

Data-Driven Design and Experimental Validation of High-Precision

Ni-Co Bimetallic Compound-Based Pseudocapacitor Models

Li Zhang^{1,2#*}, Zhuo Zhao^{1#}, Tianhui Dang¹, Xiao Yang¹, Yongzhi Lan¹, Rui Cao¹

¹School of Science, Lanzhou University of Technology, Lanzhou, 730050, Gansu, China.

²State Key Laboratory of Advanced Processing and Recycling of Nonferrous Metals, Lanzhou University of Technology, Lanzhou, 730050, Gansu, China.

*Correspondence should be addressed to Li Zhang: zhangli@lut.edu.cn

[#] Li Zhang and Zhuo Zhao contributed equally to this manuscript.

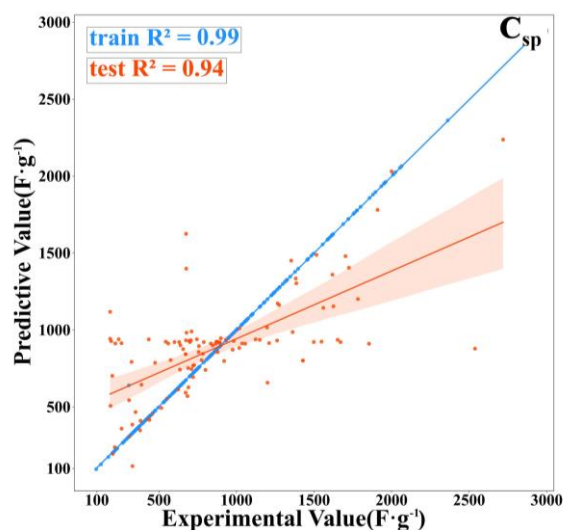


Figure. S1. Regression scatter plot of SVM optimal model

Table S1. Summary of the performance of the various models built in Weka

Weka (5 folds)	Train set				Test set			
Evaluation factor	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>RAE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>RAE</i>
RF	0.77	342.01	213.66	50.58	0.66	363.68	257.78	65.59
ANN (MLP)	0.77	374.87	297.65	70.46	0.38	650.89	423.28	107.69

Note: *R* is the correlation coefficient, assessing the same role as the R^2 value, and *RAE* is the relative absolute error.

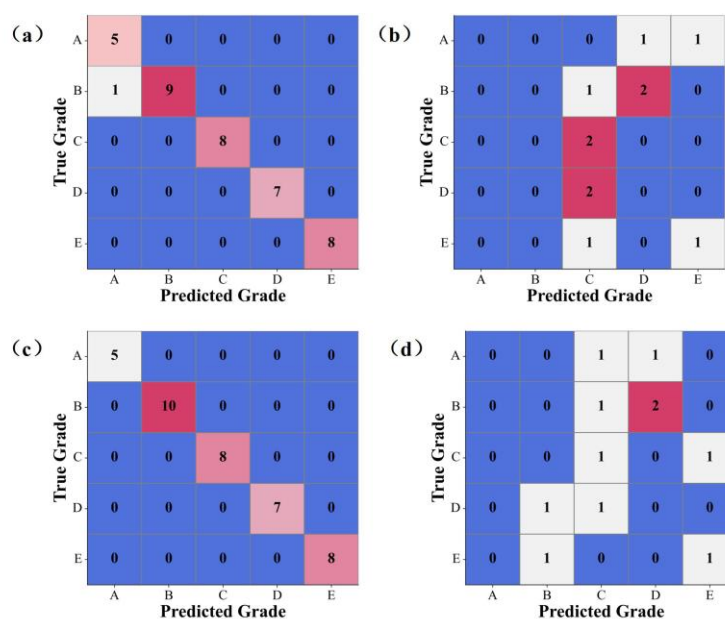


Figure. S2. Confusion matrix for the RFC model (a) training set (b) test set and for the GBC

model (c) training set (d) test set

Note: Confusion matrix clearly and intuitively shows the sample counts comparison between the actual and predicted grades. For example, grade B in Fig. S2 (a) : 9 samples are correctly predicted as B and 1 sample is incorrectly predicted as A, which corresponds to the Precision (accuracy) of grade “B” on the training set in Table 2.1 being 0.92; As shown in Fig. S2 (b), grade A: 0 samples were correctly predicted as A, 1 sample was incorrectly predicted as D, and 1 sample was incorrectly predicted as E. Grade B: 1 sample is incorrectly predicted as C and 2 samples are incorrectly predicted as D. Grade C: 2 samples are correctly predicted to be C, and so on.

