Pioneer. Adv. Mater. 2026, 2(1): 1-9 DOI: 10.48308/piadm.2026.245349.1028

Shahid Beheshti University

Pioneering Advances in Materials

Journal homepage: http://pioneering.sbu.ac.ir

ISSN: 3115-7114

Research paper

Machine Learning-Assisted Development and Optimization of Zr-Doped NiCo LDH Catalyst for Oxygen Evolution

Ali Soltani[1], Mahboubeh Tasviri[1]*, Neda Askari[2], Parvaneh Rahimi[3]

[1]*Department of Physical Chemistry, Faculty of Chemistry and Petroleum Sciences, Shahid Beheshti University, Tehran, Iran*
[2]*Department of Physics, Sharif University of Technology, Tehran, Iran*
[3]*Institute of Nanoscale and Biobased Materials, Faculty of Materials Science and Technology, Technische Universitat Bergakademie Freiberg, Freiberg, Germany*

*m_tasviri@sbu.ac.ir

Article info:

Article history:

Received: 12/07/2026
Accepted: 09/08/2026

Keywords:
Layered double hydroxides, Oxygen evolution reaction, Machine learning, AdaBoost regressor, Electrocatalytic water splitting, Zr-doped NiCo LDH.

Abstract

Electrocatalytic water splitting offers a promising path toward sustainable hydrogen production, yet the oxygen evolution reaction (OER) remains a major bottleneck due to its inherently slow kinetics. While layered double hydroxides (LDHs) have emerged as effective non-precious-metal electrocatalysts, optimizing their multicomponent compositions is difficult given the vast number of compositional and synthetic parameters involved. To address this, an AdaBoost-based ensemble regression model (ABR) was built to predict OER overpotential at 10 mA·cm-2 using compositional descriptors, synthesis routes, and electrochemical parameters as inputs. The model enabled efficient screening of candidate LDH compositions and facilitated the identification of promising catalyst designs. Shapley Additive Explanations (SHAP) were then applied to pinpoint the compositional and synthetic factors most responsible for catalytic performance and to provide insights into the relative importance of different descriptors. Despite the small tabular dataset, the ABR model demonstrated satisfactory performance (MAE ≈ 33 mV) and recommended a nickel foam substrate, hydrothermal synthesis, and optimized LDH molar fractions. Based on these findings, a Zr-doped NiCo LDH ($Ni_{0.75}Co_{0.20}Zr_{0.05}$) was engineered, achieving an overpotential of 234.5 mV at 10 mA·cm-2 along with a Tafel slope of 100.3 mV·dec-1. The experimentally measured performance confirmed the effectiveness of the machine learning-guided design strategy. Together, this machine learning-experimental strategy provides a systematic pathway for optimizing LDH-based electrocatalysts in sustainable energy conversion applications.

1. Introduction

There is a pressing need to transition away from fossil fuels because of their significant driver of global warming and climate change. Rising greenhouse gas emissions are causing severe environmental consequences, including higher global temperatures, more frequent extreme weather events, and continued sea-level rise. In this context, renewable hydrogen has emerged as a viable alternative, offering a clean and sustainable energy carrier for future energy systems [1]. Hydrogen acts as an efficient energy carrier and, when produced from renewable sources, offers an attractive solution to environmental challenges [2]. Electrocatalytic water splitting (WS) is a promising route for hydrogen production, involving the hydrogen evolution reaction (HER) and the oxygen evolution reaction (OER). However, the slow kinetics and substantial overpotential of the OER significantly limit the overall efficiency of water electrolysis [3, 4]. To address this challenge, numerous compounds, including oxides, hydroxides, sulfides, and phosphides, have been investigated as potential OER electrocatalysts, many of which exhibit promising catalytic activity. Among these materials, layered double hydroxides (LDHs) have attracted considerable attention

due to their layered structure, tunable composition, and excellent catalytic performance in water oxidation [5, 6, 7].

The development of cost-effective catalysts based on transition metals has become a primary focus in recent years. Machine learning (ML) has rapidly evolved as a powerful tool to accelerate material screening. By revealing relationships between structural features and electrocatalytic activity, ML enables the rational design of advanced materials. A wide range of models, from classical estimators like Random Forest to

advanced algorithms such as Artificial Neural Networks (ANNs), have shown strong performance in predicting key parameters, including OER overpotential (η_{OER}) and double-layer capacitance (C_{dl}), in WS systems [8, 9]. Density Functional Theory (DFT) is a key method for modeling the atomic-scale behavior of catalysts and understanding their reaction mechanisms. Combining DFT with ML has recently become a powerful approach for the fast screening of large material spaces. For instance, a hybrid DFT-ML framework has improved the fabrication of graphene nanoribbons for the HER [10]. Beyond discovery, ML also aids in assessing nanocomposite stability for WS [11]. Moreover, in the pursuit of sustainable energy, ML-guided band engineering and solid-solution alloying approaches are expediting the exploration of multicomponent alloy photocatalysts for WS, efficiently navigating their vast compositional space [12].

While ML is increasingly used in electrocatalysis, mixing different material types (like layered LDHs and non-layered oxides) into one dataset can be problematic. If the model does not receive specific information about these structural differences, it cannot accurately learn how structure affects performance. Furthermore, using Label Encoding (assigning numbers like 1, 2, 3 to different phases) can mislead the model. Since these numbers are arbitrary, the algorithm might incorrectly assume a mathematical order or priority between categories (e.g., thinking 3 is 'greater' than 2), leading to predictions based on false numerical patterns rather than actual chemical principles[8, 13].

