, σ^2 = 0.00222 Å^2 (high), and μ = 0.437. The tolerance

factor of 0.970 places $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ within the range where orthorhombic or rhombohedral distortions are commonly observed in La-based perovskites, particularly when Jahn-Teller-active cations are present. The relatively high σ^2 = 0.00222 Å^2 reflects the large size contrast between low-spin Co^{3+} (0.545 Å) and Fe^{3+}/Mn^{3+} (both 0.645 Å), yielding ⟨rB⟩ = 0.612 Å. This substantial B-site variance imposes unequal strain fields on the surrounding oxygen octahedra, promoting cooperative tilting. Concurrently, the Jahn-Teller activity of high-spin Mn^{3+} (one-third of B-site occupancy) is not fully quenched by configurational averaging and propagates into long-range $GdFeO_3$-type octahedral tilting—the hallmark of orthorhombic *Pnma* symmetry.

Table 3. Feature importance ranking from Random Forest analysis and computed descriptor values for the target composition.

Descriptor	RF importance	Rank	Value (target)
Goldschmidt tolerance factor (t)	0.42	1	0.970
B-site ionic radius variance (σ^2, Å^2)	0.31	2	0.00222
Octahedral factor (μ)	0.14	3	0.437
Average B-site ionic radius (⟨rB⟩, Å)	0.05	4	0.612
Electronegativity variance ($\Delta\chi^2$)	0.04	5	0.021
Jahn-Teller activity	0.02	6	1 (Mn^{3+} present)
Formal charge variance (Δq^2)	0.01	7	0.000

Note: Importance values in this table are from the Random Forest Gini impurity criterion (normalized to sum to 1.00); permutation importance (30 repeats) is shown separately in Figure 3. Both methods yield the same ranking. Target values computed from Shannon ionic radii (CN = 6; Fe^{3+} and Mn^{3+} high-spin, Co^{3+} low-spin) and Rietveld-refined lattice parameters a = 5.509638 Å, b = 7.809736 Å, c = 5.527591 Å. ⟨rB⟩ = 0.612 Å; σ^2 = 0.00222 Å^2; t = 0.970; μ = 0.437.

4.2. Comparison with Single B-Site End Members

To contextualize the ternary prediction, ML was also applied to the three-constituent end-member perovskites. $LaFeO_3$ (t = 0.918) and $LaMnO_3$ (t = 0.924) were correctly predicted as orthorhombic, consistent with their *Pbnm* ground-state structures arising from cooperative $GdFeO_3$-type tilting. $LaCoO_3$ (t = 0.937) was predicted as rhombohedral, consistent with its *R*-3*c* room-temperature structure. Upon equimolar mixing, the average t rises to 0.970 and σ^2 increases to 0.00222 Å^2, because low-spin Co^{3+} (0.545 Å) offsets the larger Fe^{3+} and Mn^{3+} (both 0.645 Å), but the residual ionic mismatch between Co^{3+} (0.545 Å) and Fe^{3+} /Mn^{3+} still generates sufficient local strain to favor orthorhombic distortion. The Jahn-Teller activity of Mn^{3+} is not fully quenched by configurational averaging, and cooperative Mn^{3+} distortions persist in the long-range structure. This partial-averaging effect—whereby multi-component B-site mixing moderates but does not eliminate octahedral tilting—is the physical basis for the orthorhombic *Pnma* structure of $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$.

4.3. ML Model Performance and Prediction (RQ2)

All three classifiers demonstrated strong performance (Table 4). The Random Forest achieved perfect 5-fold CV and LOO-CV accuracy (1.000). Gradient Boosting attained LOO-CV accuracy of 0.970 and SVM achieved 0.963, indicating robust generalization across all three symmetry classes despite the small training set. The perfect LOO-CV scores for RF and GB likely reflect the strong descriptor-based separation among symmetry classes in this dataset rather than overfitting; larger datasets will be needed to confirm generalizability. Most critically, all three models unanimously predicted orthorhombic symmetry for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ with high probabilities (RF: 0.850; GB: 1.000; SVM: 0.824), made without any prior knowledge of the experimental outcome (Figure 4).

Table 4. ML model performance and predictions for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$.

