Digest Journal of Nanomaterials and Biostructures

Dig. J. Nanomater. Bios. **2026**; 21(3): 57746
https://doi.org/10.31083/DJNB57746

Original Research

Machine Learning–Assisted Exploration of the Structural, Dielectric, and Ferroelectric Properties of $CaTiO_3$ Nanoceramics

Dheeraj Kumar[1], Subhash Chandra Pandey[1,*], Paramjit Kour[2,*]

[1]Department of Computer Science and Engineering (CSE), Birla Institute of Technology, 835215 Ranchi, India
[2]Department of Physics, Birla Institute of Technology, 835215 Ranchi, India
*Correspondence: s.pandey@bitmesra.ac.in (Subhash Chandra Pandey); paramjit.kour@bitmesra.ac.in (Paramjit Kour)
Academic Editor: Carmen-Georgeta Ristoscu
Submitted: 29 January 2026 Revised: 29 April 2026 Accepted: 11 May 2026 Published: 4 September 2026

Abstract

Background: The structural and electrical properties of Pb-free calcium titanate ($CaTiO_3$) nanoceramics were investigated using a combined experimental and machine learning approach. **Methods:** X-ray diffraction analysis with Rietveld refinement revealed the formation of an orthorhombic perovskite phase with the Pnma space group and the influence of annealing temperature on crystallite size and lattice strain. Dielectric measurements revealed strong frequency-dependent dispersion, while polarization–electric field hysteresis loops showed an increase in remanent polarization and coercive field with increasing annealing temperature. **Results:** Machine learning models were used to predict key structural and electrical properties. Among the evaluated models, XGBoost achieved the highest predictive performance for ferroelectric properties, with a coefficient of determination (R^2) of 0.992, while the remaining machine learning models also demonstrated strong predictive capability with R^2 values exceeding 0.96. **Conclusions:** These predictions strongly agreed with experimental observations, which demonstrated the ability of machine learning models to capture complex structure–property relationships and guide the optimization of synthesis conditions for $CaTiO_3$ nanoceramics.

Keywords: $CaTiO_3$ ceramics; dielectric properties; ferroelectric hysteresis; machine learning model; structure–property correlation

1. Introduction

Nowadays, researchers have explored the use of perovskite oxides in various fields such as microelectronics, energy storage devices, and sensors etc. [1,2,3]. This interest arises from their structural flexibility and their unique dielectric and ferroelectric properties. Among these materials, calcium titanate ($CaTiO_3$) has emerged as promising candidate because of its favorable dielectric behavior and environmentally friendly nature [4]. In addition, lattice distortion in $CaTiO_3$ may induce ferroelectricity under suitable condition [5]. However, achieving reliable control over its dielectric and ferroelectric performance at the nanoscale remains challenging [6]. Cu-doped $CaTiO_3$ has also been reported to exhibit a very high dielectric constant. Furthermore, $CaTiO_3$ is highly sensitive to minor structural modifications, making it a suitable material for advanced technological applications [7].

The integration of data-driven approaches and smart optimization tools has led to recent advancements in materials engineering. This helps in the discovery of new materials as well as improves the performance of existing materials. Further, the data-driven machine learning (ML) model has been successful in revealing the hidden correlations within the complicated datasets and thus helps optimize material systems. For example, data-driven methods have been used to analyze fuel injection fluctuations and optimize the performance of dual-fuel engines through intelligent control strategies [8]. Similarly, ML methods are also effective to find the best performing functional molecules for biomedical and optical uses much faster by moving systematically through vast chemical design spaces [9]. Such research demonstrates how ML is becoming an indispensable tool to analyze the complex multivariable systems and recognize the set of design parameters that lead to the best performance.

The dielectric and ferroelectric features of $CaTiO_3$ strongly depend on its crystal structure and microstructure. Further, at room temperature, it crystallizes in an orthorhombic phase. Previous studies have reported that different factors like lattice parameters, octahedral tilting and grain boundary density collectively and individually play a crucial role in determining the dielectric response and ferroelectric switching [10,11]. Furthermore, various extrinsic influences (grain boundaries, internal stresses and defect dipoles) and intrinsic contributions (electronic polarization, lattice distortions) also contribute to dielectric relaxation and polarization in $CaTiO_3$ [12].

Recent research has highlighted the role of microstructural engineering and advanced materials optimization in enhancing the performance of functional materials. For example, studies on BiVO4-based materials have shown how microstructural evolution and kinetic processes can notably affect the optical and structural performance [13]. Likewise, alloy design methods based on interfacial engineering have been used to improve tribo-corrosion resistance of refractory alloys [14]. In addition, advanced modeling and

optimization methods have also been applied in energy systems, such as proton exchange membrane fuel cells. These methods have been employed to understand the characteristics of dynamic coupling and to improve the operational stability [15]. These studies show the increasing importance of integrating materials design, micro-structural control, and data-driven modeling to achieve improved functional performance. These developments support the application of ML approaches to investigate the structure-property relationships in $CaTiO_3$ nanoceramics.

These effects are strongly interrelated. Therefore, a deeper understanding beyond conventional measurements is required. It is true that traditional approaches are primarily based on detailed experimental work and first-principles simulations and provide important insights. However, these approaches are often associated with drawbacks such as they are often slow, labor-intensive, and computationally costly [16]. In recent years, ML has emerged as a powerful alternative approach. It is because that ML can capture complex correlations explicitly in large datasets. In addition, ML techniques enable faster and more accurate prediction of dielectric and ferroelectric properties in ceramic systems [17,18,19].

ML techniques contribute substantially to predicting the behavior of complex systems. Moreover, it is also used to make intelligent decisions in various fields. For instance, advanced ML models like Extreme Gradient Boosting (XGBoost) models combined with multi-feature fusion techniques have been effectively used to predict the hemodynamic conditions in biomedical systems [20]. Similarly, in [21], multimodal ML frameworks have been used to improve the accuracy of medical diagnosis. Further, in [22], data-driven methods have been used to optimize the smart energy systems. Indeed, these studies show the ability of ML methods to identify non-linear relationships in large datasets. These findings further support the application of ML techniques to deal with complex material systems, including perovskite oxides.

ML algorithms possess several features that improve predictive capability of the machine such as the ability to identify complex and non-linear relationships of the material properties and structural features. Therefore, as a result, ML approaches can reduce the experimental and computational demands by rapid screening and reliable prediction of functional behavior [23]. In perovskite oxides like $CaTiO_3$, ML has been effective in revealing structure-property relationships. Moreover, ML has also clarified how variations in different synthesis conditions, microstructural characteristics, and composition control the dielectric and ferroelectric behavior [24,25,26,27].

Further, researchers have also explored ensemble learning methods in this pursuit. It has been observed that ensembles based on Random Forest (RF) and gradient boosting demonstrate high accuracy for predicting dielectric constants. Previous studies have shown that Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) have been successfully applied to model the ferroelectric hysteresis behavior [28,29]. Furthermore, the integration of rich and effective datasets with ML techniques can enable robust data-driven frameworks for different applications such as material design. This integration speeds up the discovery and optimization of advanced ceramic systems [30,31].

Previous ML studies on perovskite materials [32,33] have predominantly used large datasets with electronic structure descriptors, compositional variables, and density functional theory (DFT)-derived features to predict material properties. Although these methods give valuable insights, they usually require costly computations or very large materials databases. In contrast, the present study focuses on structural and processing descriptors that is directly measurable in experiments and thus enables the development of predictive models that are readily applicable to experimental materials optimization. The present approach of leveraging experimental structural parameters along with ML aims to build a workable data-driven toolkit for interpreting and forecasting the dielectric and ferroelectric properties of $CaTiO_3$ nanoceramics.

However, the exploration of $CaTiO_3$ nanoceramics has remained relatively limited despite recent progress in ML-assisted materials research. The majority of research focuses on broader perovskite families. Thereby creating a gap in targeted predictions and mechanistic understanding of $CaTiO_3$-specific dielectric and ferroelectric behavior [34]. In this context, the present work aims to systematically investigate the structural, dielectric, and ferroelectric properties of $CaTiO_3$ nanoceramics. The objective of present work is to extend the application of ML algorithms in the domain of model prediction and to impart insights into structure-property relationships. Indeed, the objective of this study is twofold. These are listed below:

- To develop and train the ML models capable of reliably predicting structural, dielectric, and ferroelectric responses.
- To evaluate the performance of different ML algorithms and interpret their relevance to the underlying physical mechanisms.

Integrating experimental materials research with ML-based predictive modeling is a very promising approach for rapid materials optimization. However, many existing studies mainly focus on large computational datasets or electronic structure descriptors, derived from DFT calculations and may not be easily accessed by experimental researchers [35,36]. On the other hand, the present study is centered on experimentally measurable structural, microstructural, and processing-related descriptors to prepare predictive ML models for $CaTiO_3$ nanoceramics. The work harnesses the power of combining experimentally obtained parameters with several ML algorithms to develop an effective framework to interpret the structure-property interrela-

tions governing the dielectric and ferroelectric properties. This method represents a data-driven tool to guide the tuning of the processing conditions and the functional properties of perovskite ceramics.

The novelty of this work lies in the prediction of $CaTiO_3$ with the help of ML-based approach and to establish a ML-based framework that can be extended to other functional oxides. In addition, this study bridges experimental investigations with data-driven modeling. The findings could enable rational design and optimization of dielectric and ferroelectric ceramics. This will enhance the applicability of such materials in different domains including capacitors, memory devices, and tunable microwave components [37,38].

2. Materials and Methods

The theoretical framework of this study is presented in this section. Further, basic concepts related to material are provided to facilitate understanding of structure-property relationships lucid. The polarization mechanisms and dielectric behavior in the material are also described. These concepts help explain the results of the different ML models. Moreover, this section also describes the relationship between microstructure and material properties. Subsequently, the next subsection explains the basic features of $CaTiO_3$.

2.1 Theoretical Framework

2.1.1 Fundamentals of $CaTiO_3$

As stated earlier, $CaTiO_3$ is a member of the perovskite oxide family. Its general formula is ABO_3. Here, Ca^{2+} is at the A-site, Ti^{4+} is at the B-site [5,39]. In [10], it has been reported that at room temperature $CaTiO_3$ crystallizes in an orthorhombic structure (space group *Pbnm*), featuring tilted TiO_6 octahedra and distorted Ca-O coordination polyhedra. Further, in [12], studies have shown that these lattice distortions crucially influence the electronic, dielectric, and ferroelectric behavior of the material. Mathematically, the structural stability of $CaTiO_3$ can be evaluated using the Goldschmidt tolerance factor (t). This factor is defined in Eqn. 1 as:

$$t = \frac{r_A + r_O}{\sqrt{2}\,(r_B + r_O)} \tag{1}$$

In [40], it has been given that r_A, r_B, and $r_O$denotes the ionic radii of the A-site cation, B-site cation, and oxygen anion, respectively. Further, in [41], it is reported that for $CaTiO_3$, the tolerance factor t <1, which reflects the octahedral tilting and stabilization of the orthorhombic phase. Furthermore, variations in t can be induced by doping or strain engineering. In [42], it is stated that such variations can also trigger phase transitions and thus influences dielectric and ferroelectric properties. Moreover, in [43], it is reported that at the nanoscale additional structural distortions may arise.

2.1.2 Dielectric and Ferroelectric Theories

In [44], it is reported that the multiple polarization mechanisms contribute to govern the dielectric behavior of $CaTiO_3$. These mechanisms include electronic, ionic, dipolar, and interfacial contributions. Further, in [45], it is reported that the frequency dependence of the dielectric constant (ε_r) can be described using Debye-type relaxation models. In this model, polarization lags behind the applied alternating electric field. Moreover, in [46], it is revealed that at low frequencies interfacial phenomena such as space-charge accumulation dominates the dielectric response. In contrast, primarily electronic and ionic polarization contribute at high frequencies [46]. The ferroelectricity in perovskites originates from non-centro-symmetric distortions of the crystal lattice [2]. This results in spontaneous polarization, which can be reversed by an external electric field [2]. In [4,47], it is stated that in $CaTiO_3$, intrinsic ferroelectricity is suppressed due to pronounced octahedral tilting. However, it can be induced by various factors like epitaxial strain, substitutional doping, or nanoscale size confinement. Furthermore, in [28], the Landau–Ginzburg–Devonshire (LGD) phenomenological theory provides a widely used framework for modeling ferroelectric phase transitions. In this theory, the free energy expansion is expressed in terms of polarization (P) as given in Eqn. 2:

$$F(P) = \alpha P^2 + P^4 + P^6 - EP \tag{2}$$

Here, α, β, γ are material-dependent coefficients and E is the applied external electric field [1]. This free energy model provides insight into the stability of polarization states and the magnitude of coercive fields as well. Further, the ferroelectric hysteresis loop is characterized by remnant polarization (Pr) and coercive field (Ec). These serve as a fundamental experimental measure of ferroelectric behavior [6]. Further, in [48], it is reported that the extrinsic factors play a significant role in determining the shape and stability of ferroelectric hysteresis loops in $CaTiO_3$ nanoceramics. These factors include wall motion, defect dipoles, and grain boundary effects.