Methodologically, feature selection and model evaluation are frequently biased toward linear dependencies. Although non-linear algorithms (e.g., XGBoost) are routinely employed to capture complex catalytic behavior, they are often assessed using linear statistical measures such as the Pearson correlation coefficient. This practice risks eliminating physically meaningful descriptors that exhibit non-linear or non-monotonic relationships despite weak linear correlation [9, 14].

Moreover, the reliability of the coefficient of determination (R^2) is often compromised in small-scale studies. Due to its susceptibility to artificial inflation, R^2 may capture spurious correlations rather than genuine physical trends, leading to an overestimation of generalization capability [8, 9, 15, 16]. To address these limitations, a more rigorous evaluation framework is essential. This requires prioritizing scale-dependent metrics, such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), which offer a more transparent assessment of predictive accuracy. Furthermore, to ensure robust performance on limited datasets, Nested Cross-Validation (NCV) should be employed instead of conventional train-test splits or simple K-fold validation. By separating the hyperparameter tuning phase from the model evaluation phase, NCV effectively mitigates bias and provides a more reliable estimate of how the model will perform on unseen materials.

In response to these shortcomings, we curated a refined and well-preprocessed dataset of nickel (Ni)-based LDHs to train several supervised regression models, including AdaBoost, XGBoost, and CatBoost, for predicting η_{OER}. The models used key descriptors such as chemical composition, support type, preparation method, and other relevant parameters. AdaBoost was selected as the primary predictive model in this study. This choice was justified by its superior accuracy, achieving an MAE of approximately 33 mV. Given that the target overpotential range spans from 172 to 475 mV, this reflects a relative error of nearly 11%, which is considered acceptable for screening electrocatalysts. Additionally, AdaBoost's requirement for fewer hyperparameters simplified the tuning process, reducing the risk of overfitting while maintaining predictive power. The top candidate was subsequently synthesized, characterized, and evaluated for its electrocatalytic performance [17, 18]. Notably, this approach led to the discovery of a high-performance Zr-doped NiCo LDH ($Ni_{0.75}Co_{0.20}Zr_{0.05}$), which exhibited a low overpotential of 234.5 mV at 10 $mA.cm^{-2}$. The close

agreement between the predicted and experimental results supports the utility of this data-driven framework for the efficient screening of advanced electrocatalysts.

2. ML Section

2.1 Preparation of Data

A dataset of 104 binary and ternary Ni-based LDHs was compiled from over 50 literature sources. The target variable for our prediction modeling was η_{OER} (at 10 $mA.cm^{-2}$). The input features comprised both material and experimental factors, including morphology, substrate type, synthesis method, elemental composition (represented by molar fractions and atomic numbers), and linear sweep voltammetry (LSV) scan rates.

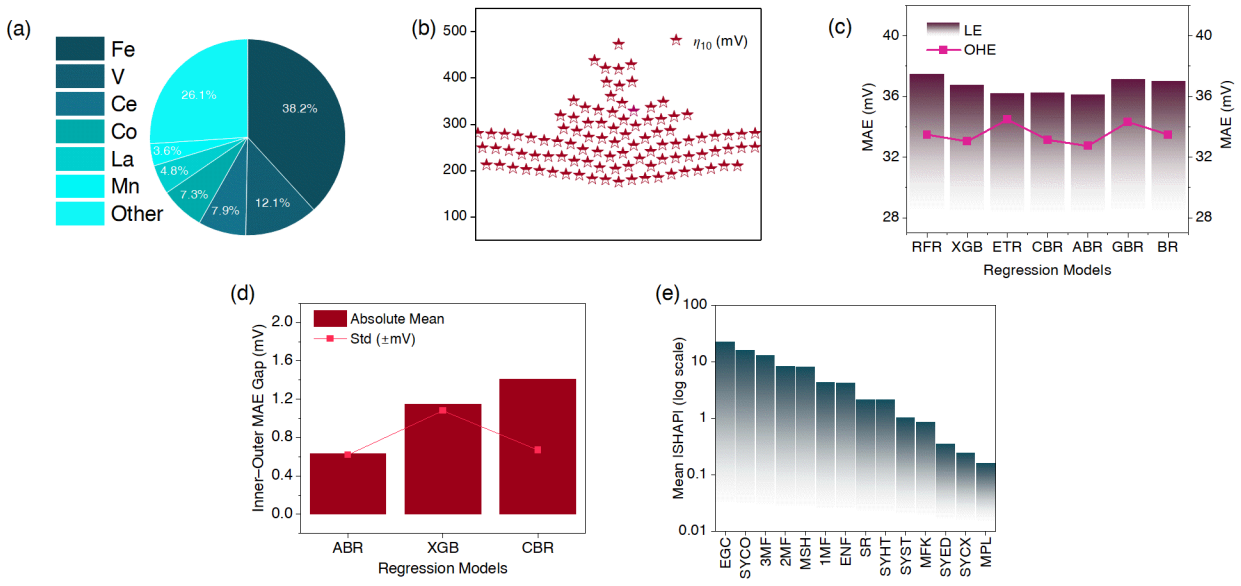

Figure 1. An overview of the dataset and model performance; (a) Distribution of elements; (b) Distribution of the overpotential (target value); (c) Comparison of predictive performance for seven regression models across Label Encoding (LE) and One-Hot Encoding (OHE) strategies. Performance is evaluated using mean Outer MAE. All results are averaged over three independent seeds using an NCV framework on the full dataset. The evaluated models include Random Forest (RFR), Extreme Gradient Boosting (XGB), Extra Trees (ETR), CatBoost (CBR), AdaBoost (ABR), Gradient Boosting (GBR), and Bagging (BR) regressors; (d) Consistency analysis. Average differences between inner and outer MAE scores for ABR, XGB, and CBR models (averaged over seven seeds). The standard deviation line represents the variability of the MAE gap across these seven seeds; (e) Mean SHAP analysis of top non-elemental features; EGC (glassy carbon), SYCO (coprecipitation), 3MF (third molar fraction), 2MF (second molar fraction), MSH (sheet-like), 1MF (first molar fraction), ENF (nickel foam), SR (LSV scan rate), SYHT (hydrothermal), SYST (solvothermal), MFK (flake-like), SYED (electrodeposition), SYCX (cation-exchange), MPL (plate-like).