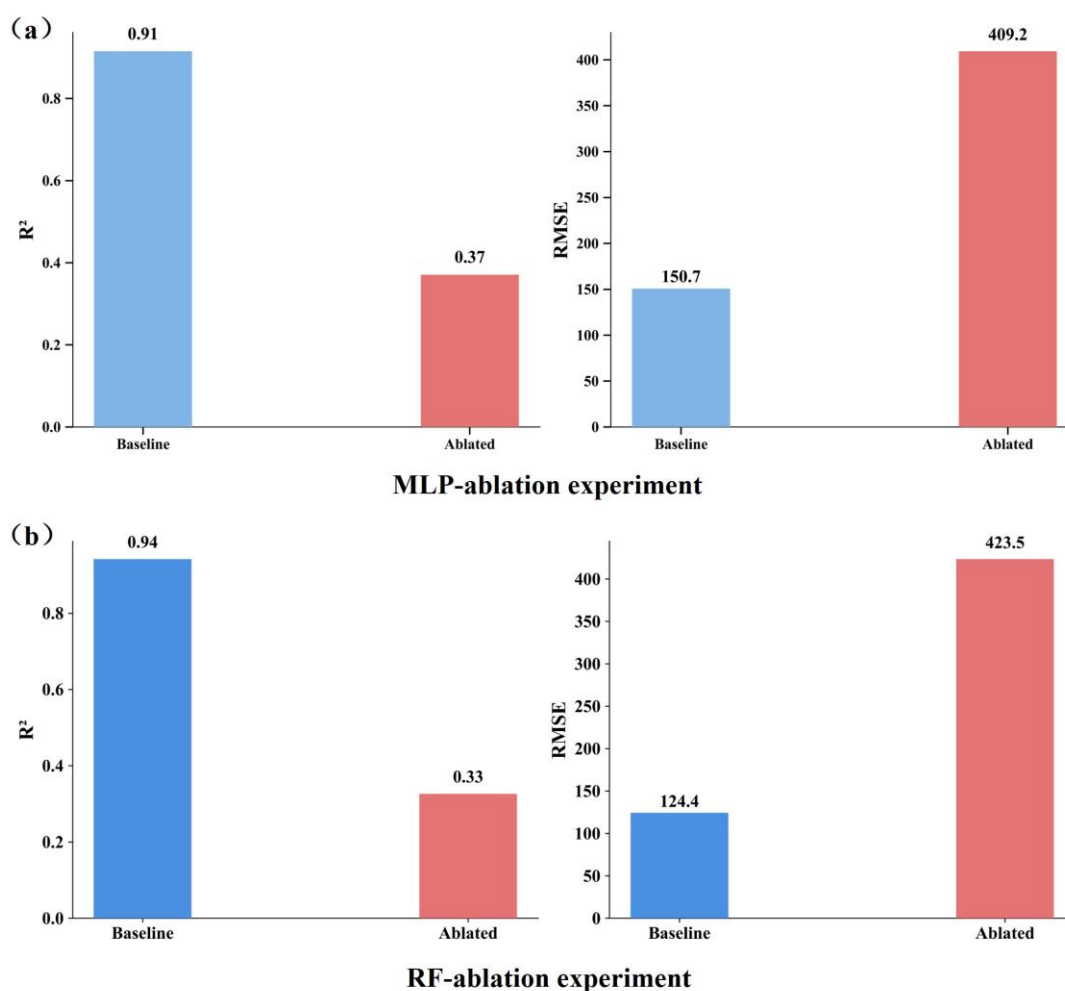


Figure. S3. Two-feature ablation experiments of BET to Ni-Co ratio for (a) MLP model and (b) RF model