2.2 Machine Learning Algorithms for Materials Prediction

ML has become increasingly important for predicting material properties and is now widely used in materials research. Moreover, it complements traditional methods and also provides a new approach for understanding materials. ML approaches complement first-principles calculations and also support phenomenological modeling [34]. ML models enable rapid prediction and can learn relationships between input and output data. The inputs include ionic radii and electronegativity differences. In addition,

lattice parameters, and processing conditions are also used as inputs. Further, dielectric constant, polarization, and band gap are commonly used as output parameters [26].

2.2.1 Supervised Learning Methods

Supervised learning techniques are widely used in materials research. This learning approach is especially useful for functional ceramics. Support Vector Machines (SVMs) are commonly used techniques of this paradigm and assist in predicting material properties. It has been reported in [28] that SVMs are effective in handling the nonlinear relationships and can establish the relationship between structural descriptors and dielectric constants. Moreover, ensemble learning methods are also used for prediction. These include Random Forest (RF) and Gradient Boosting Decision Trees (GBDT). The ensemble methods improve model reliability and they also reduce the risk of overfitting [25]. Additionally, these methods work well with high-dimensional datasets. Further, Artificial Neural Networks (ANNs) are another powerful approach [29]. Previous studies have shown that ANNs can model complex nonlinear behavior such as ferroelectric hysteresis loops [28].

2.2.2 Unsupervised Learning and Dimensionality Reduction

Unsupervised learning is an important part of ML and is also used in materials research. These techniques assist in analyzing material properties. K-means clustering is a commonly used unsupervised learning technique. Moreover, Principal Component Analysis (PCA) is also widely used. These methods help in finding hidden patterns in data and work well without labeled outputs. Further, PCA is useful for reducing data dimensions and to make data visualization easier [49]. Previous studies have shown that unsupervised learning techniques can also group materials with similar dielectric behavior [50].

2.2.3 Deep Learning and Graph-Based Models

Deep learning approaches are also play an important role in prediction of material properties. Different deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are effective methods for analyzing the sequential and spatial correlations in the dataset [51]. In [23], it is stated that Crystal Graph Convolutional Neural Networks (CGCNNs) represent the atomic structures as graphs. Thus, they enable the direct prediction of material properties from atomic arrangements without the need for handcrafted descriptors. Further, this method displays substantial potential to elucidate the structure and property based co-relationship in $CaTiO_3$ nanoceramics.

2.2.4 Model Evaluation and Interpretability

Different statistical metrics are used to evaluate the accuracy of ML-based property predictions. Previous studies commonly employ statistical metrics [52]. These are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). In addition, in [53], some interpretability techniques have also been proposed. These are Shapley additive explanations (SHAP) and feature importance ranking. These approaches provide valuable insights into the structural and compositional factors influencing material properties [53].

2.2.5 Integrating Physics and Data-Driven Approaches

Integration of physics and data-driven approaches also faces challenges. A major challenge in applying the ML techniques in material science is to ensure the physical interpretability as well as the model's generalizability [31]. These issues have been addressed in recent advancements by integrating the ML techniques with physics-based frameworks. One example of such an approach is the integration of Landau–Ginzburg–Devonshire (LGD) theory and DFT. This integration shows considerable promise to improve the prediction accuracy of the model [54]. Further, such integration incorporates domain knowledge into data-driven models. Thus, it enhances the prediction accuracy as well as reduces the risk of overfitting [55]. Further, physics-informed ML provides distinct advantages for $CaTiO_3$. This is because of the fact that the dielectric and ferroelectric behavior of $CaTiO_3$ results owing to a complex interplay between intrinsic lattice distortions, and extrinsic microstructural factors. It is also stated in [27] that combining experimental measurements, phenomenological modeling, and ML predictions offers a framework that bridges the current knowledge gaps. It also supports the rational design of advanced $CaTiO_3$-based dielectric and ferroelectric nanoceramics.

2.3 Methodology

A systematic and rigorous approach is essential for investigating the structural, dielectric, and ferroelectric properties of $CaTiO_3$ nanoceramics. Moreover, it is also crucial to effectively integrate ML techniques to predict material properties. In the present study, advanced ML algorithms are used to develop data-driven models of structural and functional properties of $CaTiO_3$ nanoceramics.

2.3.1 Dataset Preparation for Machine Learning

The dataset preparation process consists of structural and electrical properties. These are lattice parameters, crystallite size, micro-strain, dielectric constant (ε_r), tan δ, and ferroelectric parameters (P_r, P_s, E_c). In this dataset, each sample was represented by a feature vector that includes synthesis parameters. The feature vectors were prepared as reported in [56]. The synthesis parameters included are annealing temperature, and crystallite size, as well as the corresponding material properties. Further, data preprocessing has also been done and it includes normalization, outlier removal, and feature engineering. In addition, k-nearest

neighbor (k-NN) is also used to address the issue of missing values. Further, correlation matrices are constructed to evaluate interdependencies among input features in a way reported in [57]. In [58], the dataset was randomly partitioned into training (70%), validation (15%), and test (15%) subsets to ensure unbiased model evaluation. The same data split strategy was adopted in this study.

The complete list of descriptors used in the machine learning models is provided in **Supplementary Tables 1–5** of the Supplementary Information.

Dataset Size, Feature Structure, and Preprocessing Steps. A robust ML model depends strongly on the quality of dataset it is trained on. The dataset should adequately represent the main structural, microstructural, and processing variables that determine the material properties. In this work, the dataset is derived from the experimental work on $CaTiO_3$ nanoceramic samples prepared under different processing conditions. The dataset consists of 96 experimental samples. Each sample represents a unique combination of synthesis parameters along with the measured properties. For every single sample, feature vectors were constructed to capture variations in synthesis conditions, physical characteristics of the structure, and actual properties. The main descriptors used in the dataset are as follows:

- Annealing temperature.
- Annealing time.
- Crystallite size.
- Lattice strain.
- Lattice parameters (a, b, c).
- Dielectric constant (ε_r).
- Dielectric loss tangent (tan δ).
- Ferroelectric parameters such as remanent polarization (P_r), saturation polarization (P_s), and coercive field (E_c).

The descriptors listed above define the input features. These features are used in the ML models. These descriptors represent the important characteristics of the data and each feature has a clear physical significance. Moreover, they strongly affect the dielectric and ferroelectric properties of $CaTiO_3$ nanoceramics. In addition, these parameters are experimentally measurable. Calcium nitrate tetrahydrate (Merck KGaA, Darmstadt, Germany) and titanium isopropoxide (Sigma-Aldrich, St. Louis, MO, USA) were used as received. Lot numbers were not available because the chemicals were procured previously and the original packaging was no longer available.

Data preprocessing steps are completed before model training. The step helps in improving the model's robustness and accuracy. In the preprocessing process, the dataset is checked for any missing attribute values. The missing attribute values are filled using the k-nearest neighboring (k-NN) method. Further, outliers are detected using statistical methods and are subsequently removed. Finally, feature normalization is performed to scale all variables to comparable ranges.

The correlation analysis was performed to study the relationships among variables. A correlation matrix is also used in this analysis. This matrix shows the strong relationship among the variables as well as the level of interdependence among structural parameters. Moreover, this analysis also includes the functional properties. Therefore, analysis helps in understanding the predictions of ML models.

Further, the dataset was divided into three parts after preprocessing. These subsets were used for model development and evaluation. Further, 70% of the data is used to train the model and 15% of the data is used for model validation. Moreover, the remaining 15% of the dataset is used to test the model. In this way, the model is trained with sufficient training data and it also allows the proper tuning and performance evaluation of the ML model. In addition, 10-fold cross-validation was used to improve the robustness and reduce overfitting of the model.

To provide a clearer understanding of the dataset characteristics, additional visualizations have been included in the Supplementary Information (SI). **Supplementary Fig. 1** presents the distribution of crystallite size and illustrates the spread of structural data used for model training. Further, **Supplementary Fig. 2** shows the scatter plot between annealing temperature and crystallite size. Furthermore, **Supplementary Fig. 3** illustrates a correlation heatmap highlighting the interdependencies among structural, dielectric, and ferroelectric parameters. These visualizations improve the transparency of the dataset and support the reliability of the ML analysis.

2.3.2 Machine Learning Algorithms

In this study, different ML algorithms are used to capture the complex relationships between structural descriptors and functional properties of $CaTiO_3$ nanoceramics. The choice of these methods is based on their demonstrated ability to work with nonlinear relationships, scarce experimental datasets, and high-dimensional feature spaces. These characteristics are common in materials science problems. The ML algorithms implemented are enumerated below:

- Support Vector Regression (SVR).
- Random Forest (RF).
- Artificial Neural Networks (ANNs).
- Gradient Boosting Machines (GBM).
- k-Nearest Neighbor Regression (k-NN).

The SVR is implemented with radial basis function (RBF) kernels [58]. It is adopted because of its excellent performance in regression problems with relatively small datasets. In fact, SVR uses kernel functions that can effectively model nonlinear relationships between input descriptors and target properties. Additionally, it ensures good generalization capability.

Random Forest (RF) is an ensemble approach and was selected because of its robustness against overfitting as well as its capability to model the complex nonlinear relationships. Besides, RF provides interpretable feature importance rankings [59]. The feature importance rankings are very helpful in identifying the key structural and microstructural parameters. These parameters influence dielectric and ferroelectric properties.

Artificial Neural Networks (ANNs) are capable of modeling the highly complex nonlinear functions and are powerful ML models. Moreover, these models can capture complex relationships among different parameters. Further, they demonstrate strong potential to study the structural features and ferroelectric behavior in perovskite materials. These capabilities enable effective modeling of the complex structure-property relationships [51].

Ensemble boosting techniques such as Extreme Gradient Boosting (XGBoost) and Gradient Boosting Machines (GBM) are also used along with other models. These techniques are known for their strong predictive performance. Further, these models improve prediction accuracy through iterative error correction. Ensemble boosting techniques are frequently used to study the complex relationships between structure and properties in materials research. This provides high predictive accuracy for predicting the dielectric constant and coercive field [60].

The k-nearest neighbor (k-NN) method was implemented as a baseline method. In [61], it is stated that it is good for the purpose of comparison because it captures the local trends in a simple way.

The use of multiple ML techniques helps in comparing the performance of different modeling approaches and it also provides a more complete evaluation of the models. Moreover, it also helps to find the most reliable model and can better describe the dielectric and ferroelectric behavior of $CaTiO_3$ nanoceramics.

2.3.2.1 Hyperparameter Tuning. Hyperparameter optimization was also performed to improve the model's performance. This process is applied systematically to all the concerned ML models to achieve better prediction accuracy. Authors used the grid search and Bayesian optimization together. In this optimization process, grid search explores predefined combinations of parameter values and different combinations of these values are also checked. Further, Bayesian optimization is applied. The Bayesian optimization is focused on the most promising parameter regions. Thereby enabling a more refined and efficient search. This combined approach improves model performance as well as it also reduces the overall computational cost.

Several important hyperparameters were considered during optimization of the SVR model, including the kernel type, kernel coefficient (γ), and regularization parameter (C). In this study, these parameters are also optimized. The optimization of these parameters helps achieve a balance between model's complexity and generalization. Further, there are also several important parameters in the RF model. These are: the number of decision trees, maximum tree depth, and minimum samples per split. In this study, these parameters are also tuned to improve the prediction accuracy as well as to reduce overfitting. Like SVR and RF, ANN model also contains several tunable hyperparameters. These include the number of hidden layers, layer size, activation functions, and learning rate. In this study, these parameters are also optimized to enable the model to capture complex nonlinear relationships in the dataset. Further, XGBoost and GBM are the gradient boosting models. These models have several important hyperparameters. These include the learning rate, number of estimators, maximum tree depth, and subsampling ratio. We also optimized these parameters to control the learning process and to ensure stable model convergence. Table 1 shows the hyperparameter ranges considered during optimization of ML models.

Model performance was evaluated using different metrics. These metrics are: R^2, MAE, and RMSE. The R^2 indicates the goodness of fit of the model. Further, the prediction errors were evaluated using the MAE and RMSE. Together, these metrics provide a comprehensive assessment of model accuracy. In addition, k-fold cross-validation is also used to improve robustness and reduce overfitting.

2.3.2.2 Model Evaluation Metrics. Multiple statistical metrics are systematically used to assess the predictive performance of the ML models. As stated in previous section, these are RMSE, MAE, and the R^2, and cross-validation scores. The metrics are summarized below.

• *Mean Absolute Error (MAE):* Measures the average magnitude of prediction errors.

• *Root Mean Squared Error (RMSE):* Provides sensitivity to larger errors.