All η_{OER} values were documented under uniform conditions using a 1.0 M KOH aqueous electrolyte. Fig. 1a, b illustrates the statistical distributions of the descriptive characteristics alongside the target variable.

2.2 Feature Engineering

The choice between LE and one-hot encoding (OHE) can significantly influence model performance [8, 9, 13]. While LE may introduce artificial ordinal relationships among categorical variables, OHE avoids this issue by representing each category independently; however, it increases the feature space and may lead to higher dimensionality. To investigate this effect, a comparative analysis was conducted using both encoding approaches. Model functionality was evaluated using MAE. The results indicated that OHE consistently outperformed LE, leading us to adopt it for all subsequent modeling tasks. As shown in Fig. 1c, the MAE scores for all algorithms vary with the encoding strategies.

Additionally, the molar fractions were normalized to a range between 0 and 1 to minimize the bias associated with independently feeding unscaled compositional data into the model. For example, the compositions nickel:iron:vanadium (Ni:Fe:V) (0.5:0.3:0.2) and Ni:Fe:V (5:3:2) represent the same relative ratios; however, without normalization, the model would treat them as distinct.

2.3 Model Selection

For this study, several regression estimators were employed. The specific algorithms evaluated included Random Forest Regressor (RFR), Extreme Gradient Boosting Regressor (XGB), Extra Trees Regressor (ETR), CatBoost Regressor (CBR), Adaptive Boosting Regressor (ABR), Gradient Boosting Regressor (GBR), and Bagging Regressor (BR).

2.4 Cross Validation and Hyperparameter Optimization (HPO)

Achieving high accuracy often requires careful optimization strategies, particularly when working with small datasets [19, 20]. Model performance was evaluated using three statistical metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). MAE was considered the primary metric because it provides a direct and scale-dependent measure of the average prediction error in mV, which is particularly informative for small datasets. RMSE was also reported as a complementary metric because it gives greater weight to larger deviations.

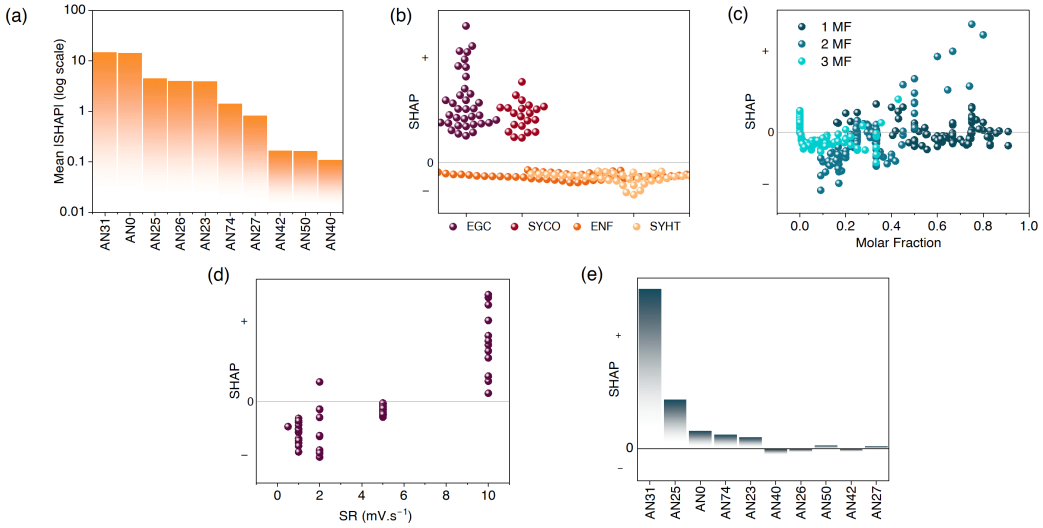

Figure 2. (a) Mean SHAP analysis of top elemental features; AN31 (Ga), AN0 (binary LDHs), AN25 (Mn), AN26 (Fe), AN23 (V), AN74 (W), AN27 (Co), AN42 (Mo), AN50 (Sn), AN40 (Zr); (b, c, d and e) SHAP dependence plots for the (b) ENF, EGC, SYHT and SYCO (c) 1MF, 2MF, and 3MF (d) SR; (e) elemental features.

In contrast, R^2 was interpreted more cautiously, as it can vary substantially in limited datasets depending on the target distribution across folds.

To obtain an unbiased estimate of model generalization, NCV was employed instead of a conventional 80/20 train-test split. Hyperparameter optimization was performed exclusively in the inner loop using Optuna, while final model performance was assessed on the unseen outer folds. All reported metrics were calculated on the outer folds and averaged across seven independent random seeds.