Note: We conducted virtual ablation experiments on two highly interactive features-

“Nickel-cobalt ratio” (feature index 0) and “BET” (index 7)-using two different models (MLP and RF), and employed SHAP analysis to validate the quantitative consistency of feature interaction importance. For the MLP model, the baseline test performance reached $R^2 = 0.9147$ and $RMSE = 150.67$. When both features were set to zero, model performance markedly deteriorated to $R^2 = 0.3708$ and $RMSE = 409.23$, resulting in $\Delta R^2 = 0.5439$ and $\Delta RMSE = 258.56$. Meanwhile, the sum of the absolute SHAP values for these two features across the top 50 samples, computed using KernelExplainer, was 145.79-closely matching the increase in RMSE-indicating that SHAP contributions can approximate the impact of feature interaction on prediction error. The ablation results for the Random Forest model showed a similar trend. The baseline performance was $R^2 = 0.9419$ and $RMSE = 124.37$, which dropped sharply to $R^2 = 0.3263$ and $RMSE = 423.47$ after the two features were removed, leading to $\Delta R^2 = 0.6156$ and $\Delta RMSE = 299.10$. These results further confirmed the critical importance of these two features within the RF model. The outcomes of both ablation experiments and SHAP-based interaction analysis were highly consistent at the quantitative level-the SHAP interaction intensity (~ 145.8) and the corresponding increases in RMSE (≈ 258.6 and ≈ 299.1) were strongly aligned, substantiating the reliability of SHAP interaction importance in interpretable machine learning. This experiment clearly demonstrates the pivotal role of the “Nickel – Co ratio” and BET surface area as high-interaction features in predicting specific capacitance (Csp). These findings provide a robust basis for experimental decision-making guided by model outputs and offer quantitative direction for researchers in selecting key adjustable parameters and conducting targeted experiments during practical material design and optimization.

Table S2. Optimal hyperparameter configurations of various models in Anaconda

MLP	GB	RF	SVM	RFC	GBC
<i>hidden_layer_sizes:</i>					
(150)					
<i>activation:'logistic'</i>	<i>n_estimators</i>	<i>max_depth:4</i>			<i>n_estimators:</i>
	:461	9	<i>C:10000</i>		100
<i>solver:'lbfgs'</i>			<i>gamma:</i>	<i>n_estimators:1</i>	
<i>max_iterations:135</i>	<i>learning_rate</i>	<i>min_samples</i>	'auto'	12,	<i>learning_rate:</i>
	<i>e:0.3</i>	<i>_split:2</i>		<i>max_features:</i>	0.01
00			<i>kernel:'rb</i>		
<i>learning_rate:0.000</i>	<i>max_depth:1</i>	<i>n_estimators:</i>	<i>f</i>	None	<i>max_depth:3</i>
	1	15		<i>max_depth:60</i>	<i>min_samples_</i>
<i>tol:1e-07</i>	<i>min_samples</i>	<i>min_samples</i>	<i>epsilon:1e</i>	<i>min_samples_</i>	<i>split:2</i>
	<i>_split:3</i>	<i>_leaf:1</i>	-05	<i>split:4</i>	<i>min_samples_</i>
<i>alpha:1e-08</i>	<i>min_samples</i>	<i>bootstrap:Fa</i>	<i>coef0:0.1</i>	<i>min_samples_l</i>	<i>leaf:1</i>
<i>early_stopping:Fals</i>	<i>_leaf:2</i>	<i>lse</i>	<i>shrinking:</i>	<i>eaf:1</i>	<i>subsample:0.4</i>
<i>e</i>			True		
<i>validation_fraction:</i>	<i>max_feature</i>	<i>max_features</i>	<i>tol:0.0001</i>	<i>bootstrap:True</i>	<i>max_features:</i>
	<i>s:'sqrt'</i>	<i>: 'sqrt'</i>			1.0
0.1					
<i>n_iter_no_change:5</i>					