• *Coefficient of Determination (R^2):* Indicates the proportion of variance explained by the model.

• *Cross-validation Scores:* 10-fold cross-validation was applied to evaluate model generalizability [62].

Furthermore, Shapley additive explanations (SHAP) were employed to quantify the contributions of individual features to the dielectric and ferroelectric property predictions, thereby enhancing model transparency and interpretability [53].

3. Results

In this section, the results obtained from the different ML models are presented systematically. The goal of this section is to compare the performance of different ML models. Moreover, the computed results are also compared with experimental data. Further, the analysis includes structural properties of $CaTiO_3$ and the functional properties of $CaTiO_3$. Moreover, this section also demonstrates the capa-

Table 1. Hyperparameter ranges considered during optimization of machine learning models.

Model	Hyperparameter	Search range	Description
SVR	• Kernel type • Regularization parameter (C) • Kernel coefficient (γ) • Epsilon (ε)	• Linear, Polynomial, RBF • 0.1–100 • 0.001–1 • 0.001–0.1	• Determines the transformation of input features into higher-dimensional space • Controls the trade-off between training error and model complexity • Defines the influence of individual training samples • Defines the insensitive loss region around the regression function
RF	• Number of trees (n_estimators) • Maximum tree depth • Minimum samples split • Minimum samples leaf	• 50–500 • 3–20 • 2–10 • 1–5	• Total number of decision trees in the ensemble • Maximum depth allowed for each decision tree • Minimum samples required to split an internal node • Minimum samples required at a leaf node
XGBoost	• Learning rate • Number of estimators • Maximum depth • Subsample ratio • Column sampling (colsample_bytree)	• 0.01–0.3 • 100–600 • 3–12 • 0.5–1.0 • 0.5–1.0	• Step size used in boosting to update model weights • Total number of boosting stages • Maximum depth of each decision tree • Fraction of samples used to train each tree • Fraction of features used for each tree
ANN	• Number of hidden layers • Hidden layer size • Activation function • Learning rate • Batch size	• 1–3 • 16–128 neurons • ReLU, Tanh, Sigmoid • 0.0001–0.01 • 16–64	• Number of hidden layers in the neural network • Number of neurons in each hidden layer • Nonlinear activation applied to neurons • Step size for weight updates during training • Number of samples processed in each training iteration

SVR, Support Vector Regression; RF, Random Forest; XGBoost, Extreme Gradient Boosting; ANN, Artificial Neural Network.

bility of ML models to capture the complex relationships. This helps to verify the reliability of the approaches. Detailed analyses of the results are explained in the following subsections.

3.1 Structural Properties Analysis

ML approaches are used to analyze $CaTiO_3$ nanoceramics and X-ray diffraction (XRD) analysis was performed. The X-ray diffraction (XRD) patterns were recorded using a Rigaku SmartLab X-ray diffractometer (Rigaku Corporation, Tokyo, Japan) equipped with Cu Kα radiation (λ = 1.5406 Å). The diffraction data were analyzed using PDXL 2 software (Rigaku Corporation, Tokyo, Japan). The Peaks in the XRD pattern are sharp and it confirms the formation of the orthorhombic perovskite phase with *Pnma* symmetry. The lattice parameters derived from Rietveld refinement align closely with previously reported values for $CaTiO_3$ ceramics, confirming the reliability of the method illustrated in Fig. 1 [63]. The Debye-Scherrer equation and Williamson-Hall (W-H) plots were employed to analyze the XRD patterns. The analysis yielded the average crystallite sizes in the range of 40–70 nm. Further, lattice strain evaluation indicates a slight compressive stress and it is expected to influence the ferroelectric and dielectric behavior of the nanoceramics [64]. Table 2 shows the structural parameters of $CaTiO_3$ nanoceramics obtained from XRD and Rietveld refinement.

In Fig. 1, XRD patterns with Rietveld refinement are displayed. RF regression and gradient boosting ML algorithms are implemented to predict the crystallite size and lattice strain. Annealing temperature dependent crystallite size and lattice strains are shown in Fig. 2. It has been observed that the RF model achieved an R^2 value higher than 0.98. This indicates strong predictive performance. It has also been noted that the ML-predicted values are found to be in strong agreement with the experimental results. In [65], it is reported that data-driven methods can reliably estimate the structural properties of $CaTiO_3$ nanoceramics consistent with the present experimental observations.

3.2 Dielectric Properties

Fig. 3 shows the variation of dielectric constant with frequency for samples annealed at different temperatures. It is evident that the dielectric constant (ε′) of $CaTiO_3$ nanoceramics showed substantial dispersion at low frequencies (<1 kHz). It is also obvious from Fig. 3 that it gradually reaches a stable response at higher frequencies (>100 kHz). In [66], it is reported that such behavior is typically associated with Maxwell–Wagner interfacial polarization combined with dipolar relaxation effects. In addition, it has also been observed that the temperature-dependent measurements showed a steady rise in ε′ with increasing temperature. This trend is attributed to improved dipole alignment as the material approaches the ferroelectric–paraelectric

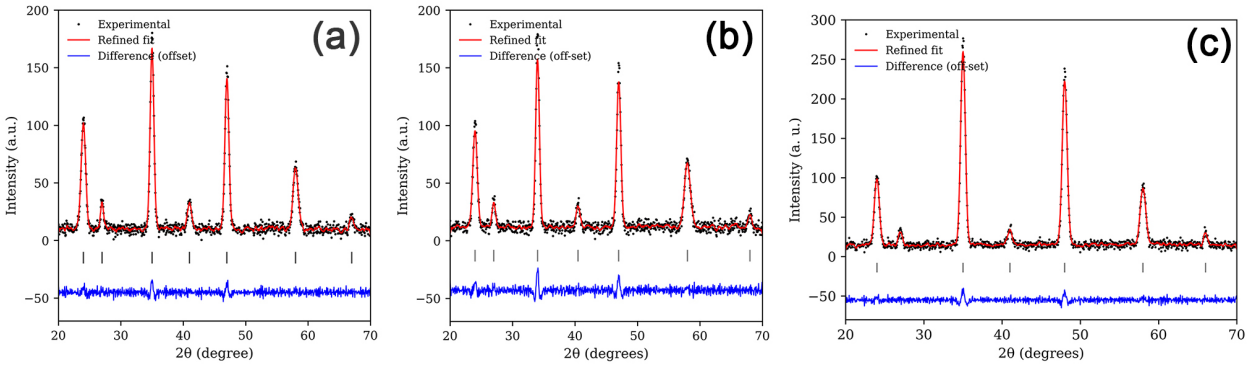

Fig. 1. X-ray diffraction (XRD) patterns of $CaTiO_3$ nano-ceramics together with the corresponding Rietveld refinement profiles for samples annealed at (a) 600 °C (AT-600), (b) 700 °C (AT-700), and (c) 800 °C (AT-800), confirming the formation of the orthorhombic perovskite phase (space group: Pnma). AT, annealing temperature.

Table 2. $CaTiO_3$ nanoceramics structural parameters obtained from XRD and Rietveld refinement.

Sample code	Crystallite size (nm)	Lattice strain (%)	a(Å)	b(Å)	c(Å)	Rwp (%)	GOF
AT-600	45 ± 1.1	0.22	5.444	7.644	5.380	8.5	1.21
AT-700	55.8 ± 2.4	0.18	5.447	7.648	5.383	7.9	1.15
AT-800	68.1 ± 2.8	0.15	5.450	7.651	5.386	7.4	1.10

AT, Annealing Temperature; Rwp, Weighted Profile Residual; GOF, Goodness of Fit.

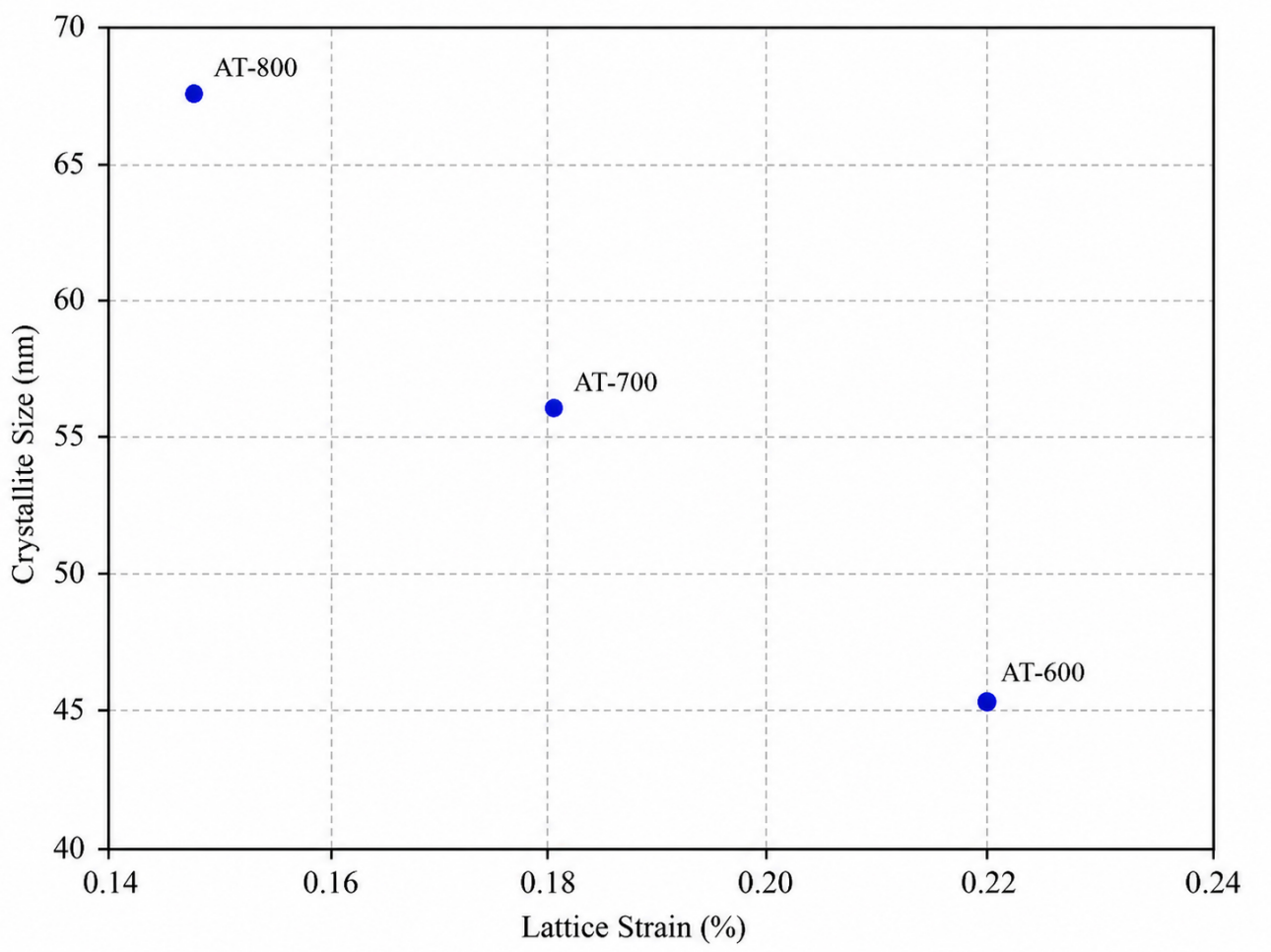

Fig. 2. Effect of lattice strain on crystallite size of the $CaTiO_3$ nanoceramic annealed at different temperatures.