The top three models (ABR, XGB, and CBR) were initially selected based on their lower outer-fold MAE values. For final model selection, we additionally examined the difference between inner-loop and outer-loop MAE, averaged across the seven seeds, as an indicator of overfitting sensitivity. ABR exhibited the smallest mean inner-outer MAE gap (±0.6 mV), compared with XGB (±1.2 mV) and CBR (±1.4 mV), suggesting more consistent generalization behavior. Moreover, ABR showed the lowest variability in outer-fold MAE across seeds, further supporting its stability. Although the ABR model performed slightly better than XGB and CBR from a statistical standpoint (Table 1), the differences were relatively small; therefore, ABR was preferred because its lower hyperparameter complexity enables a simpler and more robust optimization process.

3. Experimental Section

3.1 Catalyst Synthesis

A 1 × 1 cm^{-2} Ni foam was pretreated by sequential immersion in 2 M HCl, ethanol, and deionized water (DIW) for 10 min each. The sample was synthesized through a one-step hydrothermal method. In a typical procedure, the required metal precursors together with 4 mmol urea were dissolved in 70 mL of DIW, and the resulting solution was transferred into a 100 mL Teflon-lined autoclave containing the pretreated Ni foam. The hydrothermal reaction was conducted at 120 , followed by natural cooling to room temperature. The obtained product was collected, thoroughly washed with ethanol and DIW, and dried at 60 for 6 h. $Ni_{0.75}Co_{0.20}Zr_{0.05}$ was synthesized under these conditions with a reaction time of 15 h, where the subscripts denote the nominal molar amounts of the corresponding metal precursors used in the synthesis.

3.2 Material Characterization

The prepared LDH was studied by Philips X'Pert diffractometer (Cu Kαradiation, λ = 0.15406 nm) to investigate the structural characteristics. For morphological analysis, field-emission scanning electron microscopy (FE-SEM) with a MIRA3 FE-SEM microscope was utilized.

3.3 Electrochemical Measurements

Electrochemical measurements were conducted using an IVIUM potentiostat (Netherlands) in a conventional three-electrode setup. The catalyst-coated Ni foam served as the working electrode, while a Pt wire and an Ag/AgCl (3.0 M KCl) electrode were used as the counter

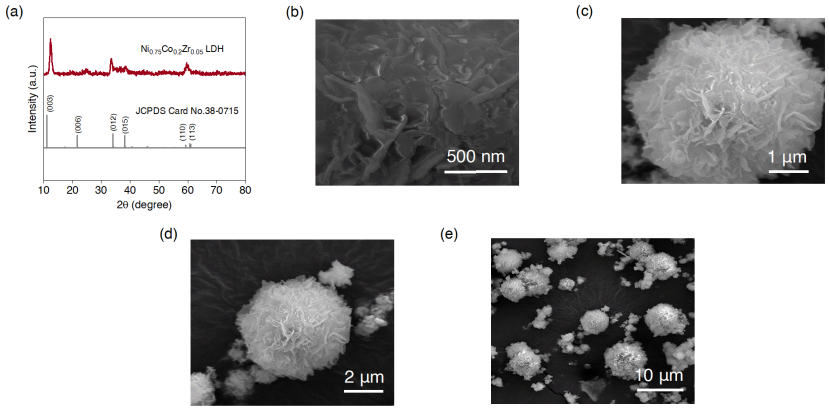

Figure 3. (a) XRD patterns of $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH; (b, c, d and e) SEM image showing the flower-like morphology of the synthesized LDH.

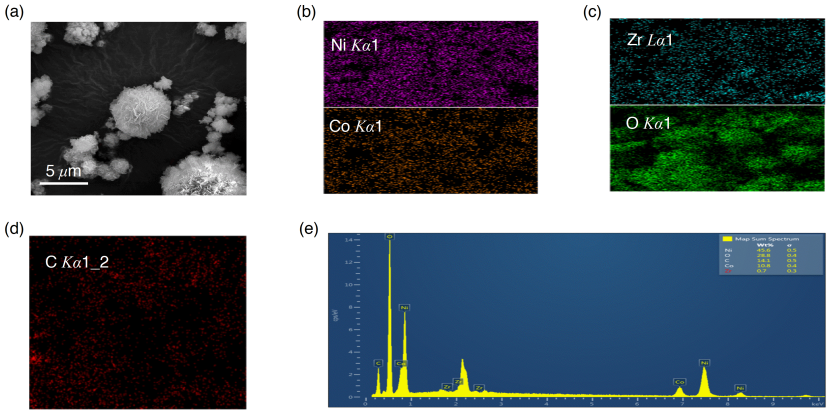

Figure 4. (a) SEM; (b, c, and d) elemental mapping of $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH showing the distribution of Ni, Co, Zr, O, and C elements; (e) EDS spectrum $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH confirming the presence of Ni, Co, Zr, O, and C elements.

Table 1. Comparison of the predictive performance of the ABR, XGB, and CBR models in terms of MAE, RMSE, and R^2.

Model	MAE (mV)	RMSE (mV)	R^2
ABR	33.1	42.2	0.73
CBR	33.7	42.4	0.69
XGB	33.5	41.7	0.70

Note: All results were averaged across seven random seeds and were computed based on the outer folds of the NCV procedure.

and reference electrodes, respectively. All experiments were carried out in 1.0 M KOH aqueous electrolyte (pH = 13.89) at ambient temperature. LSV measurements were performed at a scan rate of 5 $mV.s^{-1}$. The η_{OER} values were determined at a current density of 10 $mA.cm^{-2}$ under 100 % iR compensation, and Tafel slope values were calculated from the corresponding LSV curves. Electrochemical impedance spectroscopy (EIS) measurements were carried out in the frequency range of 100 kHz to 0.01 Hz with an AC amplitude of 10 mV. The electrochemical capacitance of the $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH was evaluated using cyclic voltammetry at scan rates of 10, 20, 30, 40, 50, 60, 70, and 80 $mV.s^{-1}$. Chronoamperometry was subsequently performed to assess the stability of the catalyst.