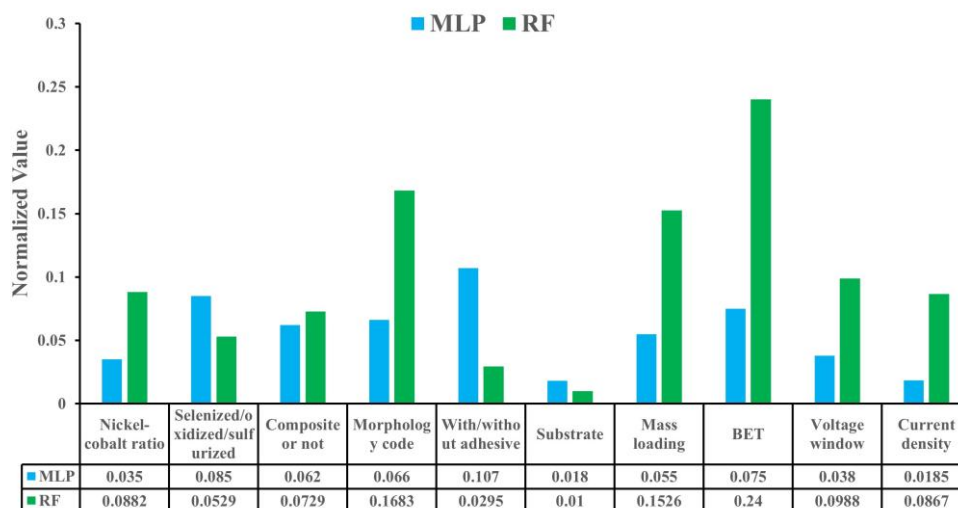


Figure. S4. Summary of MLP and RF importance rankings and corresponding histograms

Note: The relative importance values of the attributes in the MLP model are divided

by 10 to normalize them against the importance values of the attributes in the RF model.

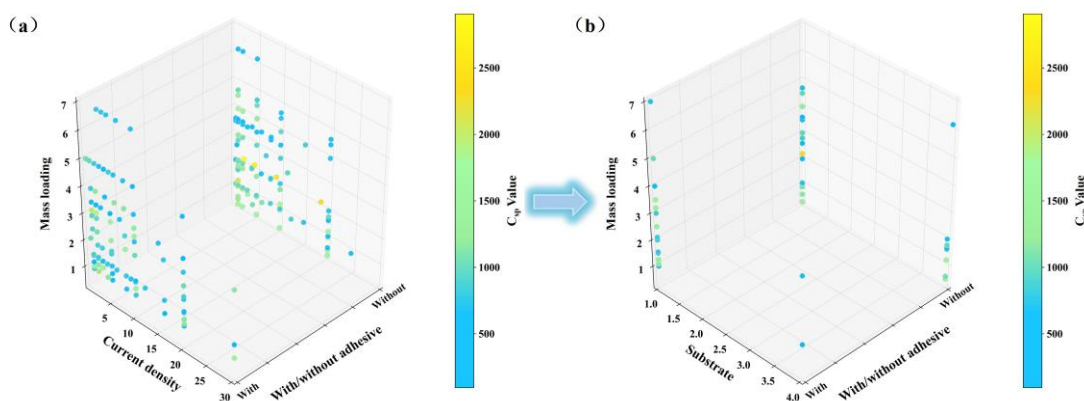


Figure. S5. (a) Trivariate-3D plots of with/without adhesive, current density vs. mass loading, and (b) trivariate-3D plots of with/without adhesive, substrate vs. mass loading

Note: As shown in Fig. S4(a) and (b), the following observations can be made: (i) conventional nickel foam substrates exhibit a higher potential for accommodating greater mass loading; (ii) binder-free fabrication methods allow for a higher upper limit of mass loading, which is consistent with the improved loading capacity enabled by binder-free strategies; (iii) regardless of binder usage, the specific capacitance initially increases and then decreases with increasing mass loading, aligning with the understanding that optimal electrochemical performance requires a reasonable loading level; (iv) when binder is used, the rate capability first increases and then decreases with mass loading. This is attributed to the fact that at low mass loading, the electrode is thinner, facilitating easier electrolyte penetration and enhancing the contact efficiency between ions and electrolyte, which favors rate performance; whereas at high mass loading, the electrode becomes thicker, hindering electrolyte infiltration and reducing the accessibility of inner active sites, thereby lowering rate capability; (v) under binder-free conditions, the rate capability exhibits the same trend as with binder usage, further confirming the validity and reliability of the modeling data.