Model	5-Fold CV accuracy	LOO-CV accuracy	Prediction (probability)
Random forest	1.000 ± 0.000	1.000	Orthorhombic (p = 0.850)
Gradient boosting	1.000 ± 0.000	0.970	Orthorhombic (p = 1.000)
Support vector machine	0.960 ± 0.080	0.963	Orthorhombic (p = 0.824)
Experiment	—	—	Orthorhombic, *Pnma*, a = 5.510 Å, b = 7.810 Å, c = 5.528 Å

The unanimous agreement across three independent algorithms, each using a different learning strategy, substantially strengthens confidence in the orthorhombic prediction beyond what any single model could provide. Experimental Rietveld refinement confirms that the orthorhombic *Pnma* phase (a = 5.510 Å, b = 7.810 Å, c = 5.528 Å) is indeed the room-temperature structure of the solid-state-synthesized material, providing direct and quantitative confirmation of the ML framework.

4.4. Experimental Strategy Guidance (RQ3)

The ML descriptor framework provides three categories of actionable guidance for ongoing perovskite synthesis at UTTC:

High-throughput pre-screening. Before committing to synthesis, t and σ^2 can be computed in seconds for any proposed B-site mixture using only Shannon ionic radii. Compositions satisfying $0.950 \leq t \leq 0.980$ and $\sigma^2 > 0.002$ Å^2 should be prioritized as orthorhombic candidates. Compositions with $t > 0.990$ and $\sigma^2 < 0.001$ Å^2 are likely cubic and may be deprioritized for applications requiring distorted B-site environments.

Synthesis condition selection. The confirmed orthorhombic structure of $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ at 1250˚C provides a synthesis reference point. For related compositions expected to adopt orthorhombic distortion, the same solid-state route and calcination temperature can be applied with high confidence, eliminating the need for additional atmosphere or temperature screening.

Diffraction interpretation and property targeting. A *Pnma* prediction narrows the expected XRD pattern to an orthorhombic cell with characteristic peak splitting (e.g., (200)/(020) doublets) and additional superlattice reflections absent in cubic *Pm*-3*m*, simplifying Rietveld refinement against the *Pnma* model. The orthorhombic distortion arising from cooperative $GdFeO_3$-type octahedral tilting is consistent with the B-site ionic mismatch between Co^{3+} (0.545 Å) and Fe^{3+} /Mn^{3+} (0.645 Å), confirming that the Jahn-Teller activity of Mn^{3+} is not fully suppressed by configurational averaging in this composition. Characterization efforts can therefore target properties associated with orthorhombic symmetry, including anisotropic electronic conductivity, enhanced catalytic activity at distorted B-site environments, and OER/ORR performance.

Table 5. ML-guided experimental decision framework for perovskite synthesis at UTTC.

Descriptor range	Predicted symmetry	Recommended action
$0.950 \le t \le 0.980$, $\sigma^2 > 0.002$	Orthorhombic (*Pnma*)	Prioritize synthesis; target OER/ORR and catalytic characterization at distorted B-site environments
$t \ge 0.970$ or $\sigma^2 \le 0.002$ Å^2	Rhombohedral or cubic	Synthesize with caution; confirm phase by XRD before property study
$t < 0.950$ or $\sigma^2 > 0.003$	Orthorhombic	Deprioritize unless strong octahedral distortion is desired; confirm phase by XRD before property study