4. Result and Discussion

4.1 ML Results

4.1.1 Catalyst Discovery

To support candidate selection, SHAP analysis was used to determine the direction of each descriptor's contribution to the target value.

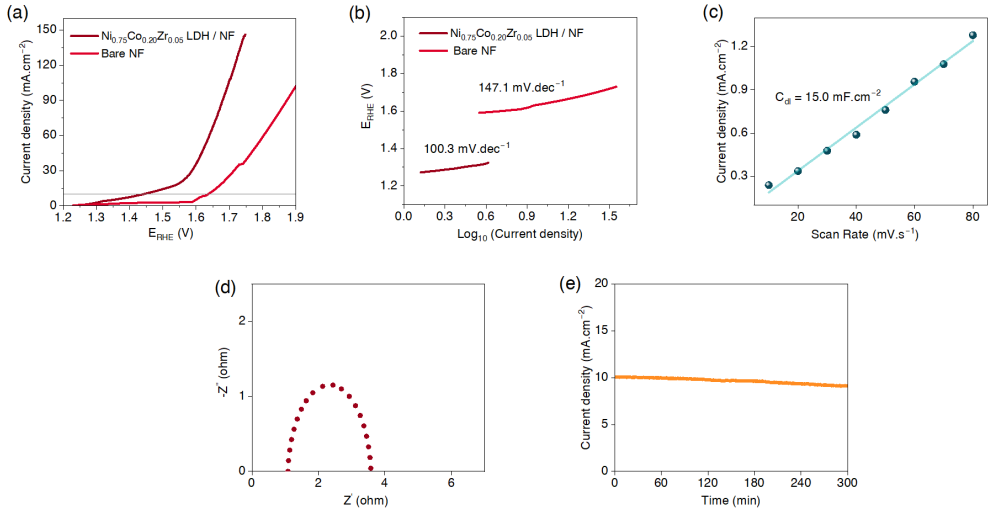

Figure 5. (a) Linear Sweep Voltammetry (LSV) polarization curves comparing the Oxygen Evolution Reaction (OER) performance of the LDH; (b) Corresponding Tafel plots derived from the LSV data, used to evaluate the OER kinetics. (c) Determination of the C_{dl} for the optimal $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH; (d) EIS Nyquist plot; (e) Chronoamperometry test showing long-term operational stability at 1.48 V (vs. RHE).

Table 2. Comparison of experimental and ABR-predicted η_{OER} value, including absolute and relative prediction errors for the synthesized LDH sample.

LDH	Exp. η_{OER} (mV)	Pred. η_{OER} (mV)	Abs. Error (mV)	Rel. Error (%)
$Ni_{0.75}Co_{0.20}Zr_{0.05}$	234.5	225.0	9.5	4.0

4.1.2 Top Features Based on SHAP

Mean SHAP values were employed for each feature to assess its overall impact on the target (η_{OER}). The results shown in Fig. 1e and Fig. 2a highlight the most impactful features.

4.1.3 Feature Analysis and Experimental Guide

Based on the mean SHAP values, the most important features were classified into two main categories: non-elemental features (Fig. 1e) and elemental features (Fig. 2a). Among the non-elemental features, the most influential descriptors were associated with electrode characteristics, molar fractions, and synthesis methods. In particular, the two most important electrode-related features were EGC and ENF, while SYHT and SYCO ranked as the two leading synthesis-related descriptors.

However, mean SHAP values alone indicate only the magnitude of feature importance and do not reveal whether a given feature contributes to an increase or decrease in the target variable, (η_{OER}). Therefore, SHAP dependence analysis is required to determine the direction of each feature's effect. As shown in Fig. 2a, ENF and SYHT exhibit negative contributions, indicating that the presence of these features is associated with lower overpotentials and thus more favorable OER performance. For 1MF (associated with Ni), SHAP values were predominantly negative in the range of 0.7-0.8. For 2MF, most data points were concentrated between 0.1 and 0.3, whereas for 3MF, the majority of the values were below 0.2. Based on these observations, representative values of 0.75 and 0.20, corresponding to the averages of the selected ranges, were chosen for 1MF and 2MF, respectively. The value of 3MF was then determined from the normalization constraint, namely, 3MF = 1 − 1MF − 2MF, since the three fractions had been normalized such that, their sum was equal to 1 (Fig. 2c). In addition, the morphology was fixed as MSH (sheet-like), as it was identified as the most important descriptor within the morphology category.

To ensure cost-effectiveness, precious elements such as Pt, Pd, and Au were excluded from the elemental descriptor set. The elemental features with the highest mean SHAP values are presented in Fig. 2a, with the most influential descriptors corresponding to AN31 (Ga), AN0 (binary LDHs), AN25 (Mn), AN26 (Fe), AN23 (V), AN74 (W), AN27 (Co), AN42 (Mo), AN50 (Sn), and AN40 (Zr).

As shown by the SHAP dependence plots in Fig. 2e, AN31 (Ga), AN25 (Mn), AN74 (W), and AN23 (V) are associated with substantial increases in overpotential, indicating an unfavorable effect on η_{OER}. A similarly unfavorable trend is observed for AN0, suggesting that

Table3. EIS fitting parameters for the Zr-doped NiCo LDH.

LDH	R_s (Ω)	R_{ct} (Ω)	CPE
$Ni_{0.75}Co_{0.20}Zr_{0.05}$	1.08	2.50	0.95

binary LDHs also tend to increase the target value. By contrast, the remaining elemental features exhibit relatively small SHAP contributions.