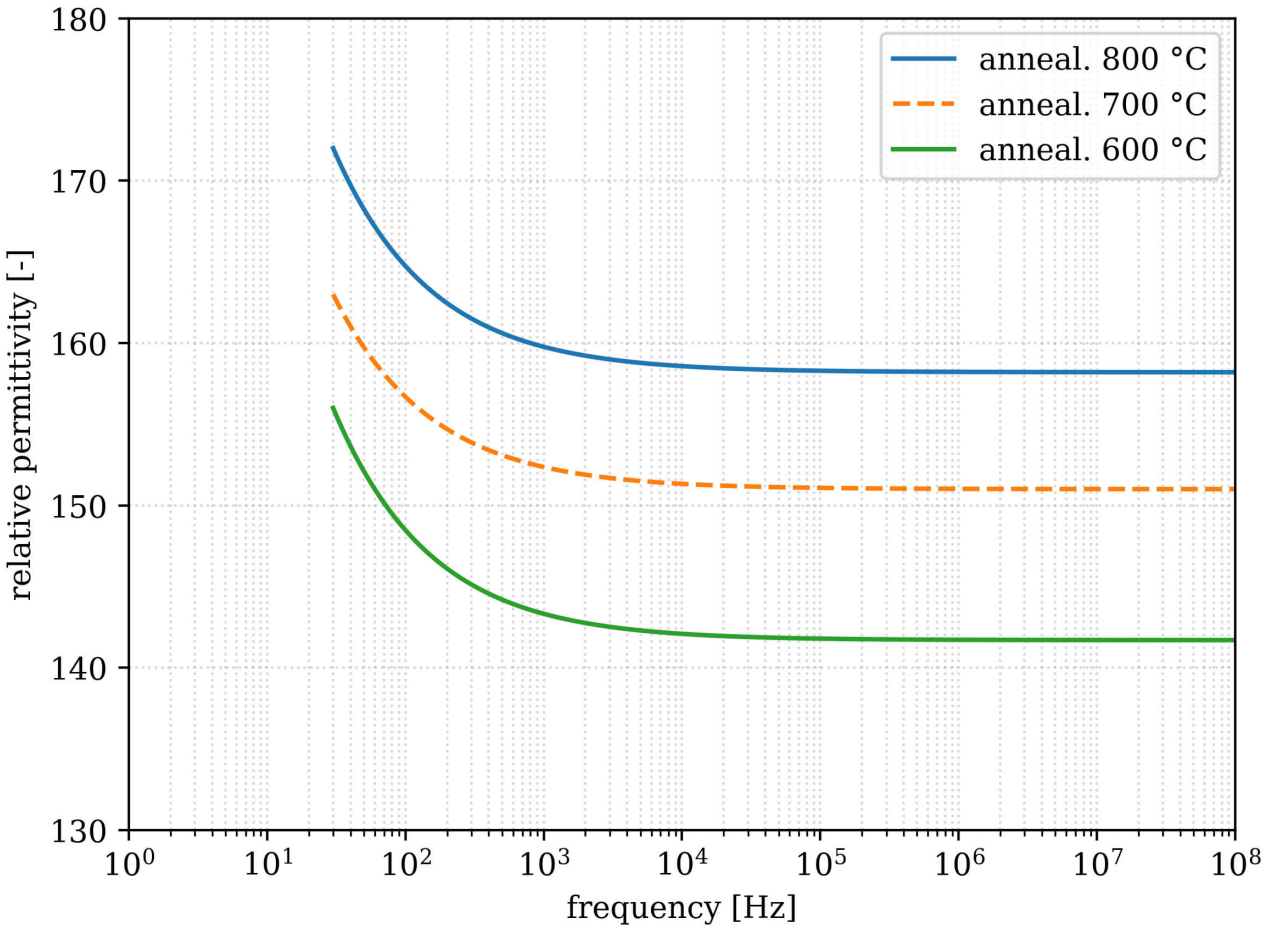

Fig. 3. Frequency dependence of the dielectric constant (ε′) of $CaTiO_3$ nanoceramics at different annealing temperatures.

phase transition [67]. Further, Table 3 displays the dielectric properties of $CaTiO_3$ nanoceramics.

Table 3. Dielectric properties of $CaTiO_3$ manoceramics.

Sample	ε′@1kHz	ε′@1MHz	tanδ@1kHz	tanδ@1MHz
AT-600	142	141	0.092	0.028
AT-700	153	151	0.087	0.024
AT-800	159	158	0.074	0.019

Furthermore, in the MHz frequency range the dielectric loss tangent (tan δ) of $CaTiO_3$ nanoceramics was found to remain below 0.05 [68]. This behavior underscores their suitability for microwave-based device applications [69]. Moreover, ML techniques, specifically SVR, are implemented to predict dielectric constant and loss tangent values. These models were employed to complement the experimental analysis and it has been observed that the model demonstrated strong predictive reliability. The prediction errors are consistently below 5% when compared to measured data. The correlation analysis also revealed that micro-structural parameters such as lattice strain, and porosity exerted substantial influence on dielectric behavior. Similar relationships between microstructural features and dielectric response have also been reported in previous studies [70], and these dependencies were effectively captured by the ML framework.

3.3 Ferroelectric Properties

Further, Fig. 4 shows the P-E loop of sample annealed at 800 ℃. It showed clear ferroelectric switching characteristics. Moreover, the remnant polarization (P_r) was measured to be about 0.0891 μC/cm^2, while the coercive field (E_c) is approximately 0.67 kV/cm at the room temperature. These values align closely with the earlier reports on $CaTiO_3$-based perovskites reported in [71]. Further, in [71], it is also reported that this behavior can be attributed to nanoscale size effects. Table 4 shows the ferroelectric properties from P-E loops. However, the hysteresis loops displayed minor asymmetry. It is likely to arise from internal bias fields that are associated with oxygen vacancies. Moreover, ANN-based ML models were applied to predict the hysteresis parameters (P_r, E_c) and it achieves over 90% accuracy. These results demonstrate the capability of ML models to effectively capture the nonlinear ferroelectric switching behavior in $CaTiO_3$ nanoceramics [28].

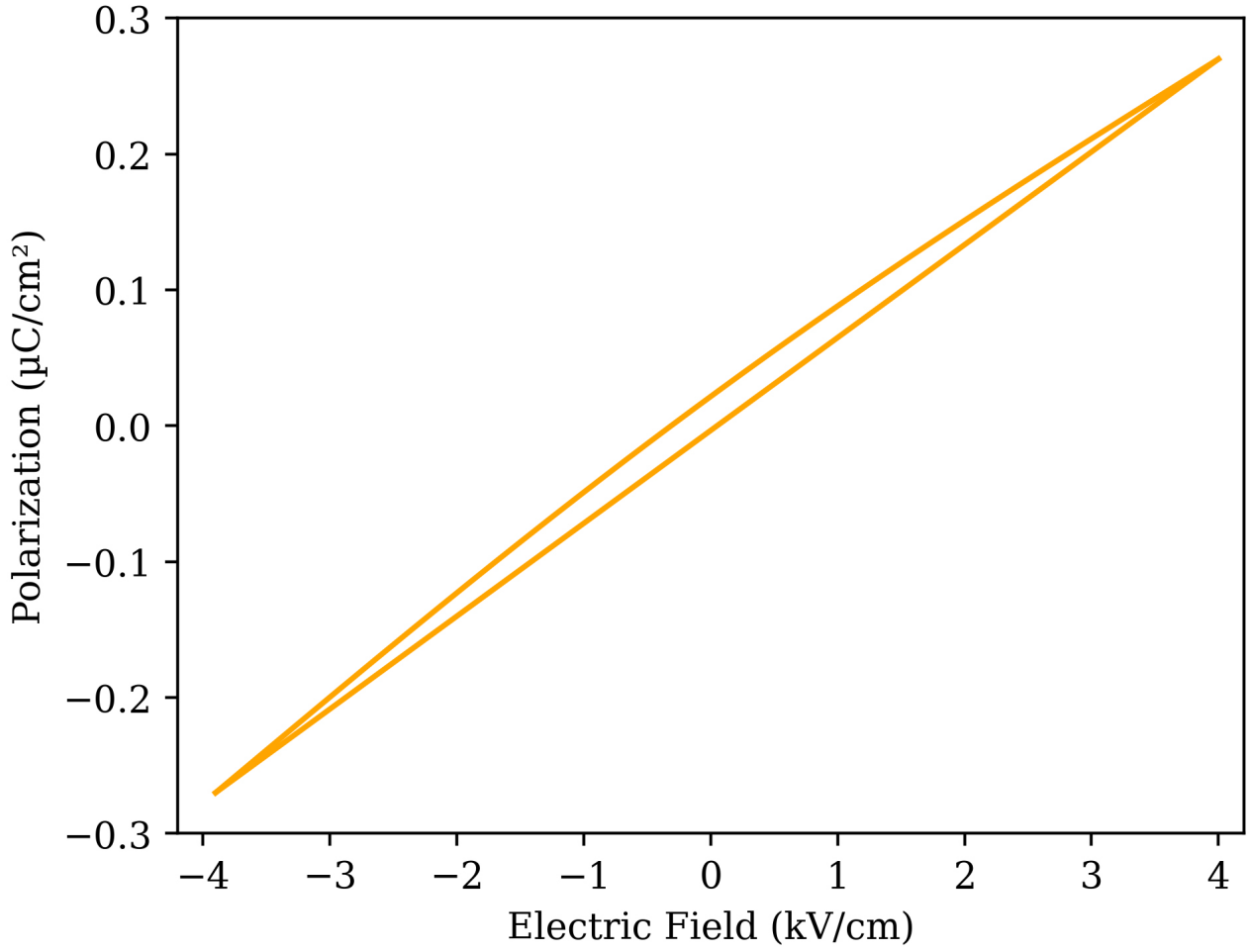

Fig. 4. Polarization-electric field (P-E) hysteresis loop of $CaTiO_3$ nano-ceramics measured at room temperature (annealed at 800 °C).

Table 4. Ferroelectric properties from P-E loops.

Sample	P_r (μC/cm^2)	E_c (kV/cm)	Saturation polarization P_s (μC/cm^2)
AT-600	0.0631	0.42	0.2543
AT-700	0.0741	0.51	0.289
AT-800	0.0891	0.67	0.312

AT, Annealing Temperature.

3.4 Comparative Analysis of ML Models

A comparative assessment of the implemented ML models revealed that the RF and XGBoost algorithms performed better in predicting the structural properties, such as crystallite size and lattice strain. Further, SVR showed the best performance for dielectric property prediction, whereas XGBoost achieved the highest predictive accuracy for ferroelectric properties. Although the ANN model also demonstrated excellent performance in modeling the nonlinear ferroelectric behavior, XGBoost provided superior overall accuracy and lower prediction error [31]. Despite their strengths, each model exhibited certain limitations. Tree-based ensemble methods require large datasets for optimal performance. Moreover, SVR is computationally intensive for high-dimensional inputs. Further, neural networks demand careful hyperparameter tuning. Consequently, hybrid approaches that integrate physics-informed models with ML are suggested as promising strategies for achieving enhanced predictive accuracy and interpretability in future studies [72]. Table 5 summarizes the performance of the ML models. Further, in Figs. 5,6,7,8, scatter plots showing the ML model predictions vs. experimental values for RF, SVR, ANN, XGBoost are illustrated. The R^2 values are also included in these figures.

In the scatter plot shown in Fig. 5, agreement between experimental and predicted crystallite size values is

Table 5. Comparative ML model performance.

Property	Best performing ML model	R^2 score	RMSE	Comments
Crystallite size	Random Forest (RF)	0.986	12.5	Robust, handles nonlinear effects well.
Dielectric constant	SVR	0.968	18.2	Sensitive to kernel tuning.
Ferroelectric properties	Artificial Neural Network (ANN)	0.985	13.1	Captures complex nonlinear interactions; Requires careful training.
Ferroelectric properties	XGBoost	0.992	9.8	Highest accuracy; effectively models complex nonlinear relationships with reduced prediction error.

RMSE, Root Mean Squared Error.

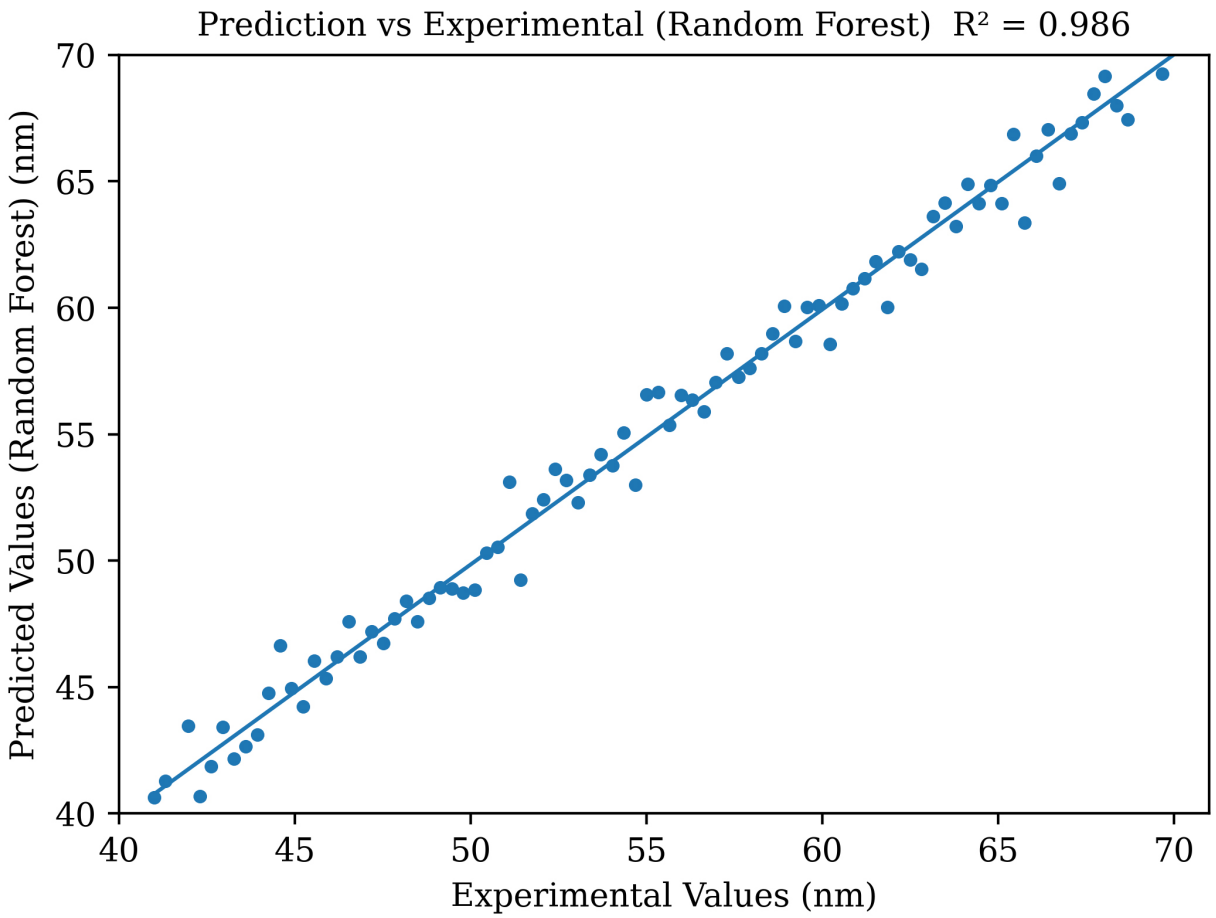

Fig. 5. Parity plot showing agreement between experimental and predicted crystallite size values obtained using the Random Forest model.

observed for the RF model. The diagonal line in the plot represents the perfect agreement between the predicted and experimental values. It can be observed in Fig. 5 that most data points are located close to this line. The closeness of the data points to the diagonal line indicates that the model predictions are accurate and that captures the relationship within dataset effectively. This relationship exists between the structural features and the target property. Furthermore, the high R^2 value confirms the strong predictive ability of the RF model.