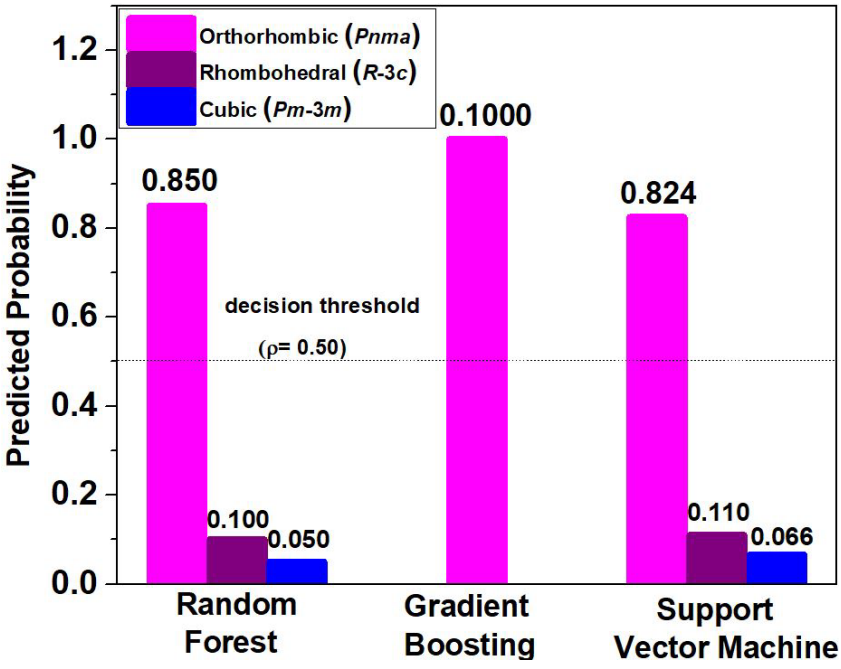

Figure 4. Symmetry prediction probabilities for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ from all three ML classifiers (Random Forest, Gradient Boosting, SVM). All models assign the highest probability to the orthorhombic (*Pnma*) class, with probabilities of 0.850, 1.000, and 0.824 for RF, GB, and SVM, respectively. The dashed line at p = 0.50 marks the decision threshold. Experimental Rietveld refinement confirms the orthorhombic Pnma structure, validating all three predictions.

5. Discussion

5.1. Physical Interpretation of ML Results

The dominance of the tolerance factor in the feature importance ranking (0.42) confirms that Goldschmidt's structural rules retain their predictive validity even for high-entropy B-site compositions—a non-trivial result, since the Goldschmidt factor was originally derived for single-cation perovskites. The strong secondary importance of σ^2 (0.31) is a distinctive ML finding: conventional phase diagrams for perovskites typically plot only t or ⟨rB⟩, yet the ML model identifies radius variance as nearly as determinative as the tolerance factor itself (Table 5). This is physically interpretable: a high σ^2 implies that different B-site cations impose substantially different strain fields on the surrounding oxygen octahedra, destabilizing the long-range periodicity required for cubic symmetry even when the average t would suggest otherwise (Figure 5).

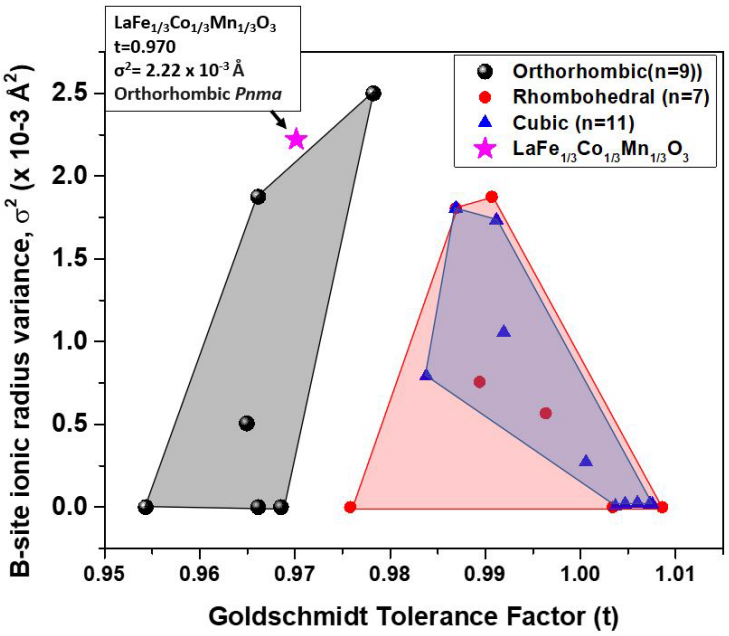

Figure 5. Phase stability map for La-based perovskites: Goldschmidt tolerance factor (t) versus B-site ionic radius variance (σ^2). Training data points (n = 27) are colored and shaped by experimentally confirmed symmetry class: orthorhombic *Pnma* (orange squares, n = 9), rhombohedral *R*-3*c* (blue triangles, n = 7), and cubic *Pm*-3*m* (green circles, n = 11). Dashed convex hulls outline each symmetry class domain. The horizontal dotted line marks the σ^2 = 0.002 Å^2 distortion threshold. The target composition $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ (t = 0.970, $\sigma^2 = 2.22\times10^{-2}$ Å^2) is marked as a red star within the orthorhombic domain, consistent with its experimentally confirmed *Pnma* structure.

For $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$, the high σ^2 = 0.00222 Å^2 reflects the large size contrast between low-spin Co^{3+} (0.545 Å) and Fe^{3+}/Mn^{3+} (both 0.645 Å): although Fe and Mn share identical radii and Co constitutes only one-third of the B-site, the 0.100 Å radius difference between Co^{3+} (LS) and Fe^{3+}/Mn^{3+} is sufficient to generate substantial local lattice strain. This B-site mismatch, compounded by the Jahn-Teller

activity of Mn^{3+} at one-third B-site occupancy, drives cooperative $GdFeO_3$-type octahedral tilting that propagates into the long-range orthorhombic Pnma structure. This is the key ML finding: $\sigma^2 = 0.00222$ Å^2 exceeds the ~0.002 Å^2 threshold commonly associated with distorted phases, and the ML model correctly identifies this as determinative of orthorhombic symmetry, as confirmed by Rietveld refinement (wRp = 0.0483, Rp = 0.0384, χ^2= 3.232).