Among the features showing potentially favorable effects, AN26 (Fe) and AN27 (Co) are particularly noteworthy, as both elements are well known for their catalytic activity and are commonly associated with low Tafel slope values [21, 22, 23]. These compositions were therefore identified as promising candidates for further investigation. In addition, as shown in Fig. 2e, AN40 (Zr) exhibits a negative SHAP contribution, indicating a tendency to reduce the target value relative to the other remaining elements.

Based on these observations, NiFeZr and NiCoZr emerged as the most plausible final LDH compositions. After considering the compositions already represented in the original dataset and aiming to preserve the novelty of the proposed material, NiCoZr was selected for further investigation. Accordingly, the final candidate composition can be approximated as $Ni_{0.75}Co_{0.20}Zr_{0.05}$.

4.2 Experimental Results

4.2.1 Experimental Validation

To identify the ML-suggested candidate, the structural integrity of the synthesized LDH was first verified. As shown in Fig. 3a, the diffraction pattern of NiCoZr LDH generally matches JCPDS card No. 38-0715; however, noticeable peak shifts were observed. These shifts arise from the incorporation of Zr^{4+} into the LDH lattice, where it replaces Ni^{2+}/Co^{2+} in the brucite-like layers. The higher charge of Zr^{4+} increases the layer charge density and strengthens interlayer electrostatic interactions, leading to a contraction of the interlayer spacing, confirming its successful structural incorporation into the LDH [24].

The morphology of $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH was examined. The micrographs (Fig. 3b-e) reveal a well-defined flower-like structure. The chemical composition of as-obtained LDH was further determined using EDS. Fig. 4b-d presents the elemental mapping of the Zr-doped NiCo LDH, showing that Ni, Co, and Zr are homogeneously distributed across the sample surface. Fig. 4e also displays the presence of Ni, Co, and Zr which is consistent with the mapping results.

The LSV curves (Fig. 5a) demonstrate that Zr-doped NiCo LDH exhibits an efficient electrocatalytic performance relative to the bare Ni foam substrate, which required a substantially higher overpotential to reach the same current density. The average overpotential (±Std) required to achieve a current density of 10 $mA.cm^{-2}$ is 234.5 ±11.6 mV with a Tafel slope of 100.3 $mV.dec^{-1}$ (Fig. 5b), markedly lower than the 147.1 $mV.dec^{-1}$ obtained for bare NF, confirming the intrinsic catalytic contribution of the $Ni_{0.75}Co_{0.20}Zr_{0.05}$ LDH beyond that of the underlying substrate. The experimental η_{OER} value was compared with that predicted by the ABR model (Table 2).

From the slope of the current density versus scan rate plot, a C_{dl} value of 15.0 $mF.cm^{-2}$ was obtained (Fig. 5c), indicating a high electrochemically active surface area. EIS was performed to evaluate the charge-transfer resistance (R_{ct}) during the OER process. The Nyquist plot (Fig. 5d) shows a small semicircle diameter for NiCoZr LDH, indicating efficient charge transfer and low interfacial resistance. The measured R_{ct} value is 2.50 Ω(Table 3), which is consistent with the excellent OER activity observed in the LSV and Tafel analyses. Moreover, as shown in Fig. 5e, NiCoZr LDH maintained a steady current density over 5 hours at a constant potential of 1.48 V (vs. RHE) in 1.0 M KOH, demonstrating its excellent durability under OER operating conditions.

Despite these encouraging results, several challenges were identified. A primary limitation of this work is the dataset size (n = 104). Although this scale is consistent with the current state of reported LDH electrochemical data in the open literature, it restricts the representativeness of the feature space, particularly for compositional combinations that are underrepresented or absent from the training set [25, 26, 27]. These considerations reinforce why the present framework should be regarded as a screening and hypothesis-generation tool rather than a stand-alone predictive engine, and why experimental validation remains an indispensable complement to the ML component. The absence of essential features such as impedance data (e.g., charge-transfer resistance) and morphological parameters like particle size or nanosheet thickness further limits model performance. Morphological characteristics are highly sensitive to hydrothermal synthesis conditions, including reaction temperature and duration, which are difficult to accurately model [7, 28]. Consequently, while the model recommended a sheet-like morphology as favorable, the synthesized catalyst instead developed a flower-like structure, underscoring that morphology remains a synthetic outcome rather than a directly controllable design parameter within the

current framework. These limitations highlight the need for current predictive regressors to be closely guided by experimental evidence. Nonetheless, several advantages were observed. Averaging overpotential values across multiple syntheses also enhances result reliability. SHAP analysis revealed that substrate type plays a crucial role, consistent with experimental findings. For instance, the model confirmed the superior performance of NF compared to GC substrates. This aligns with experimental observations showing that in-situ growth on NF improves conductivity, whereas using GC with a binder such as Nafion introduces additional resistance [29, 30, 31]. Moreover, compared to the conventional co-precipitation method, the hydrothermal synthesis approach offers several advantages. The controlled temperature and pressure conditions within the autoclave enable regulated crystallization process, resulting in the formation of highly homogeneous and well-crystallized LDHs [32, 33]. However, certain features, such as the LSV scan rate, exhibited non-monotonic trends. Addressing these challenges underscores the critical importance of comprehensive data reporting in literature, particularly the inclusion of detailed morphological, electrochemical, and synthesis-related parameters to enable the development of more accurate, reliable, and generalizable ML models in materials research.

