Chinese Chemical Letters 36 (2025) 109870

Contents lists available at ScienceDirect

Chinese Chemical Letters

journal homepage: www.elsevier.com/locate/cclet

Identifying key determinants of discharge capacity in ternary cathode materials of lithium-ion batteries

Xiangyue Li[a,1], Dexin Zhu[b,1], Kunmin Pan[a,*], Xiaoye Zhou[c], Jiaming Zhu[d,e], Yingxue Wang[f,*], Yongpeng Ren[a], Hong-Hui Wu[b,d,*]

[a] School of Materials Science and Engineering, Henan University of Science and Technology, Luoyang 471003, China
[b] Beijing Advanced Innovation Center for Materials Genome Engineering, Institute for Carbon Neutrality, University of Science and Technology Beijing, Beijing 100083, China
[c] Department of Materials Science and Engineering, Shenzhen MSU-BIT University, Shenzhen 518172, China
[d] Institute of Materials Intelligent Technology, Liaoning Academy of Materials, Shenyang 110004, China
[e] School of Civil Engineering, Shandong University, Ji'nan 250061, China
[f] National Engineering Laboratory for Risk Perception and Prevention, Beijing 100041, China

ARTICLE INFO

Article history:
Received 15 March 2024
Accepted 8 April 2024
Available online 9 April 2024

Keywords:
LNCM ternary cathode material
Discharge capacity
Feature engineering
Machine learning
SHAP analysis

ABSTRACT

Although lithium-ion batteries (LIBs) currently dominate a wide spectrum of energy storage applications, they face challenges such as fast cycle life decay and poor stability that hinder their further application. To address these limitations, element doping has emerged as a prevalent strategy to enhance the discharge capacity and extend the durability of Li-Ni-Co-Mn (LNCM) ternary compounds. This study utilized a machine learning-driven feature screening method to effectively pinpoint four key features crucially impacting the initial discharge capacity (*IC*) of Li-Ni-Co-Mn (LNCM) ternary cathode materials. These features were also proved highly predictive for the 50th cycle discharge capacity (*EC*). Additionally, the application of SHAP value analysis yielded an in-depth understanding of the interplay between these features and discharge performance. This insight offers valuable direction for future advancements in the development of LNCM cathode materials, effectively promoting this field toward greater efficiency and sustainability.

Lithium-ion batteries (LIBs) dominate the energy storage market due to their high energy density, long cycle life, and environmental compatibility [1–6]. Among cathode materials, Li-Ni-Co-Mn (LNCM) ternary materials, comprising Ni, Co, and Mn, are distinguished for their high energy density, durability, and environmental sustainability [7–10]. The initial discharge capacity (*IC*) of LNCM materials is a critical performance metric, affecting cycle life and indicative of overall performance [11,12]. Moreover, the discharge capacity at the 50th cycle (*EC*) stands out as a critical parameter indicative of a battery's cycle endurance. Hence, the pursuit and discovery of strategies to enhance the *IC* and *EC* performances have become prominent subjects of inquiry in the realm of battery.

Discharge capacity, a pivotal metric for lithium-ion ternary battery performance, mirrors the available electric charge [13,14]. Doping is a prevalent approach to augment discharge capacity by optimizing LNCM material properties [15–18], and optimizing crystal-electronic structures to enhance capacity and cycle stability [19–21]. It increases specific surface area, boosting active material utilization and, consequently, battery discharge capacity [21,22]. Doping also optimizes the electronic structure, quickening lithium-ion diffusion, decreasing internal resistance, and uplifting overall battery efficacy [19]. The recent synthesis of AI and materials science has sparked keen interest due to their synergistic significance [23,24]. Machine learning, a key component of AI, offers pathways for optimizing and developing material properties [25,26]. Neural networks, for example, fulfill the need for high-efficiency solid-state electrolytes (SSE) materials [27] and accurately forecast polymer electrolyte ion conductivity [28]. Similarly, machine learning can predict the discharge capacity of LNCM ternary materials, guiding improvements in LIBs' performance.