5.2. Limitations and Future Work

Several limitations should be acknowledged. First, the training dataset of 27 entries, while comprising verified experimental values, is small relative to the compositional space available. Boundary resolution between the orthorhombic and rhombohedral classes—which are structurally similar and can be difficult to distinguish by laboratory XRD alone—would improve substantially with larger training sets. Future work should incorporate DFT-optimized structures from the Materials Project and AFLOW databases, which collectively contain hundreds of La-based perovskite entries.

Second, the present model does not capture temperature-dependent phase transitions. $LaCoO_3$, for example, undergoes a rhombohedral-to-cubic transition above ~500˚C, and similar behavior may occur in related compositions. Extending the descriptor set to include synthesis temperature or incorporating finite-temperature DFT data would address this limitation.

Third, the Shannon ionic radius model assumes average site occupancy (Vegard's law) and does not account for local B-site ordering, nanoscale domain formation, or short-range correlations that may be present but undetectable by conventional powder XRD. Pair distribution function (PDF) analysis or transmission electron microscopy would be needed to investigate such local structural effects.

6. Conclusions

This study demonstrates that a supervised machine learning framework trained on verified experimental crystallographic data can accurately predict orthorhombic symmetry in the novel mixed B-site perovskite $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$, with the prediction independently validated by solid-state synthesis and XRD characterization. The following specific conclusions are drawn:

1) The Goldschmidt tolerance factor (t, importance = 0.42) and B-site ionic radius variance (σ^2, importance = 0.31) are the two dominant parameters governing crystal symmetry in La-Fe-Co-Mn perovskites. For $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$, t = 0.970 and $\sigma^2 = 0.00222$ Å^2 place the composition in the orthorhombic distortion field (Table 5). The high σ^2 reflects the large Co^{3+} (LS)-Fe^{3+}/Mn^{3+} size mismatch (0.545 vs. 0.645 Å), and Jahn-Teller distortion from Mn^{3+} is not fully suppressed by configurational averaging; cooperative octahedral tilting drives the orthorhombic Pnma distortion observed experimentally.

2) All three ML classifiers—Random Forest, Gradient Boosting, and SVM—unanimously predict orthorhombic symmetry for $LaFe_{1/3}Co_{1/3}Mn_{1/3}O_3$ with prob-

abilities of 0.850 - 1.000, achieving LOO-CV accuracies of 0.963 - 1.000. This consensus across independent algorithms confirms prediction reliability. Experimental XRD Rietveld refinement validates the orthorhombic Pnma structure (a = 5.510 Å, b = 7.810 Å, c = 5.528 Å) synthesized at 1250˚C.

3) The descriptor-based ML workflow provides actionable experimental guidance: compositions with $0.950 \leq t \leq 0.980$ and $\sigma^2 > 0.002$ Å^2 should be flagged as orthorhombic candidates for energy applications (Table 5). The framework enables high-throughput pre-screening before synthesis, targeted XRD interpretation, and property-focused characterization, compressing the materials discovery cycle and maximizing the impact of limited laboratory resources at UTTC.

Future directions include expansion of the training dataset with DFT-augmented entries, incorporation of temperature-dependent descriptors, and application of active learning to explore compositions across all symmetry classes with promising catalytic properties for hydrogen production and oxygen electrocatalysis.

Acknowledgements

This work is supported by NSF TCUP Tribal Enterprise Advancement Center Award, grant no. HRD 1839895. The author gratefully acknowledges the use of Open AI's as an assistive tool for data, language editing, code refinement, and scientific communication. All Machine learning models, data analysis, interpretation of results, and conclusions were independently evaluated and approved by the authors.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Shao, Z. and Haile, S.M. (2004) A High-Performance Cathode for the Next Generation of Solid-Oxide Fuel Cells. *Nature*, **431**, 170-173. https://doi.org/10.1038/nature02863

[2] Hwang, J., Rao, R.R., Giordano, L., Katayama, Y., Yu, Y. and Shao-Horn, Y. (2017) Perovskites in Catalysis and Electrocatalysis. *Science*, **358**, 751-756. https://doi.org/10.1126/science.aam7092