5. Conclusion

In conclusion, this study provides a systematic data-driven approach for the design and optimization of Ni-based LDH electrocatalysts. The scientific contribution of this work is centered on three key aspects. First, it demonstrates that utilizing a nested cross-validation (NCV) optimized AdaBoost (ABR) model allows for extracting reliable trends from relatively small experimental datasets, maintaining a practical predictive accuracy (MAE of 33 mV). Second, the study elucidates the role of Zr-doping in the LDH framework; the experimental validation of the predicted $Ni_{0.75}$ $Co_{0.20}$ $Zr_{0.05}$ /NF catalyst, which showed an overpotential of 234.5 mV at 10 $mA{\cdot}cm^{-2}$, confirms that Zr^{4+} incorporation can effectively modulate the electrochemical response of NiCo-based structures. Finally, this integrated workflow offers a more efficient alternative to traditional trial-and-error methods, providing a structured methodology for the development of OER catalysts. These findings suggest that the synergy between optimized ML algorithms and targeted synthesis can streamline the discovery process for advanced electrochemical materials.

Conflicts of Interest

There are no conflicts to declare.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors thank the Research Council of Shahid Beheshti University for its financial support.

References

[1] G. Glenk, S. Reichelstein, Economics of converting renewable power to hydrogen, Nat Energy 4 (2019) 216–222. doi:10.1038/s41560-019-0326-1 .

[2] V. R. Stamenkovic, D. Strmcnik, P. P. Lopes, N. M. Markovic, Energy and fuels from electrochemical interfaces, Nature Mater 16 (2017)57–69. . doi:10.1038/nmat4738

[3] H. Ding, H. Liu, W. Chu, C. Wu, Y. Xie, Structural Transformation of Heterogeneous Materials for Electrocatalytic Oxygen Evolution Reaction, Chem. Rev. 121 (2021) 13174–13212. doi:10.1021/acs.chemrev.1c00234 .

[4] C. Spöri, J. T. H. Kwan, A. Bonakdarpour, D. P. Wilkinson, P. Strasser, The Stability Challenges of Oxygen Evolving Catalysts:Towards a Common Fundamental Understanding and Mitigation of Catalyst Degradation, Angew Chem Int Ed 56 (2017) 5994–6021. . doi:10.1002/anie.201608601

[5] A. Hameed, M. Batool, Z. Liu, M. A. Nadeem, R. Jin, Layered Double Hydroxide-Derived Nanomaterials for Efficient Electrocatalytic Water Splitting: Recent Progress and Future Perspective, ACS Energy Lett. 7 (2022) 3311–3328. doi: 10.1021/acsenergylett.2c01362 .

[6] M. Xu, M. Wei, Layered Double Hydroxide-Based Catalysts: Recent Advances in Preparation, Structure, and Applications, Adv Funct Materials 28 (2018) 1802943. doi: 10.1002/adfm.201802943 .

[7] A. A. Altalhi, E. A. Mohamed, N. A. Negm, Recent advances in layered double hydroxide (LDH)-based materials: Fabrication, modification strategies, characterization, promising environmental catalytic applications, and prospective aspects, Energy Adv. 3 (2024) 2136–2151. doi: 10.1039/D4YA00272E .

[8] X. Jiang, Y. Wang, B. Jia, X. Qu, M. Qin, Using Machine Learning to Predict Oxygen Evolution Activity for Transition Metal Hydroxide Electrocatalysts, ACS Appl. Mater. Interfaces 14 (2022) 41141– 41148. doi: 10.1021/acsami.2c13435 .

[9] C. Wei, D. Shi, F. Zhou, Z. Yang, Z. Zhang, Z. Xue, T. Mu, Analysis of the oxygen evolution activity of layered double hydroxides (LDHs) using machine learning guidance, Phys. Chem. Chem. Phys. 25 (2023) 7917–7926. doi: 10.1039/D2CP06052C .

[10] S. Ram, A. S. Lee, S.-C. Lee, S. Bhattacharjee, Advanced Multi-functional Electrocatalysts: Integrating DFT and Machine Learning for OER, HER, and ORR Reactions, Chem. Mater. 37 (2025) 3608– 3621. doi: 10.1021/acs.chemmater.4c03213 .

[11] P. A. Koyale, S. V. Mulik, J. L. Gunjakar, T. D. Dongale, V. B. Koli, N. B. Mullani, S. S. Sutar, Y. G. Kapdi, S. S. Soni, S. D. Delekar, Synergistic Enhancement of Water-Splitting Performance Using MOF-Derived Ceria-Modified g-C3 N4 Nanocomposites: Synthesis, Performance Evaluation, and Sta-

bility Prediction with Machine Learning, Langmuir 40 (2024) 13657–13668. doi:10.1021/acs.langmuir.4c01336

[12] S.-H. V. Oh, S.-H. Yoo, W. Jang, Small dataset machine-learning approach for efficient design space exploration: Engineering ZnTe-based high-entropy alloys for water splitting, npj Comput Mater 10(2024) 166. doi: 10.1038/s41524-024- 01341-3

[13] M. Benjamin, Handling Categorical Variables in Ensemble Algorithms, Ensemble Algorithms (2025).

[14] A. Holzinger, R. Goebel, R. Fong, T. Moon, K.-R. Müller, W. Samek (Eds.), xxAI - Beyond Explainable AI: International Workshop, Held in Conjunction with ICML 2020, July 18, 2020, Vienna, Austria, Revised and Extended Papers, number 13200 in Lecture Notes in Computer Science; Lecture Notes in Artificial Intelligence, Springer Nature, 2022.