This study aims to identify key material descriptors influencing *IC* and *EC* of LNCM ternary materials using machine learning. Data on *IC* and *EC* for LNCM ternary materials, doped with various single elements, were collected from academic publications. Subsequently, 229 material descriptors were constructed using do-

* Corresponding authors.
E-mail addresses: pankunming2008@haust.edu.cn (K. Pan), wangyingxue@cetc.com.cn (Y. Wang), wuhonghui@ustb.edu.cn (H.-H. Wu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cclet.2024.109870
1001-8417/

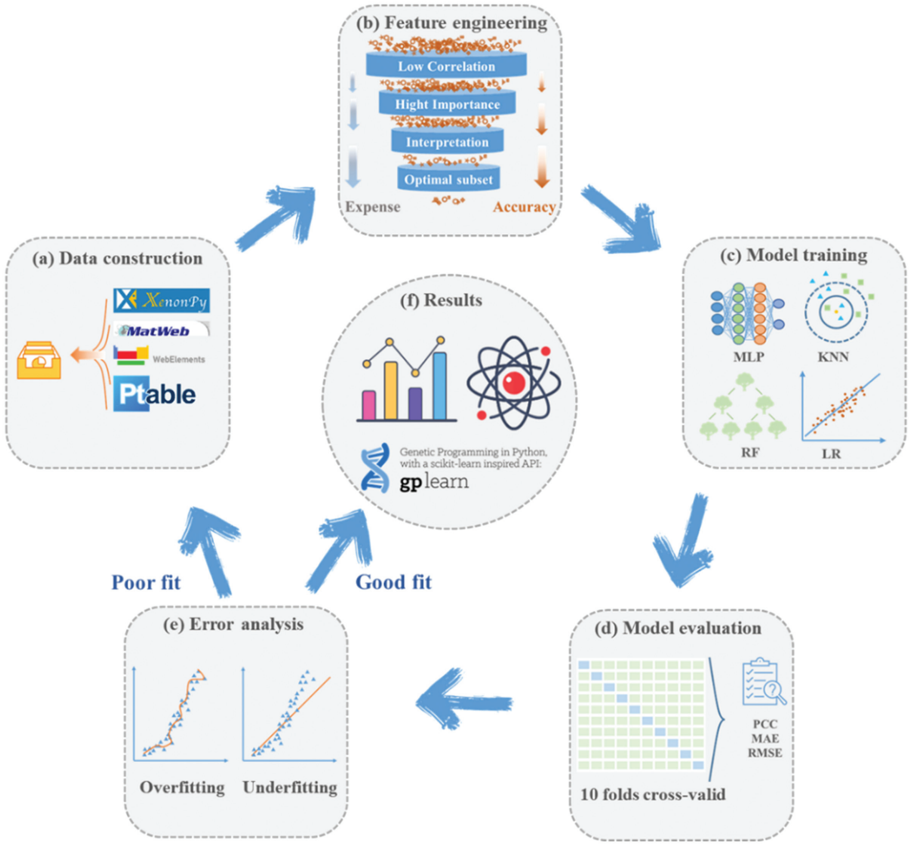

Fig. 1. A workflow of screening the key factors in the discharge capacity of LNCM.

main knowledge and XenonPy [29]. An effective feature screening method was developed, integrating the Pearson correlation coefficient (*PCC*), importance coefficient, optimal subset method [30], and SHapley Additive exPlanations (SHAP) plots. This method pinpointed key features with high prediction accuracy for both *IC* and *EC* of LNCM materials [31], the workflow was shown in Fig. 1. SHAP plots elucidated potential relationships between these features and the target performance.

This work integrated dual feature sets originating from XenonPy calculations and domain-specific knowledge, summarized in Tables S1 and S2 (Supporting information). This study compiled a dataset of 237 single-element doped LNCM materials, as shown in Tables S3 and S4 (Supporting information), and normalized the feature set by Eq. S8 (Supporting information) [32–34]. After screening for feature independence, importance, and interpretation (Figs. S1 and S2 in Supporting information), we distilled down to 13 vital features. To derive the optimal subset, the optimal subset method is adopted for screening. As shown in Fig. 2a, observing a declining trend in *PCC* beyond four features, we concluded that the optimal subset comprises four attributes: *var:ground_state_magnetic_moment* (*VMM*), *MTE*, *Vmax*, and *CD*.

According to the "no free lunch" law, it is impractical to use a single algorithmic model to solve all machine learning problems [30,35-37]. Therefore, this study selects 5 typical models by ten-fold cross-validation [38–40], as shown in Fig. 2b. A series of indexes, *PCC*, *MAE*, and *RMSE*, were employed to gauge model performance [41–44]. Inspection of the results in Fig. 2b demonstrates that the XGB model consistently exhibits optimal performance across all three indexes, with *PCC* at 0.9034 and *MAE* and *RMSE* measured at 13.8033 and 18.9776 mAh/g, respectively. This indicates that through its exploration of the underlying relationship between the optimal set of features and the target perfor-