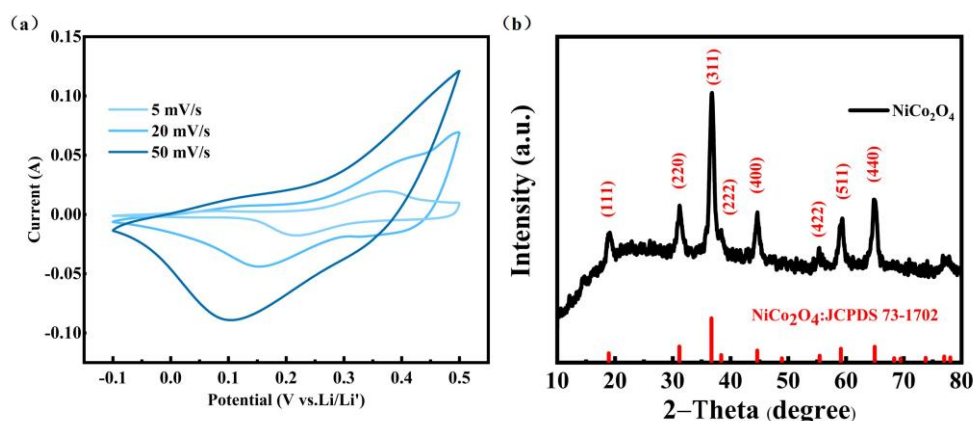


Fig. S6 (a) CV electrochemical testing and (b) XRD diffraction pattern characterization of NiCo₂O₄ samples

Note: As shown in Fig. S5(a), the NiCo₂O₄ sample exhibits distinct pseudocapacitive behavior at moderate to low scan rates (5 mV s⁻¹–20 mV s⁻¹), with evident oxidation and reduction peaks observed at (0.2 V, 0.35 V) and (0.15 V, 0.4 V), respectively. In Fig. S5(b), the XRD pattern of the prepared sample displays characteristic diffraction peak positions and intensities that are consistent with the standard reference card for NiCo₂O₄: JCPDS PDF#73-1702, present a typical Fd $\bar{3}$ m disorderly spinel structure.

Table S3. Comparison of properties of various NiCo₂O₄ electrode materials

Materials	Specific Capacitance	Cyclic Stability	Ref.
NiCoO ₂ /G@NF	1220 F g ⁻¹ at 1 A g ⁻¹	80% after 5000 cycles at 10 A g ⁻¹	1
N-NiCoO	945.79 F g ⁻¹ at 1 A g ⁻¹	93.3% after 5000 cycles at 5 A g ⁻¹	2
Porous NiCoO nanowire	1191.4 F g ⁻¹ at 1 A g ⁻¹	86% after 3000 cycles at 10 A g ⁻¹	3
NiCoO-EGO	530 F g ⁻¹ at 1 A g ⁻¹	82% after 10000 cycles at 5 A g ⁻¹	4
Nanostructured spinel NiCoO urchin-like nanostructure	730 F g ⁻¹ at 1 A g ⁻¹	92.23% after 5000 cycles at 50 mV ⁻¹	5
NiCoO urchin-like NCO-D	423.9 F g ⁻¹ at 1.5 A g ⁻¹	94% after 1500 cycles at 7.5 A g ⁻¹	6
E-NCO nanoarrays	580 F g ⁻¹ at 1 A g ⁻¹	≈97% after 1000 cycles at 15 A g ⁻¹	7
Chestnut shell spherical NiCo ₂ O ₄	1360.7 F g ⁻¹ at 1 A g ⁻¹	87.2% after 10000 cycles at 40 A g ⁻¹	8
	1538 F g ⁻¹ at 1 A g ⁻¹	90.3% after 10000 cycles at 10 A g ⁻¹	This work