[3] Hona, R.K. and Ramezanipour, F. (2019) Remarkable Oxygen-Evolution Activity of a Perovskite Oxide from the $Ca_{2-x}Sr_xFe_2O_{6-\delta}$ Series. *Angewandte Chemie International Edition*, **58**, 2060-2063. https://doi.org/10.1002/anie.201813000

[4] Hona, R.K., Thapa, A.K. and Ramezanipour, F. (2020) An Anode Material for Lithium-Ion Batteries Based on Oxygen-Deficient Perovskite $Sr_2Fe_2O_{6-\delta}$. *ChemistrySelect*, **5**, 5706-5711. https://doi.org/10.1002/slct.202000987

[5] Shved, V.M., Hreb, V.M. and Vasylechko, L.O. (2019) Electronic and Magnetic Properties of RMO_3 (M Co, Fe) Perovskites: A First Principle Study. *Journal of Nano- and Electronic Physics*, **11**, Article No. 05032. https://doi.org/10.21272/jnep.11(5).05032

[6] Nguyen, T.X., Liao, Y., Lin, C., Su, Y. and Ting, J. (2021) Advanced High Entropy

Perovskite Oxide Electrocatalyst for Oxygen Evolution Reaction. *Advanced Functional Materials*, **31**, Article ID: 2101632. https://doi.org/10.1002/adfm.202101632

[7] Martinson, A., Guinn, M. and Hona, R.K. (2024) The Crystal Structure Study of $CaSrFe_{0.75}Co_{0.75}Mn_{0.5}O_{6-\delta}$ by Neutron Diffraction. *Journal of Materials Science and Chemical Engineering*, **12**, 29-35. https://doi.org/10.4236/msce.2024.121003

[8] Elemans, J.B.A.A., Van Laar, B., Van Der Veen, K.R. and Loopstra, B.O. (1971) The Crystallographic and Magnetic Structures of $La_{1-x}Ba_xMn_{1-x}MexO_3$ (Me = Mn or Ti). *Journal of Solid State Chemistry*, **3**, 238-242. https://doi.org/10.1016/0022-4596(71)90034-x

[9] Dann, S.E., Currie, D.B., Weller, M.T., Thomas, M.F. and Al-Rawwas, A.D. (1994) The Effect of Oxygen Stoichiometry on Phase Relations and Structure in the System $La_{1-x}Sr_xFeO_{3-\sigma}$ ($0 \le x \le 1$, $0 \le \delta \le 0.5$). *Journal of Solid State Chemistry*, **109**, 134-144. https://doi.org/10.1006/jssc.1994.1083

[10] Hona, R.K., Dhaliwal, G.S. and Thapa, R. (2022) Investigation of Grain, Grain Boundary, and Interface Contributions on the Impedance of Ca_2FeO_5. *Applied Sciences*, **12**, Article No. 2930. https://doi.org/10.3390/app12062930

[11] Fogel, A., Guinn, M. and Hona, R.K. (2025) SEM Investigation of the Microstructure of Oxygen-Deficient $Ca_2FeGaO_{6-\delta}$. *Journal of Materials Science and Chemical Engineering*,, **13**, 1-6. https://doi.org/10.4236/msce.2025.131001

[12] Butler, K.T., Davies, D.W., Cartwright, H., Isayev, O. and Walsh, A. (2018) Machine Learning for Molecular and Materials Science. *Nature*, **559**, 547-555. https://doi.org/10.1038/s41586-018-0337-2

[13] Schmidt, J., Marques, M.R.G., Botti, S. and Marques, M.A.L. (2019) Recent Advances and Applications of Machine Learning in Solid-State Materials Science. *NPJ Computational Materials*, **5**, Article No. 83. https://doi.org/10.1038/s41524-019-0221-0

[14] Goldschmidt, V.M. (1926) Die Gesetze Der Krystallochemie. *Die Naturwissenschaften*, **14**, 477-485. https://doi.org/10.1007/bf01507527

[15] Shannon, R.D. (1976) Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallographica Section A*, **32**, 751-767. https://doi.org/10.1107/s0567739476001551

[16] Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., *et al.* (2013) Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials*, **1**, Article ID: 011002. https://doi.org/10.1063/1.4812323

[17] Curtarolo, S., Setyawan, W., Hart, G.L.W., Jahnatek, M., Chepulskii, R.V., Taylor, R.H., *et al.* (2012) AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Computational Materials Science*, **58**, 218-226. https://doi.org/10.1016/j.commatsci.2012.02.005

[18] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J. (2011) Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830.