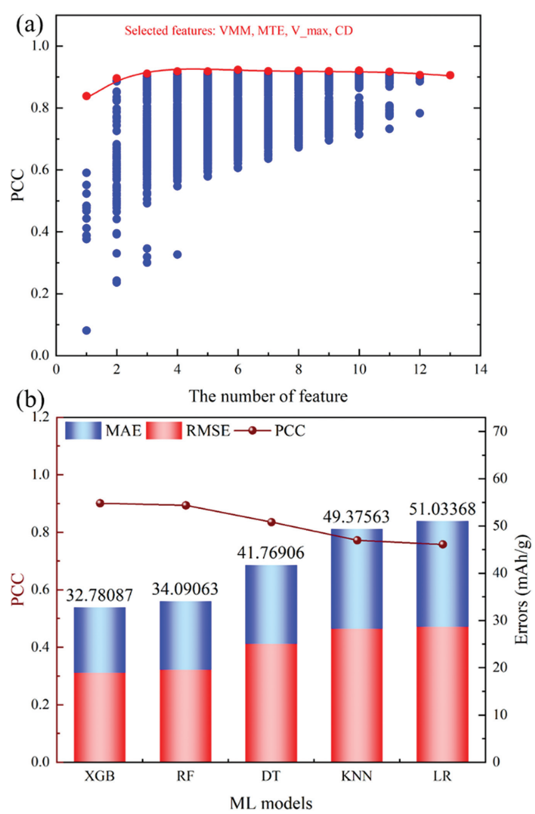

Fig. 2. Selecting the optimal feature set and regression model. (a) *PCC* value was obtained by the optimal subset selection method. (b) The performance of 5 regression models with *PCC*, *MAE*, and *RMSE* values.

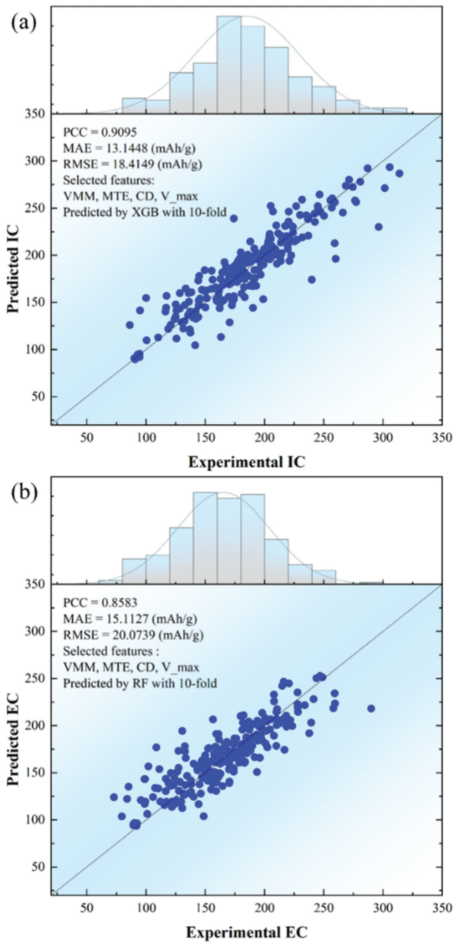

Fig. 3. The scatterplot of the *IC* optimal model and *EC* optimal model: The scatterplot of (a) *IC* model and (b) *EC* model based on the same feature set.

mance, the XGB model achieves accurate and low-bias predictions. Consequently, the XGB model is selected as the preferred output model for predicting *IC*.

After feature engineering screening, four features demonstrating high predictive accuracy for target performance were identified, and Fig. 3a visually illustrates their predictive strength. By inputting the optimal subset of features into various regression models to select the model demonstrating the best performance, as illustrated in Fig. 2a, the XGB model was chosen as the optimal output model for predicting *IC*. Similarly, following the same rationale, the RF model was selected as the top-performing model for forecasting *EC* (Fig. 3b). After 100 random tests, the target performance yielded $PCC = 0.9012$, $MAE = 13.9376$ mAh/g, and $RMSE = 19.1462$ mAh/g. Beyond the notable predictive capability for *IC*, this screening approach also demonstrates heightened performance in predicting *EC*. Following 100 random tests, the performance indexes were $PCC = 0.8461$, $MAE = 15.2549$ mAh/g, and $RMSE = 20.8239$ mAh/g, highlighting the robustness of the model.

The study further analyzed the impact of specific features on target performance using SHAP plots. Figs. 4a-d presented SHAP scatter plots, where Figs. 4a and c highlighted the positive correlation between *Vmax* and *MTE* with target performance, suggesting improved performance with higher *Vmax* and *MTE* values. Conversely, Figs. 4b and d associate *CD* and *VMM* with target performance, revealing a negative relationship as well.