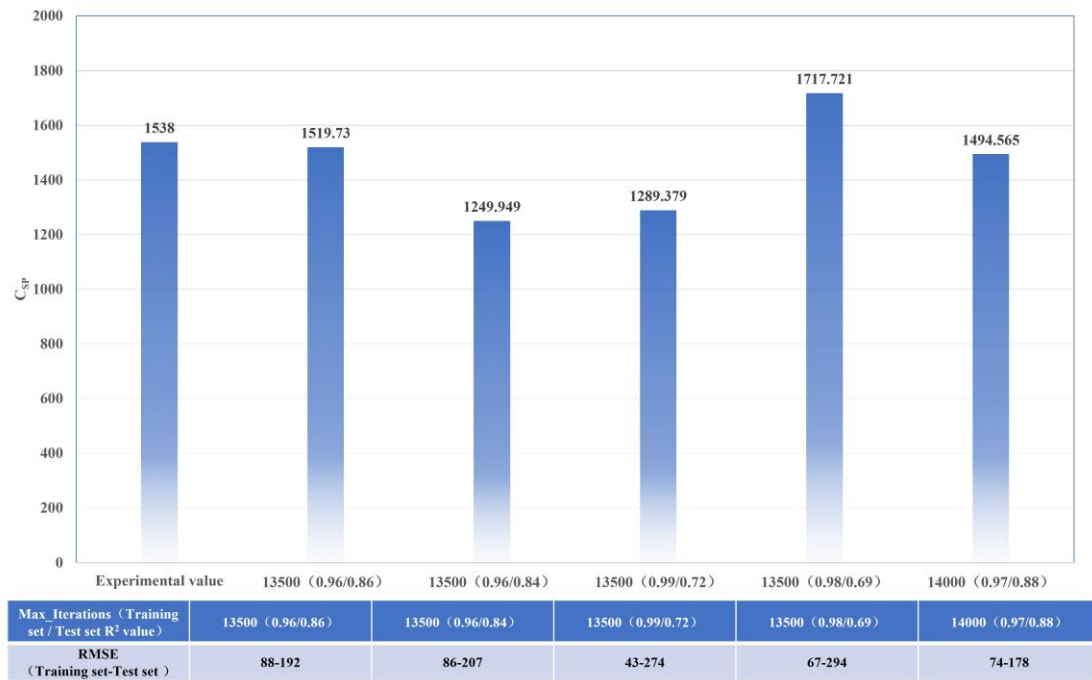


Fig. S7. Histogram of prediction results under different performances of MLP model

Note: As shown in the Fig.S7, different performance outcomes and corresponding specific capacitance predictions of NiCo₂O₄ samples were obtained from MLP models trained with either fixed or varied *Max_Iterations* hyperparameters. Specifically, for the three groups with the same hyperparameter setting (*Max_Iterations* = 13500), where the R² values for the training and test sets were (0.96/0.84), (0.99/0.72), and (0.98/0.69), the significant overfitting (R² difference > 0.1) led to noticeable deviations between the predicted and experimental values. In contrast, for the two groups with R² values of (0.96/0.86) and (0.97/0.88), corresponding to *Max_Iterations* values of 13500 and 14000 respectively, the prediction errors were smaller and the predicted values were closer to the actual value (1538), which can be attributed to the lower degree of overfitting (R² difference < 0.1) and thus a more reasonable generalization capability. These results indicate that reduced overfitting tends to enhance model prediction performance. Overall, all prediction errors fell within the RMSE ranges provided by the training and test sets. Based on this analysis, the following conclusions can be drawn: (i) the degree of overfitting between the training and test set R² values serves as a criterion to assess the reliability of the prediction results, and the same principle applies to cases of underfitting; (ii) the prediction error of the model can be estimated using the

RMSE ranges from the training and test sets, in combination with the level of overfitting or underfitting, to comprehensively evaluate the credibility of the error range. In other words, without altering or fine-tuning key hyperparameters of the model (e.g., adjusting *Max_Iterations* from 13,500 to 14,000 as shown in the figure), a smaller degree of R^2 overfitting between the training and testing sets, along with a narrower RMSE gap, indicates higher reliability and accuracy of the model in predicting new samples.