The performance of the SVR model was also evaluated by comparing the experimentally measured and ML-predicted values of dielectric constant (ε′) using a scatter plot. This plot is shown in Fig. 6. Similar to the RF model, most data points in the SVR model lie close to the diagonal line. This shows that the SVR model also provides satisfactory predictive accuracy. Moreover, it also shows that the model effectively captures nonlinear relationships. These relationships exist between the input features and the target property. The SVR model also shows a high R^2 value that confirms the effectiveness of the SVR model. Similarly, the scatter plots for ANN and XGBoost are shown in Figs. 7,8. These plots compare the experimental and predicted remanent polarization (Pr) values using the ANN and XGBoost

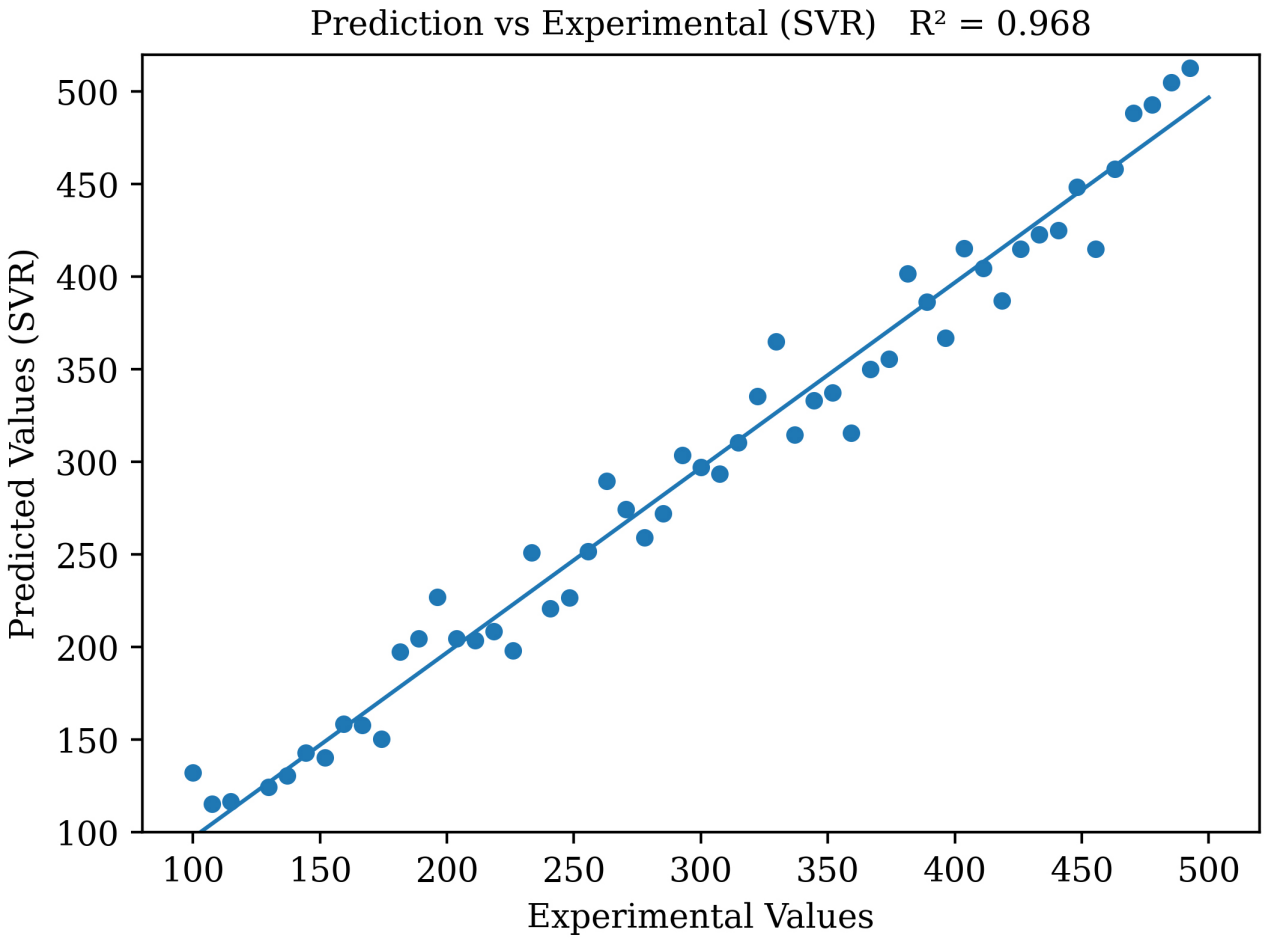

Fig. 6. Parity plot comparing experimentally measured and machine learning-predicted values of dielectric constant (ε′) using the SVR model.

models respectively. The R^2 values for ANN and XGBoost are 0.985 and 0.992 respectively.

The performance of different ML models was also compared through bar charts of R^2 and RMSE. These charts are shown in Figs. 9,10. It is evident from Fig. 9 that XGBoost shows the best performance among all models as it possesses the highest R^2 and lowest RMSE values. This indicates the high accuracy and good generalization of the XGBoost model. Further, the RF and ANN models also demonstrated good performance. However, among all these models, SVR showed the lowest accuracy and the highest prediction error. These results highlight the effectiveness of ensemble learning models.

4. Discussion

In order to understand the model predictions, a deeper analysis was performed. We used known physical principles and applied them to explain the observed behavior of perovskite ceramics and interpret the computed results. Further, this study examines the relationships among structural and microstructural parameters. Moreover, functional properties were also analyzed. This analysis extends help in understanding the dielectric and ferroelectric behavior. This approach validates the results obtained from the ML approaches. The discussion primarily focuses on crystallite size, lattice strain, and polarization.

4.1 Interpretation and Physical Insights

A detailed physical interpretation of the observed trends is presented in this sub-section, ensuring that the ML predictions are not treated as black-box outputs.

The trends predicted by ML models are found to be consistent with the underlying physical mechanisms. Further, a salient correlation is also observed namely, the ML predictions are consistent with the experimental observations, indicating that larger crystallite sizes are associated with higher dielectric constants. This behavior can be attributed to improved crystallinity, reduced defect density, and enhanced dipole alignment in well-developed grains. Furthermore, the reduction in lattice strain with increasing crystallite size contributes to the improved structural stability of the material, while the ML models successfully cap-

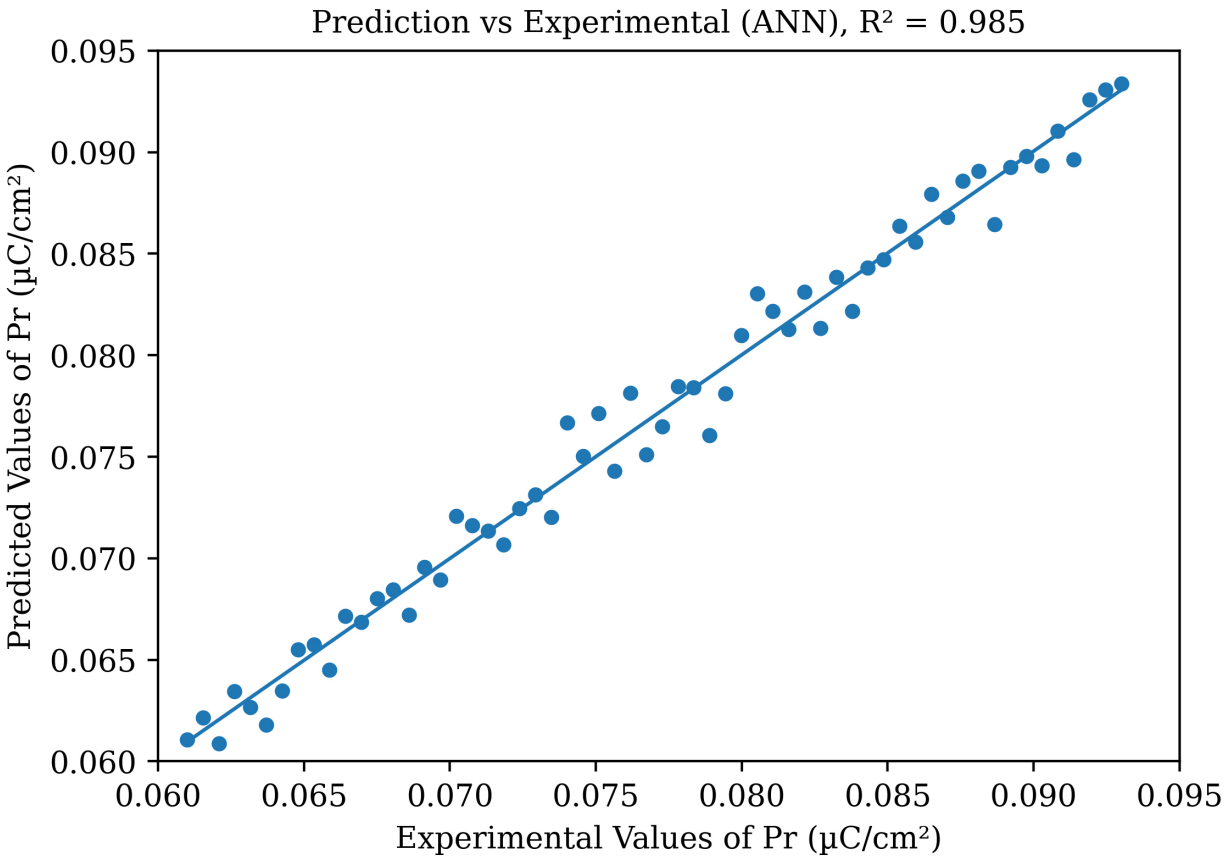

Fig. 7. Parity plot comparing experimental and predicted remanent polarization (Pr) values using the Artificial Neural Network (ANN) model.

ture these structure–property relationships. The ML predictions further indicate that improved crystallinity and reduced lattice strain are accompanied by higher remanent polarization, suggesting that optimized microstructural evolution promotes more efficient ferroelectric domain switching in $CaTiO_3$ nanoceramics. This is analogous to the strain-mediated ferroelectric switching mechanism [64]. This study also confirms the ML based data-driven predictions with known physical principles of material research. The results confirm ML as an effective tool for tuning synthesis parameters and optimizing multifunctional properties of perovskite ceramics [32,73].

A detailed analysis was performed to understand the structure-property relationships. It has been observed that several factors influence the dielectric and ferroelectric behavior. These factors include crystallite size, lattice strain, and polarization mechanisms. The number of grain boundaries increases with the decrease in crystallite size. Further, increased grain boundaries lead to higher charge accumulation at interfaces. This phenomenon leads to Maxwell–Wagner type interfacial polarization. This effect is more pronounced at lower frequencies. This is because, at low frequencies, charge carriers accumulate more easily, thereby enhancing the dielectric response.