In summary, this research successfully identified four key features significantly affecting the *IC* and *EC* of LNCM ternary materials. By examining the potential relationships between these features and the target performance using SHAP, we contribute to advancing battery technology and energy storage solutions. Furthermore, this methodology offers a basis for exploring performance prediction and optimization strategies in other battery materials.

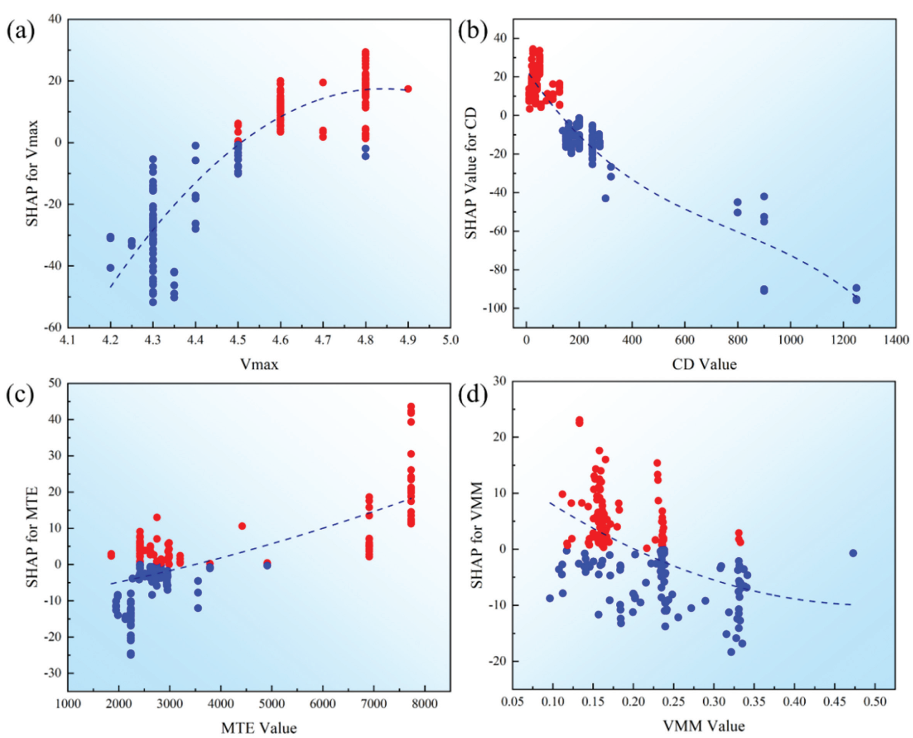

Fig. 4. The relationship between each feature and the *IC* based on the SHAP value diagram: The scatterplot of (a) *Vmax*, (b) *CD*, (c) *MTE*, and (d) *VMM* with their SHAP value, respectively.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xiangyue Li: Resources, Software, Writing – original draft, Writing – review & editing. **Dexin Zhu:** Software, Supervision, Writing – review & editing. **Kunmin Pan:** Funding acquisition, Resources, Supervision, Writing – review & editing. **Xiaoye Zhou:** Investigation, Supervision. **Jiaming Zhu:** Investigation, Validation. **Yingxue Wang:** Data curation, Software, Funding acquisition. **Yongpeng Ren:** Methodology, Visualization. **Hong-Hui Wu:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (Nos. 52122408, 52071023), the Program for Science & Technology Innovation Talents in the University of Henan Province (No. 22HASTIT1006), the Program for Central Plains Talents (No. ZYYCYU202012172), the Ministry of Education, Singapore (No. RG70/20) and the Opening Project of National Joint Engineering Research Center for Abrasion Control and Molding of Metal Materials, Henan University of Science and Technology (No. HKDNM201906).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cclet.2024.109870.