Supplementary Note 1: (i) Precision: The proportion of samples predicted to belong to a given class that actually belong to that class. For A, a precision of 0.80 means that 80% of the samples predicted as A are actually A. (ii) Recall: The proportion of actual samples from a given class that are correctly predicted by the model. For A, a recall of 0.75 means that 75% of the actual A samples are correctly predicted. (iii) F1-score: The harmonic mean of precision and recall for a given class, providing a comprehensive evaluation of model performance. Support: The actual sample count for each class; accuracy: Similar to the R^2 value. (iv) Macro avg: The average precision, recall, and F1-score across all classes, without considering the sample count of each class. (v) Weighted avg: The weighted average of precision, recall, and F1-score across all classes, considering the sample count of each class.

Supplementary Note 2: Univariate sensitivity analysis was conducted by repeatedly training the model on datasets in which the values of a single feature column were manually shuffled, and comparing the resulting R^2 scores with those obtained from the original unshuffled dataset using the same model configuration. A larger difference in R^2 score indicates a greater impact of the corresponding feature on model performance. In this study, the MLP model was used to evaluate the relative importance of the most influential feature-“With/without adhesive”-by shuffling its values in descending order. Under identical hyperparameter settings, the model’s R^2 score changed from 0.984 to 0.999. Similarly, when the moderately important feature “mass loading” and the least important feature “current density” were shuffled using the same method, the R^2 scores changed to 0.997 and 0.991, respectively. Thus, the largest R^2 variation occurred when the binder feature was shuffled. Assuming model randomness is negligible, this analysis further confirms that the “binder” feature has the highest influence on the MLP model’s prediction performance compared to the other two attributes, which is consistent with the ranking shown in Figure 5(b). Similar tests were also performed on the RF model, and the results aligned with those shown in Figure 5(a); therefore, they are not discussed in detail here.

- 1 J. Wu, Y. Sun, X. Yang, G. Long, Y. Zong, X. Li and X. Zheng, *Ceramics International*, 2018, **44**, 4875–4882.
- 2 Y. Li, Y. Shan and H. Pang, *Chinese Chemical Letters*, 2020, **31**, 2280–2286.
- 3 T. Nawaz, Y. Wen, M. Ahmad, K. Hussain, A. Ali, Q. Ullah, S. Ali Khan, S. Mohammad Wabaidur, P. Rosaiah, S. Ullah, W. U. Arifeen, T. J. Ko and I. Hussain, *Journal of Industrial and Engineering Chemistry*, 2024, **133**, 498–504.
- 4 S. Pappu, S. Anandan, T. N. Rao, S. K. Martha and S. V. Bulusu, *Journal of Energy Storage*, 2022, **50**, 104598.
- 5 R. Prakshale, S. Bangale, M. Kamble and S. Sonawale, *Micro and Nanostructures*, 2024, **189**, 207820.
- 6 Y. Y. Kannangara, S. Karunarathne, W. P. S. L. Wijesinghe, C. Sandaruwan, C. R. Ratwani, A. R. Kamali and A. M. Abdelkader, *Journal of Energy Storage*, 2024, **84**, 110717.
- 7 P. Chomkhuntod, P. Phonsuksawang, A. Waehayee, K. Ngamchuea, P. Iamprasertkun, S. Maensiri, A. Ruangvittayanon and T. Siritanon, *Journal of Energy Storage*, 2024, **86**, 111303.
- 8 P. Qi, H. Wang, M. Chen, Y. Lu, H. Wu, K. Hao, G. Liu and Y. Tang, *Applied Surface Science*, 2024, **665**, 160267.