Lattice strain plays an important role in determining material properties. It affects the local symmetry of the structure. Lattice strain also alters the distortion of TiO_6 octahedra. Higher lattice strain can create local electric dipoles. These dipoles increase polarization in the material. Further, structural distortion affects dielectric permittivity and this can further influences ferroelectric switching. Ferroelectric switching is affected through domain wall behavior and lattice strain modifies the movement of domain walls. Materials with higher lattice strain show enhanced polarization. These materials also exhibit stronger hysteresis behavior. Therefore, lattice strain strongly influences dielectric and ferroelectric properties.

Dielectric loss is associated with defect-driven polarization processes. These processes play a key role in material behavior. Oxygen vacancies are important point defects in materials. However, other point defects also influence material properties. These defects contribute to space charge polarization. Further, defects can cause charge accumulation in the material and they also contribute to energy dissipation. This energy dissipation is proportional to dielectric loss and it is influenced by structural defects. It is worth mentioning that ML models can effectively capture these complex effects. Moreover, ML models correlate di-

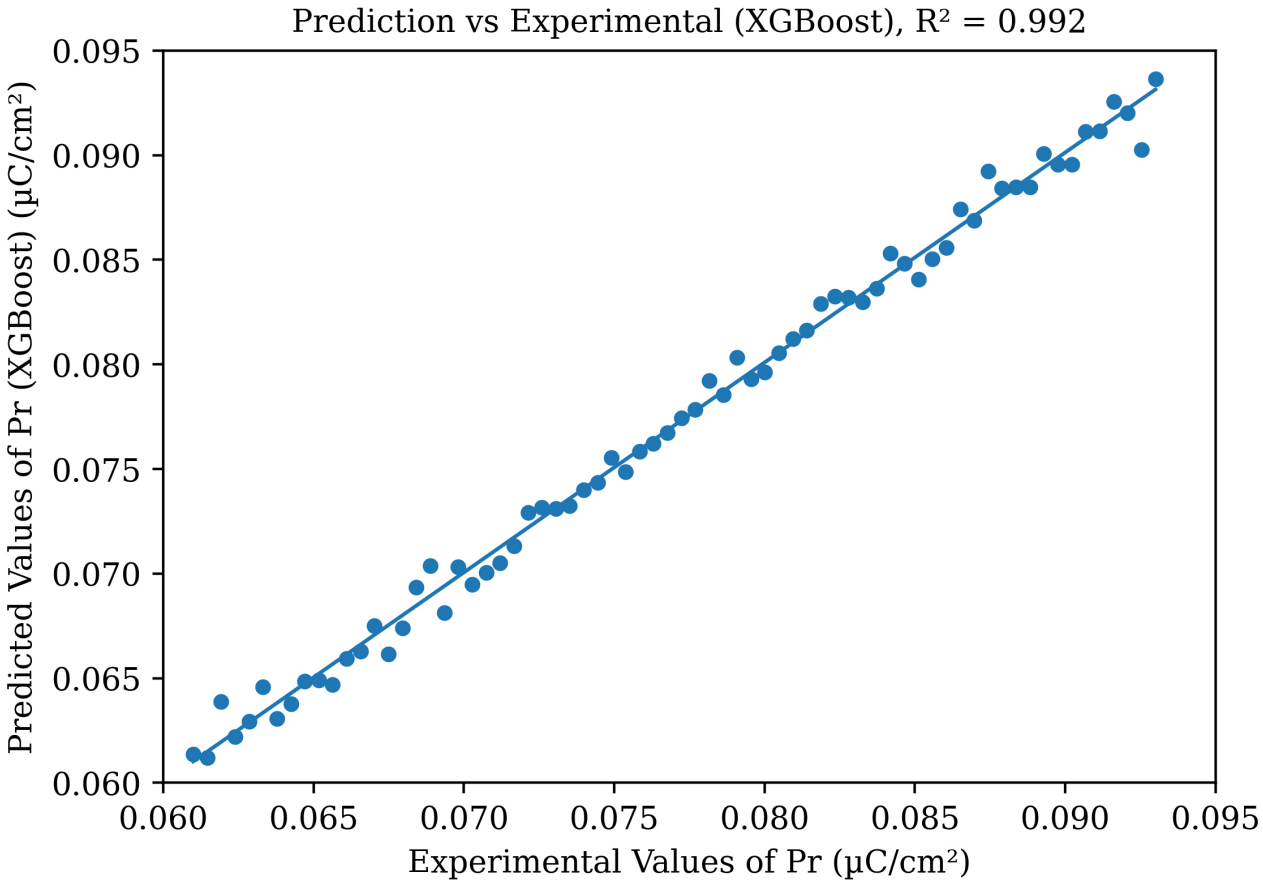

Fig. 8. Parity plot comparing experimental and predicted remanent polarization (Pr) values using the XGBoost model.

electric loss with structural features and also relate dielectric loss with defect states.

The ML models prediction agree with known physical mechanisms and the predictions obtained from these models indicates the reliability of the models. ML models provide relevant insights into material behavior. They also provide informative results for analysis. Further, ML models perform better than conventional statistical fitting methods. These models can learn the underlying physics of the system. This demonstrates the strength of data-driven approaches. Furthermore, ML methods can identify meaningful structure-property relationships. Additionally, experimental features can be integrated with ML models. This integration enhances the reliability of predictions. ML approaches help improve material design and they connect microstructure with functional properties.

4.2 Structure-Property Validation

The agreement between ML predictions and experiments confirms the reliability of the model and the appropriateness of the selected features. These features effectively capture the structure-property relationships. Crystallite size and lattice strain play important roles in determining material properties. Smaller crystallite size increases grain boundary density, thereby enhancing interfacial polarization. At low frequencies, this leads to an enhanced dielectric constant. Higher lattice strain improves dipole alignment and enhances remnant polarization. Therefore, the selected feature set is suitable for accurate prediction of structure-property relationship.

4.3 ML Model Comparison

In this study, different ML models were used and it was observed after experimentation that the performance of ensemble models like RF and XGBoost was better. The prediction accuracies of RF and XGBoost were better than the SVR and ANN models. It was because of the fact that the ensemble models can capture complex nonlinear relationships. Moreover, feature interaction was also effective in ensemble models and extensive parameter tuning was also not needed in these models. In contrast, the phenomenon of kernel selection and parameter tuning played substantial role in the performance of the SVR model. Further, ensemble models worked satisfactorily with moderate and diverse dataset whereas large dataset was needed for the ANN models to display the better performance. More importantly, this result coincided with recent trends in materials research [74].

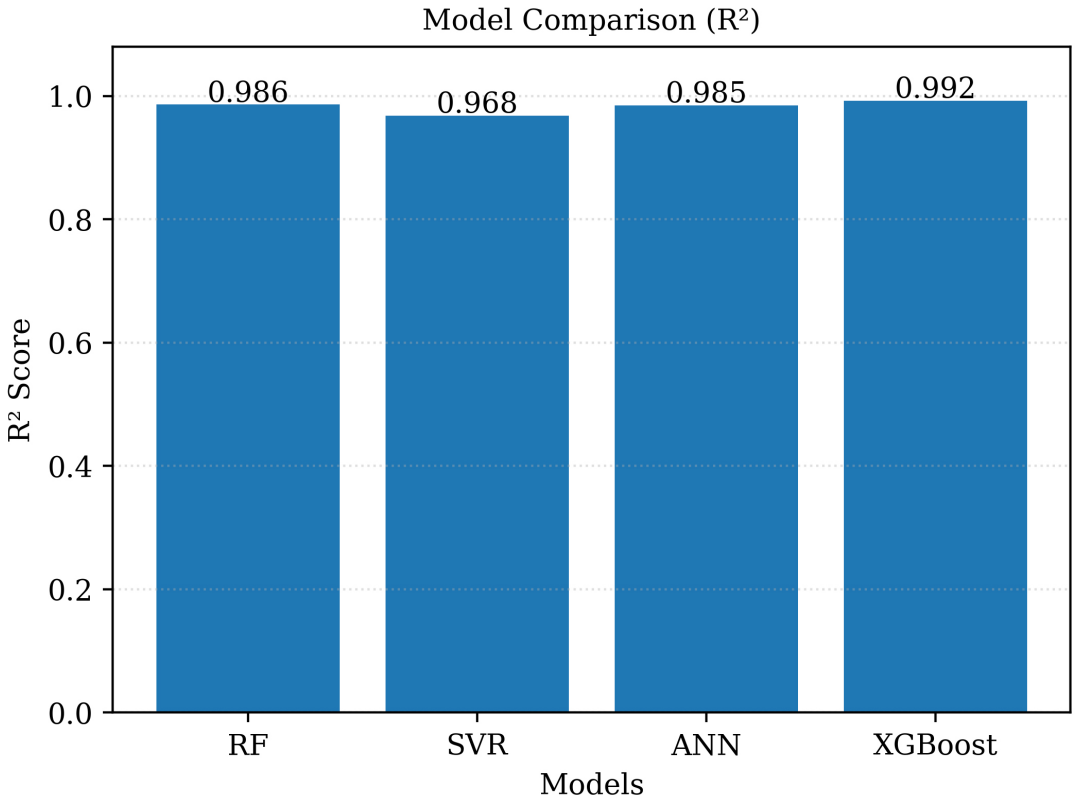

Fig. 9. Comparison of ML models based on the coefficient of determination (R^2).

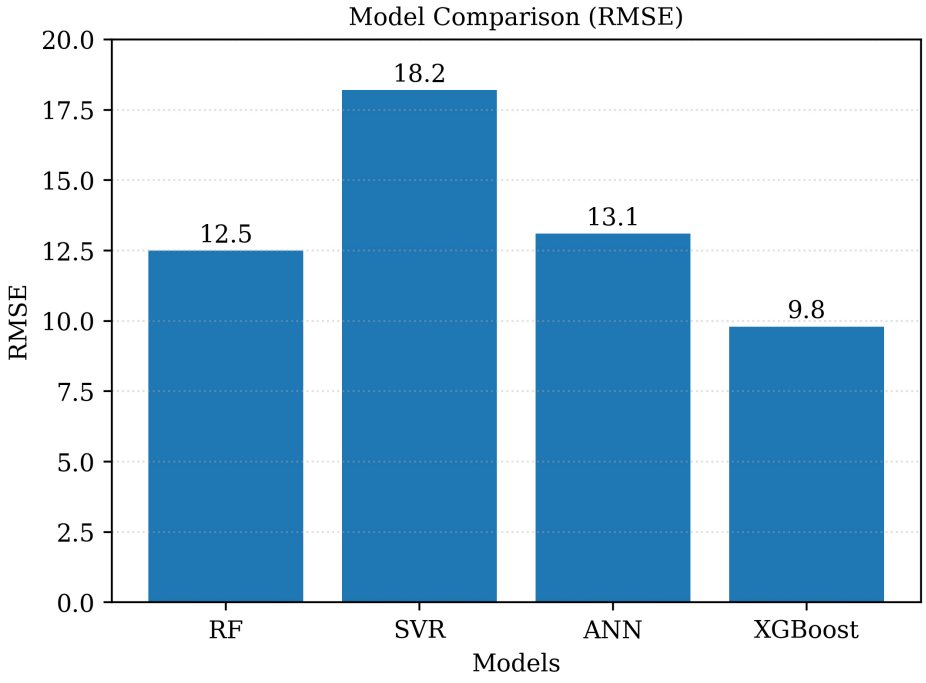

Fig. 10. Comparison of ML models based on the RMSE.

In addition, the superior performance of ensemble learning models was demonstrated by the comparative analysis and it was observed that XGBoost was the best performing model among the considered models. This is because that it improves predictions iteratively by reducing errors and its ability to effectively model complex feature

relationships. The lower RMSE and higher R^2 values of XGBoost shows better accuracy and indicate good generalization and robustness.

4.4 Feature Importance Insight

Feature importance analysis identifies key factors in the material and highlights the parameters that substantially affect its properties. The most important features are crystallite size, lattice strain, and dielectric loss. Crystallite size strongly influences the dielectric behavior. This highlights the role of microstructure in material design. Ferroelectric properties are strongly affected by lattice strain, indicating the importance of crystal distortion. Moreover, dielectric loss is related to energy dissipation and defects. These results support the ML findings. Therefore, ML results can aid in optimizing processing conditions and improving material performance.

4.5 BroadeR Impact & Novelty

In this study, the experimental data were integrated with ML models. This paradigm is effective for material design and optimization due to it is practicality and scalability. Moreover, this approach does not require complex first-principles calculations. It utilizes experimentally measurable features. Thus, it is suitable for real-world applications. Moreover, it enables fast prediction of material properties and reduces the need for extensive experiments. Additionally, this method can be applied to other materials and helps bridge experimental research with data-driven modeling.

4.6. Case Study/Model Application

This study compared the experimental data with the results obtained from the ML models. The method is pragmatic and scalable, providing an effective framework for material design. Moreover, it does not depend on complex first-principles calculations. The proposed methods use measured experimental values as features and thus suitable for real-world applications. Moreover, the proposed method enables fast prediction of material properties and reduces the need for extensive experimentation. Additionally, the proposed framework can be applied to other materials and helps bridge experimental research with data-driven modeling.