References

[1] J. Hou, M. Yang, D. Wang, et al., Adv. Energy Mater. 10 (2020) 1904152.
[2] N. Zhang, T. Deng, S. Zhang, et al., Adv. Mater. 34 (2022) 2107899.
[3] M. Armand, J.M. Tarascon, Nature 451 (2008) 652–657.
[4] M.R. Palacín, Chem. Soc. Rev. 38 (2009) 2565–2575.
[5] B. Scrosati, J. Hassoun, Y.K. Sun, Energy Environ. Sci. 4 (2011) 3287–3295.
[6] G. Zhu, K. Wen, W. Lv, et al., J. Power Sources 300 (2015) 29–40.
[7] Y. Liu, M. Liu, Int. J. Hydrogen Energy 42 (2017) 18189–18195.
[8] Y.R. Kim, Y.W. Yoo, D.Y. Hwang, et al., Solid State Ionics 389 (2023) 116108.
[9] K. Xu, Energy Environ. Mater. 2 (2019) 229–233.
[10] H. Jeong, S.J. Lim, S. Chakravarthy, et al., J. Power Sources 451 (2020) 227764.
[11] C. Wang, W. Zhang, T. He, et al., Int. J. Electrochem. 26 (2020) 12.
[12] G.G. Eshetu, H. Zhang, X. Judez, et al., Nat. Commun. 12 (2021) 5459.
[13] L. Liu, M. Li, L. Chu, et al., Prog. Mater Sci. 111 (2020) 100655.
[14] J. Zhang, X. Li, G. Zhang, et al., Int. J. Energy Res. 44 (2020) 3134–3147.
[15] A.S. Maiti, A.H. Sclar, D.R. A, et al., Energy Storage Mater. 45 (2022) 74–91.
[16] W.W. Li, X.J. Zhang, J.J. Si, et al., Rare Met. 40 (2020) 1–8.
[17] S.B. Kim, S.Y. Ahn, J.H. Kim, et al., Electrochem. Commun. 146 (2023) 107426.
[18] H. Kim, M.G. Kim, H.Y. Jeong, et al., Nano Lett. 15 (2015) 2111–2119.
[19] L. Bao, Z. Yang, L. Chen, et al., ChemSusChem 12 (2019) 2294–2301.
[20] L. Sun, X. Yi, X. Ren, et al., J. Electrochem. Soc. 163 (2016) A766–A772.
[21] G. Xue, Q. Xue, J. Li, et al., Solid State Ionics 293 (2016) 7–12.
[22] G. Wang, T. Fearn, T. Wang, et al., ACS Cent. Sci. 7 (2021) 1551–1560.
[23] R. Ramprasad, R. Batra, G. Pilania, et al., npj Comput. Mater. 3 (2017) 54.
[24] J. Liu, C.J. Tian, W. Chang, J. Chin. Ceramic Soc. 51 (2023) 6.
[25] W. Li, T. Yang, C. Liu, et al., Adv. Sci. 9 (2022) e2105550.
[26] D. Yue, Y. Feng, X.X. Liu, et al., Adv. Sci. 9 (2022) e2105773.
[27] Z. Lu, P. Adeli, C.H. Yim, et al., ACS Appl. Energy Mater. 5 (2022) 8042–8048.
[28] G. Bradford, J. Lopez, J. Ruza, et al., ACS Cent. Sci. 9 (2023) 206–216.
[29] Y.R. Yoshida, Mol. Inf. 39 (2020) 1900107.
[30] X. Li, D. Zhu, K. Pan, et al., Int. J. Refract. Met. Hard Mater. 117 (2023) 106386.
[31] S. Wu, G. Lambard, C. Liu, et al., Mol. Inf. 39 (2020) 1900107.
[32] Y. Chen, S. Wang, J. Xiong, et al., J. Mater. Sci. Technol. 132 (2023) 213–222.
[33] B. Kumarsingh, K. Verma, A.S. Thoke, Int. J. Comput. Appl. Technol. 116 (2015) 11–15.
[34] J. Li, N. Wu, J. Zhang, et al., Nano-Micro Lett. 15 (2023) 227.
[35] K. Sharma, M. Cerezo, Z. Holmes, et al., Phys. Rev. Lett. 128 (2020) 070501.
[36] C. Shang, C. Wang, H. Wu, et al., Sci. China Technol. Sci. 66 (2023) 2069–2079.
[37] F. Wang, H.H. Wu, L. Dong, et al., J. Mater. Sci. Technol. 165 (2023) 49–65.
[38] T.T. Wong, P.Y. Yeh, IEEE Trans. Knowl. Data Eng. 32 (2020) 1586–1594.
[39] R. Malhotra, S. Meena, 2021 S International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India (2021) 431–438 https://ieeexplore.ieee.org/document/9533030.
[40] R. Kohavi, IEEE Trans. Knowl. Data Eng. 32 (2019) 1586–1594.
[41] T.O. Hodson, Geosci. Model Dev. 15 (2022) 5481–5487.
[42] D. Chicco, M.J. Warrens, G. Jurman, Peer J. Comput. Sci. 7 (2021) e623.
[43] T. Chai, R.R. Draxler, Geosci. Model Dev. 7 (2014) 1247–1250.
[44] J. Adler, I. Parmryd, Cytometry Part A 77 (2010) 733–742.