[15] R. Ding, J. Chen, Y. Chen, J. Liu, Y. Bando, X. Wang, Unlocking the potential: Machine learning applications in electrocatalyst design for electrochemical hydrogen energy transformation, Chem. Soc. Rev. 53 (2024) 11390–11461. doi: 10.1039/D4CS00844H .

[16] L. Zhang, J. Hou, H. Ji, D. Meng, J. Qi, X. San, Prediction of Oxygen Evolution Activity for FeCoMn Oxide Catalysts via Machine Learning, Catalysts 14 (2024) 513. doi: 10.3390/catal14080513 .

[17] Gavin C. Cawley, Nicola L. C. Talbot, On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation, Journal of Machine Learning Research 11 (2010) 2079–2107.

[18] I. Guyon, A. Saffari, G. Dror, G. Cawley, Model Selection: Beyond the Bayesian/Frequentist Divide, Journal of Machine Learning Research 11 (2010) 61–87.

[19] J. A Ilemobayo, O. Durodola, O. Alade, O. J Awotunde, A. T Olanre- waju, O. Falana, A. Ogungbire, A. Osinuga, D. Ogunbiyi, A. Ifeanyi, I. E Odezuligbo, O. E Edu, Hyperparameter Tuning in Machine Learning: A Comprehensive Review, J. Eng. Res. Rep. 26 (2024) 388–395. doi: 10.9734/jerr/2024/v26i61188 .

[20] F. Maleki, N. Muthukrishnan, K. Ovens, C. Reinhold, R. Forghani, Machine Learning Algorithm Validation, Neuroimaging Clinics of North America 30 (2020) 433–445. doi: 10.1016/j.nic.2020.08.004 .

[21] Z. U. Abideen, M. Malik, W. Wu, Q. Hou, A. U. Rehman, T. Huang, H. Liu, NiFe layered double hydroxides as high-performance electrocatalysts for the oxygen evolution reaction: Recent developments, Nanoscale 17 (2025) 26578–26593. https://doi.org/10.1039/D5NR04120A.

[22] Z. Wang, X. Long, S. Yang, Effects of Metal Combinations on the Electrocatalytic Properties of Transition-Metal-Based Layered Double Hydroxides for Water Oxidation: A Perspective with Insights, ACS Omega 3 (2018) 16529–16541. doi: 10.1021/acsomega.8b02565 .

[23] Y.-C. Zhang, C. Han, J. Gao, L. Pan, J. Wu, X.-D. Zhu, J.-J. Zou, NiCo-Based Electrocatalysts for the Alkaline Oxygen Evolution Reaction: A Review, ACS Catal. 11 (2021) 12485–12509. 10.1021/ acscatal.1c03260.

[24] J. Poonoosamy, F. Brandt, M. Stekiel, P. Kegler, M. Klinkenberg, B. Winkler, V. Vinograd, D. Bosbach, G. Deissmann, Zr-containing layered double hydroxides: Synthesis, characterization, and evaluation of thermodynamic properties, Applied Clay Science 151 (2018) 54–65. doi: 10.1016/j.clay.2017.10.013 .

[25] P. Xu, X. Ji, M. Li, W. Lu, Small data machine learning in materials science, npj Comput Mater 9 (2023) 42. 10.1038/ s41524-023-01000-z.

[26] F. Conrad, H. Wiemer, S. Ihlenfeldt, Leveraging multi-task learning regressor chains for small and sparse tabular data in materials design, Mach. Learn.: Sci. Technol. 6 (2025) 015045. adae53. doi: 10.1088/2632-2153/

[27] C. Kuenneth, A. C. Rajan, H. Tran, L. Chen, C. Kim, R. Ramprasad, Polymer informatics with multi-task learning, Patterns 2 (2021) 100238. doi:10.1016/j.patter.2021.100238

[28] G. M. Tomboc, T. Kim, S. Jung, H. J. Yoon, K. Lee, Modulating the Local Coordination Environment of Single-Atom Catalysts for Enhanced Catalytic Performance in Hydrogen/Oxygen Evolution Reaction, Small 18 (2022) 2105680. doi:10.1002/smll.202105680 .

[29] W. Jiang, W. Lehnert, M. Shviro, The Influence of Loadings and Substrates on the Performance of Nickel-Based Catalysts for the Oxygen Evolution Reaction, ChemElectroChem 10 (2023) e202200991. doi:10.1002/celc.202200991

[30] B. You, Y. Sun, Hierarchically Porous Nickel Sulfide Multifunctional Superstructures, Advanced Energy Materials 6 (2016) 1502333. doi:10.1002/aenm.201502333

[31] P. Zhang, L. Li, D. Nordlund, H. Chen, L. Fan, B. Zhang, X. Sheng, Q. Daniel, L. Sun, Dendritic core-shell nickel-iron-copper metal/metal oxide electrode for efficient electrocatalytic water oxidation, Nat Commun 9 (2018) 381. doi:10.1038/s41467-017-02429-9 .

[32] Y. Yao, S. Sun, H. Zhang, Z. Li, C. Yang, Z. Cai, X. He, K. Dong, Y. Luo, Y. Wang, Y. Ren, Q. Liu, D. Zheng, W. Zhuang, B. Tang, X. Sun, W. W. Hu, Enhancing the stability of NiFe-layered double hydroxide nanosheet array for alkaline seawater oxidation by Ce doping, Journal of Energy Chemistry 91 (2024) 306–312. doi:10.1016/j.jechem.2024.01.011

[33] T. Zhang, Y. Li, W. Zheng, Morphological control synthesis of layered double hydroxides for energy applications, Discov Mater 5 (2025) 191. doi:10.1007/s43939-025-00393-6