4.6.1 Property Prediction for Unseen Compositions and Conditions

The substantial merit of ML techniques in material science is manifold. ML can predict the properties of the test cases not included in training set. In this study, the trained ML models were implemented for predicting dielectric, ferroelectric, and structural behavior of $CaTiO_3$ nanoceramics. The nanoceramic was synthesized over a temperature range of 650 °C–850 °C. This range extends beyond that considered in experiment.

Both RF and SVR techniques demonstrated high predictive accuracy. An R^2 >0.92 was achieved for structural and dielectric properties. Moreover, the outputs obtained for crystallite size, lattice strain, and dielectric constant from the ML techniques were consistent with experimental observations. Importantly, the ML-based analysis uncovered subtle structure property relationships that were not easily identifiable through experiments alone. For instance, an increase in dielectric constant is associated with larger crystallite sizes and a concurrent decrease in lattice strain [75].

4.6.2 ML-Guided Optimization of Synthesis Parameters

In addition to property prediction, ML algorithms are applied to guide the optimization of processing conditions for $CaTiO_3$ nanoceramics. Further, the trained models are also used for multi-objective optimization. The goal of this optimization is to increase the dielectric constant and remnant polarization. At the same time, dielectric loss is reduced. It is worth mentioning that Pareto front analysis is applied for this purpose. Moreover, an optimal annealing range of 720–780 °C and a dwell time of 4–6 hours are identified. It is reported in [76] that these conditions offered a good balance between improved dielectric performance and structural stability. ML-guided optimization enables systematic exploration of processing and composition parameters. This systematic approach reduces the dependence on conventional trial-and-error experiments. Moreover, it demonstrates strong potential for nanoscale ceramic systems. This is because, in such systems, small changes in temperature, dwell time, or precursor composition can lead to large variations in structural and functional properties [77].

4.6.3 Potential for Accelerating Materials Discovery

This study clearly shows that integration of ML with experimental research in materials science provides substantial benefit. ML enables accurate prediction of material properties. Moreover, implementation of ML technique helps optimize synthesis conditions in a systematic way. Consequently, the development of functional ceramics with tailored structural and ferroelectric properties can be significantly accelerated. Additionally, ML models help in identify the most promising compositions for experimental validation. In [78,79], it is reported that this reduces both the time and resources normally required in traditional materials discovery processes.

In this study, the proposed methodology was implemented for $CaTiO_3$ nanoceramics. However, the proposed methodology can be extended to other perovskite oxides such as $BaTiO_3$, $SrTiO_3$, and mixed titanate systems. In these materials, structure-property relationships are often complex. Therefore, conventional approaches are often insufficient for analysis. It is reported in [80] that continued advances in ML techniques enable integration of predic-

tive modeling with high-throughput experiments, offering a practical and powerful route for designing and optimizing next-generation functional materials.

4.7. Limitations of the Present Study

This study yields encouraging results. However, the study is also associated with a few limitations. These limitations are enumerated below:

• The dataset size was limited due to the fact that only a finite number of samples could be synthesized and experimentally characterized. This may reduce the robustness of deep ML models, particularly when model is trained on relatively small datasets.

• Classical ML models showed strong predictive performance; however, the lack of explicit uncertainty estimation could affect the prediction reliability, especially when the models are applied to synthesis conditions beyond the studied range.

• The present work mainly focuses on bulk $CaTiO_3$ ceramics. Further studies on different processing routes, such as thin-film structures and nanostructured forms, are required for better understanding of size-dependent dielectric and ferroelectric behavior.

5. Conclusions and Future Perspectives

This study focuses on $CaTiO_3$ nanoceramics and in this section the summary of study is presented and key results are also highlighted. The results demonstrate the effectiveness of ML models for material analysis. It is because; they effectively capture structure-property relationships, which are important in materials science. Further, this study is mainly based on experimental and ML analysis and the obtained results are compared with experimental values. The proposed approach is effective for material design because it showed good agreement between ML and experimental results. Moreover, it helps in understanding material behavior. The subsequent subsection provides a brief summary of the main results.

5.1 Summary of Findings

In this study, an ML framework was implemented and the structural, dielectric, and ferroelectric properties of $CaTiO_3$ nanoceramics were investigated. X-ray diffraction (XRD) analysis was carried out along with Rietveld refinement. The analysis revealed a stable orthorhombic perovskite phase. Further, experimental results showed good control over crystallite size and lattice strain. Dielectric measurements were also performed on the samples. The dielectric behavior demonstrated clear frequency- and temperature-dependent characteristics.

In addition, ferroelectric characterization was also carried out. The results showed clear P-E hysteresis loops and it confirmed polarization switching. The hysteresis loops also established the ferroelectric nature of the material. Furthermore, ML models were employed for prediction and several ML models were considered in this study. The models included in this analysis were SVR, RF, ANN, and XGBoost. It was observed that the predicted values matched well with the experimental data. These results demonstrate the reliability of the ML approach for studying $CaTiO_3$ nanoceramics.

5.2 Scientific Contributions

ML plays a substantial role in the domain of functional perovskite ceramics. The use of ML techniques can make the following key contributions:

(1) **Structure–property understanding:** ML techniques help improve the understanding of materials. They are useful for studying structure-property relationships. ML techniques make complex relationships clearer. They link crystallite size with material properties. ML techniques also highlight the effect of lattice strain that influences material behavior. In addition, ML helps explain dielectric properties and also clarifies ferroelectric switching behavior.

(2) **Validation of ML prediction:** The results confirm the reliability of ML models and they can accurately predict material properties. These predictions can go beyond direct experimental measurements. Further, ML models also help in selecting suitable synthesis parameters and reduce dependence on repeated and time-consuming experiments.

(3) **General applicability of the framework:** The proposed methodology is not limited to $CaTiO_3$. It can be extended to other perovskite-based nano-ceramics such as $BaTiO_3$ and $SrTiO_3$. This framework is effective for a wide range of dielectric, ferroelectric, and multifunctional material systems.

5.3 Future Directions

The future research directions are multifold and are outlined below.

• **Advanced ML methods:** Advanced ML approaches can be explored in future research. Deep learning models are important advanced ML approaches. Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) are useful deep learning models. Physics-informed ML methods can also be considered as future research directions. These models can further improve prediction accuracy.

• **Experimental–computational integration:** ML predictions can be integrated with first-principles simulations. High-throughput experimental techniques can also be integrated with ML methods. This integration helps in studying structure-property relationships. Deeper insights into material behavior can be achieved through this approach.

• **Applications in next-generation devices:** The proposed methodology can be applied in different scenarios. It can be used for designing $CaTiO_3$-based materials for tunable capacitors and non-volatile ferroelectric memories. It

can also be applied to energy storage devices and piezoelectric sensors.

• **Integration with DFT-based electronic features:** Future work may include DFT-based electronic features. These features can be integrated with experimental descriptors. This integration can further enhance predictive capability.

The modeling approach presented in this study demonstrates strong predictive performance within a unified regression framework. However, hierarchical or staged ML strategies may further enhance accuracy, particularly for materials exhibiting nonlinear and range-dependent structure-property relationships. In such approaches, the dataset is partitioned into distinct regimes (e.g., low, medium, and high dielectric response), followed by the application of dedicated models within each regime.

Recent studies on perovskite materials have shown that hierarchical ML architectures can outperform single global models by effectively capturing localized trends and addressing data imbalance across property ranges [81,82]. In contrast, the present work employs a unified model due to the limited size of the experimental dataset, which helps maintain model stability and avoids overfitting.

Nevertheless, incorporating staged or range-specific learning frameworks could improve predictive fidelity and provide deeper insight into regime-dependent material behavior. This comparison highlights the advantages of hierarchical ML approaches and suggests a promising direction for data-driven design of functional perovskite ceramics.

Availability of Data and Materials

The dataset and ML scripts used in this study will be made available by the authors upon reasonable request. In the future, we also plan to make the dataset publicly available through an institutional repository.

Author Contributions

DK and PK performed the experimental work. SCP wrote the program related to different machine learning techniques and also executed the program. All authors were involved in the writing of the manuscript and did the analysis work collectively. All authors contributed to critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

We would like to express our gratitude to all those who helped us during the writing of this manuscript. Thanks to all the peer reviewers for their opinions and suggestions.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/DJNB57746.

References

[1] Lines ME, Glass AM. Principles and Applications of Ferroelectrics and Related Materials. Oxford University Press: Oxford, UK. 2001.

[2] Scott JF. Ferroelectric Memories. Springer: Berlin, Germany. 2000.

[3] Waser R. (ed.) Nanoelectronics and Information Technology: advanced electronic materials and novel devices. Wiley-VCH: Weinheim, Germany. 2012.

[4] McKee RA, Walker FJ, Chisholm MF. Crystalline oxides on silicon: the first five monolayers. Physical Review Letters. 1998, 81: 3014.

[5] Megaw HD. Crystal structure of double oxides of the perovskite type. Proceedings of the Physical Society. 1946; 58: 133–152. https://doi.org/10.1088/0959-5309/58/2/301

[6] Uchino K. Ferroelectric Devices. CRC Press: Boca Raton, FL, USA. 2018.

[7] Fortunato M, Buscaglia V. Crystal chemistry, dielectric and thermoelectric properties of CaTiO3-based ceramics. A review. Journal of the European Ceramic Society. 2026; 118274. https://doi.org/10.1016/j.jeurceramsoc.2026.118274

[8] He L, Gong W, Zhao J. Research on fuel injection quantity fluctuation characteristics and optimization improvement of dual-fuel engines. Energy. 2025; 324: 135953. https://doi.org/10.1016/j.energy.2025.135953

[9] Huang K, Yang Y, Huang H, Tang Y, Han W, Huang J, et al. Machine Learning Accelerates Discovery of High-Performance Corrole Photosensitizers for Optical Imaging Diagnosis and Photodynamic Therapeutics of Nasopharyngeal Carcinoma. The Journal of Physical Chemistry Letters. 2025; 16: 12579–12588. https://doi.org/10.1021/acs.jpclett.5c03041

[10] Yashima M, Ali R. Structural phase transition and octahedral tilting in the calcium titanate perovskite CaTiO3. Solid State Ionics. 2009; 180: 120–126. https://doi.org/10.1016/j.ssi.2008.11.019

[11] Randall CA, Bhalla AS. Nanostructural-property relations in complex lead perovskites. Japanese journal of applied physics. 1990; 29: 327.

[12] Cohen RE. Origin of ferroelectricity in perovskite oxides. Nature. 1992; 358: 136–138. https://doi.org/10.1038/358136a0

[13] Sun Q, Yang L, Hu F, Liang C, Wang H, Qi J. Microstructural evolution and kinetics of monoclinic scheelite-type BiVO4-based materials with enhanced color performance. Ceramics International. 2025; 51: 40414–40430. https://doi.org/10.1016/j.ceramint.2025.06.276

[14] Ling Z, Li S, Liu J, Chen W, Fu Z. Critical Nb induced interfacial armoring for enhanced molten Al tribo-corrosion resistance in Ti-Nb-Cr refractory alloys. Wear. 2025; 584–585: 206439. https://doi.org/10.1016/j.wear.2025.206439

[15] Xu K, Fan L, Chen C, Shen C, Jiang Z, Wei Y. Analysis of dynamic coupling characteristics and multi-constraint optimization of a proton exchange membrane fuel cell considering membrane

degradation. Fuel. 2026; 404, 136275. https://doi.org/10.1016/j.fuel.2025.136275
[16] Haertling GH. Ferroelectric ceramics: history and technology. Journal of the American Ceramic Society. 1999; 82, 797–818. https://doi.org/10.1111/j.1151-2916.1999.tb01840.x
[17] Ward L, Agrawal A, Choudhary A, Wolverton C. A general-purpose machine learning framework for predicting properties of inorganic materials. npj Computational Materials. 2016; 2: 16028. https://doi.org/10.1038/npjcompumats.2016.28
[18] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Materials. 2013; 1: 011002. https://doi.org/10.1063/1.4812323
[19] Schütt KT, Arbabzadah F, Chmiela S, Müller KR, Tkatchenko A. Quantum-chemical insights from deep tensor neural networks. Nature Communications. 2017; 8: 13890. https://doi.org/10.1038/ncomms13890
[20] Wang X, Zhou L, He M, Peng Z, Zhang J. XGboost with multi-feature fusion for hemodynamic state prediction. Neuroscience. 2025; 591: 63–75. https://doi.org/10.1016/j.neuroscience.2025.10.054
[21] Li Z, Jiang S, Xiang F, Li C, Li S, Gao T, et al. Enhancing vitiligo stage diagnosis through a reliable multimodal model with uncertainty calibration. Applied Intelligence. 2025; 55: 986. https://doi.org/10.1007/s10489-025-06839-x
[22] Su H, Chi L, Zio E, Li Z, Fan L, Yang Z, et al. An integrated, systematic data-driven supply-demand side management method for smart integrated energy systems. Energy. 2021; 235: 121416. https://doi.org/10.1016/j.energy.2021.121416
[23] Xie T, Grossman JC. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters. 2018; 120: 145301. https://doi.org/10.1103/PhysRevLett.120.145301
[24] Peng XL, Fathidoost M, Lin B, Yang Y, Xu BX. What can machine learning help with microstructure-informed materials modeling and design? MRS Bulletin. 2025; 50: 61–79. https://doi.org/10.1557/s43577-024-00797-4
[25] Pilania G, Mannodi-Kanakkithodi A, Uberuaga BP, Ramprasad R, Gubernatis JE, Lookman T. Machine learning bandgaps of double perovskites. Scientific Reports. 2016; 6: 19375. https://doi.org/10.1038/srep19375
[26] Liu Y, Kelley KP, Vasudevan RK, Funakubo H, Ziatdinov MA, Kalinin SV. Experimental discovery of structure-property relationships in ferroelectric materials via active learning. Nature Machine Intelligence. 2022; 4: 341–350. https://doi.org/10.1038/s42256-022-00460-0
[27] Wang J, Cui Z, Zhang X, Zhao J, Li F, Zhou Z, et al. Prediction of dielectric properties of ferroelectric materials based on deep neural networks. Science Progress. 2025. https://doi.org/10.1177/00368504251320846
[28] Laosiritaworn W, Wongdamnern N, Yimnirun R, Laosiritaworn Y. Concurrent artificial neural network modeling of single-crystal and bulk-ceramics ferroelectric-hysteresis: an application to barium titanate. Ferroelectrics. 2011; 414: 90–96. https://doi.org/10.1080/00150193.2011.577313
[29] Laosiritaworn W, Yimnirun R, Laosiritaworn Y. The knowledge-based modeling of ferroelectric hysteresis area: an application to forming (1-x)PZT-(x)PZN hysteresis database. Integrated Ferroelectrics. 2015; 166: 65–73. https://doi.org/10.1080/10584587.2015.1092199
[30] Isayev O, Fourches D, Muratov EN, Oses C, Rasch K, Tropsha A, et al. Materials cartography: Representing and mining materials space using structural and electronic fingerprints. Chemistry of Materials. 2015; 27: 735–743. https://doi.org/10.1021/cm503507h
[31] Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. npj Computational Materials. 2022; 8: 59. https://doi.org/10.1038/s41524-022-00734-6
[32] Tao Q, Xu P, Li M, Lu W. Machine learning for perovskite materials design and discovery. npj Computational Materials. 2021; 7: 23. https://doi.org/10.1038/s41524-021-00495-8
[33] Li C, Hao H, Xu B, Shen Z, Zhou E, Jiang D, et al. Improved physics-based structural descriptors of perovskite materials enable higher accuracy of machine learning. Computational Materials Science. 2021; 198: 110714. https://doi.org/10.1016/j.commatsci.2021.110714
[34] Yasui K. Merits and demerits of machine learning of ferroelectric, flexoelectric, and electrolytic properties of ceramic materials. Materials. 2024; 17: 2512. https://doi.org/10.3390/ma17112512
[35] Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Computational Materials. 2017; 3: 54. https://doi.org/10.1038/s41524-017-0056-5
[36] Li C, Zheng K. Methods, progresses, and opportunities of materials informatics. InfoMat. 2023; 5: e12425. https://doi.org/10.1002/inf2.12425
[37] Bai Y, Jantunen H, Juuti J. Energy harvesting research: the road from single source to multisource. Advanced materials. 2018; 30: 1707271.
[38] He J, Li J, Liu C, Wang C, Zhang Y, Wen C, et al. Machine learning identified materials descriptors for ferroelectricity. Acta Materialia. 2021; 209: 116815. https://doi.org/10.1016/j.actamat.2021.116815
[39] Megaw HD. Ferroelectricity in Crystals. Methuen: London. 1957.
[40] Goldschmidt VM. Die Gesetze der Krystallochemie. Naturwissenschaften. 1926; 14: 477–485. https://doi.org/10.1007/BF01507527
[41] Shannon RD. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Foundations of Crystallography. 1976; 32: 751–767.
[42] Pertsev NA, Zembilgotov AG, Tagantsev AK. Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. Physical review letters. 1998; 80:1988.
[43] Frey MH, Payne DA. Grain-size effect on structure and phase transformations for barium titanate. Physical Review B. 1996; 54: 3158–3168. https://doi.org/10.1103/PhysRevB.54.3158
[44] Kittel C. Solid State Physics. 8th ed. John Wiley & Sons: Hoboken, NJ, USA. 2005.
[45] Jonscher AK. Dielectric relaxation in solids. Journal of Physics D: Applied Physics. 1999; 32: R57–R70. https://doi.org/10.1088/0022-3727/32/14/201
[46] Ben Ishai P, Talary MS, Caduff A, Levy E, Feldman Y. Electrode polarization in dielectric measurements: a review. Measurement Science and Technology. 2013; 24: 102001. https://doi.org/10.1088/0957-0233/24/10/102001
[47] Prosandeev S, Ponomareva I, Naumov I, Kornev I, Bellaiche L. Original properties of dipole vortices in zero-dimensional ferroelectrics. Journal of Physics: Condensed Matter. 2008; 20: 193201.
[48] Kuroiwa Y, Aoyagi S, Sawada A, Harada J, Nishibori E, Takata M, et al. Evidence for Pb-O covalency in tetragonal PbTiO 3. Physical review letters. 2001; 87: 217601.
[49] Jolliffe IT. Principal Component Analysis. In Lovric M (ed.) International Encyclopedia of Statistical Science. Springer: Berlin, Heidelberg. 2002. https://doi.org/10.1007/978-3-662-69359-9_483
[50] Sato N, Takahashi A, Kiyohara S, Terayama K, Tamura R, Oba F. Target material Property-Dependent cluster analysis of inorganic compounds. Advanced Intelligent Systems. 2024; 6: 2400253.

https://doi.org/10.1002/aisy.202400253
[51] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521: 436–444. https://doi.org/10.1038/nature14539
[52] Bishop CM, Nasrabadi NM. Pattern Recognition and Machine Learning. Springer: New York. 2006.
[53] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017; 30: 4765–4774.
[54] Yuan R, Wang B, Li J, Sun P, Liu Z, Ding X, et al. Machine learning-enabled design of ferroelectrics with multiple properties via a Landau model. Acta Materialia. 2025; 286: 120760. https://doi.org/10.1016/j.actamat.2025.120760
[55] Schmidt J, Marques MR, Botti S, Marques MA. Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials. 2019; 5: 83. https://doi.org/10.1038/s41524-019-0221-0
[56] Damewood J, Karaguesian J, Lunger JR, Tan AR, Xie M, Peng J, et al. Representations of materials for machine learning. Annual Review of Materials Research. 2023; 53: 399-426. https://doi.org/10.1146/annurev-matsci-080921-085947
[57] Kuhn M, Johnson K. Applied Predictive Modeling. Springer: New York. 2013.
[58] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th international joint conference on Artificial intelligence. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 1995.
[59] Breiman L. Random forests. Machine Learning. 2001; 45: 5–32. https://doi.org/10.1023/A:1010933404324
[60] Friedman JH. Greedy function approximation: A gradient boosting machine. The Annals of Statistics. 2001; 29: 1189–1232.
[61] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory. 1967; 13: 21–27. https://doi.org/10.1109/TIT.1967.1053964
[62] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer: New York, NY, USA. 2009
[63] Shahzad K, Khan MN, Bashir J. Synchrotron X-ray diffraction studies of pure and A- and B-site cation doped CaTiO3. Ferroelectrics. 2011; 414: 162–169. https://doi.org/10.1080/00150193.2011.577334
[64] Eklund CJ, Fennie CJ, Rabe KM. Strain-induced ferroelectricity in orthorhombic CaTiO3 from first principles. Physical Review B. 2009; 79: 220101. https://doi.org/10.1103/physrevb.79.220101
[65] Jarin S, Yuan Y, Zhang M, Hu M, Rana M, Wang S, et al. Predicting the crystal structure and lattice parameters of the perovskite materials via different machine learning models based on basic atom properties. Crystals. 2022; 12: 1570. https://doi.org/10.3390/cryst12111570
[66] Prodromakis T, Papavassiliou C. Engineering the Maxwell-Wagner polarization effect. Applied Surface Science. 2009; 255: 6989–6994. https://doi.org/10.1016/j.apsusc.2009.03.030
[67] Yurtseven H, Kiraci A. Temperature dependence of the polarization and the dielectric constant near the paraelectric-ferroelectric transitions in BaTiO3. Journal of Molecular Modeling. 2013; 19: 3925–3930. https://doi.org/10.1007/s00894-013-1927-4
[68] Sahoo S, Dash U, Parashar SKS, Ali SM. Frequency and temperature dependent electrical characteristics of CaTiO3 nanoceramic prepared by high-energy ball milling. Journal of Advanced Ceramics. 2013; 2: 291–300. https://doi.org/10.1007/s40145-013-0075-8
[69] Pullar RC, Penn SJ, Wang X, Reaney IM, Alford NM. Dielectric loss caused by oxygen vacancies in titania ceramics. Journal of the European Ceramic Society. 2009; 29: 419–424. https://doi.org/10.1016/j.jeurceramsoc.2008.06.019
[70] Qin J, Liu Z, Ma M, Li Y. Machine learning approaches for permittivity prediction and rational design of microwave dielectric ceramics. Journal of Materiomics. 2021; 7: 1284–1293. https://doi.org/10.1016/j.jmat.2021.02.012
[71] Tagantsev AK, Cross LE, Fousek J. Domains in ferroic crystals and thin films. Springer: New York. 2010. https://doi.org/10.1007/978-1-4419-1417-0
[72] Michaloglou A, Papadimitriou I, Gialampoukidis I, Vrochidis S, Kompatsiaris I. Physics-informed neural networks in materials modeling and design: a review. Archives of Computational Methods in Engineering. 2026; 33: 5223–5260. https://doi.org/10.1007/s11831-025-10448-9
[73] Kumari M, Shukla A. Machine learning-driven prognostication of multifunctional properties of (K,Na)NbO3-based lead-free ceramics for optimized materials design. Computational Materials Science. 2026; 263: 114437. https://doi.org/10.1016/j.commatsci.2025.114437
[74] Jain A. Machine learning in materials research: developments over the last decade and challenges for the future. Current Opinion in Solid State and Materials Science. 2024; 33: 101189. https://doi.org/10.1016/j.cossms.2024.101189
[75] Singh JP, Gautam S, Kumar P, Tripathi A, Chen JM, Chae KH, et al. Correlation between the dielectric properties and local electronic structure of copper doped calcium titanate. Journal of Alloys and Compounds. 2013; 572: 84–89. https://doi.org/10.1016/j.jallcom.2013.03.070
[76] Ni Z, Wang F, Guan Y, Hu K, Zhao F, Chen X, et al. Acceleration of optimizing the dielectric properties of (Ca0. 7Nd0. 2) 1− x (Li0. 5Nd0. 5) xTi1− y (Mg1/3Nb2/3) yO3 with the aid of machine learning. Journal of the American Ceramic Society. 2024; 107: 3380–3389. https://doi.org/10.1111/jace.19674
[77] Kumar K, Chowdhury A. Reviewing the cases of nanoscale heterogeneity in ceramics: boon or bane? Materialia. 2021; 16: 101109. https://doi.org/10.1016/j.mtla.2021.101109
[78] Duan C, Nandy A, Kulik HJ. Machine learning for the discovery, design, and engineering of materials. Annual Review of Chemical and Biomolecular Engineering. 2022; 13: 405–429. https://doi.org/10.1146/annurev-chembioeng-092320-120230
[79] Liu Y, Zhao T, Ju W, Shi S. Materials discovery and design using machine learning. Journal of Materiomics. 2017; 3: 159–177. https://doi.org/10.1016/j.jmat.2017.08.002
[80] McCullough K, Williams T, Mingle K, Jamshidi P, Lauterbach J. High-throughput experimentation meets artificial intelligence: a new pathway to catalyst discovery. Physical Chemistry Chemical Physics. 2020; 22: 11174-11196. https://doi.org/10.1039/d0cp00972e
[81] Saidi WA, Shadid W, Castelli IE. Machine-learning structural and electronic properties of metal halide perovskites using a hierarchical convolutional neural network. npj Computational Materials. 2020; 6: 36. https://doi.org/10.1038/s41524-020-0307-8
[82] Vasudevan R, Pilania G, Balachandran PV. Machine learning for materials design and discovery. Journal of Applied Physics. 2021; 129: 070401. https://doi.org/10.1063/5.0043300