Proceedings Paper:

Liu, K, Li, K orcid.org/0000-0001-6657-0522 and Chen, T (2022) Interpretable Sensitivity Analysis and Electrode Porosity Classification for Li-ion Battery Smart Manufacturing. In: Proceedings: 2021 IEEE Sustainable Power and Energy Conference (iSPEC). 2021 IEEE Sustainable Power and Energy Conference (iSPEC), 23-25 Dec 2021, Nanjing, China. IEEE , pp. 3653-3658. ISBN 978-1-6654-1439-5

https://doi.org/10.1109/iSPEC53008.2021.9735647

Interpretable Sensitivity Analysis and Electrode Porosity Classification for Li-ion Battery Smart Manufacturing

Kailong Liu
WMG
University of Warwick
Coventry, UK
kailong.liu@warwick.ac.uk;
kliu02@qub.ac.uk

Kang Li
School of Electronic and Electric Engineering
University of Leeds
Leeds, UK
k.li1@leeds.ac.uk

Tao Chen
School of Electrical Engineering
Southeast University
Nanjing, Jiangsu, China
taoc@seu.edu.cn

***Abstract*—Lithium-ion batteries have become one of the most promising sources for accelerating the development of sustainable energy, where effective cell manufacturing plays a direct role in determining battery qualities. Due to the highly complicated process and strongly coupled interdependencies of battery manufacturing, a data-driven approach that can evaluate the sensitivity of manufacturing parameters and provide the effective classification is urgently required. This paper proposes a boosting tree-based ensemble machine learning framework to analyze and predict how the battery electrode porosity varies with respect to the key parameters of both mixing and coating stages for the first time. Three boosting models including the AdaBoost, LPBoost, and TotalBoost are established and compared. Illustrative results demonstrate that the proposed ensemble machine learning framework is able to not only give effective quantification of both importance and correlations of parameters of interest but also provide satisfactory early-stage prediction. These kinds of information could benefit the monitoring and analysis of battery manufacturing chain, further help to produce high quality batteries for wider sustainable energy applications.**

Keywords—Sustainable energy, Li-ion battery, Battery manufacturing, Data analysis, Ensemble machine learning

I. Introduction

Environmental challenges including the global warming and the reduced sources of fossil fuels have increased the requirements for sustainable energy and transportation applications. Lithium-ion (Li-ion) batteries become one of the most promising energy storage sources not only for mobility electrification but many other sustainable applications, which leads to the significantly increased demand for them in recent years. However, the performance of Li-ion batteries such as capacity, service life, energy or power density, and thermal conductivity are directly and highly affected by their related production process. To optimize battery quality and consequently production costs, it is vital to understand the correlations between various production parameters and battery quality variables [1].

Unfortunately, battery production is complex with many inter-mediate stages and numerous strong-coupled process parameters. Due to the multiple disciplinary information including electrical, mechanical, and chemical operations are involved in the whole battery manufacturing chain, the analyses of importance and correlations among various intermediate parameters and battery variables still often rely on the experiment experiences, expert advice, trial and error method [2]. These methods lead to huge laborious and time consumptions, slow battery product development, inaccurate quality control and difficulty in distinguishing the products with different quality levels at earlier stages. In this context, designing an effective data analysis strategy to conduct the reliable sensitivity analyses including the importance as well as correlation quantifications of battery manufacturing parameters is urgently needed.

With the rapid development of machine learning and artificial intelligence techniques, data-driven solutions have been popular tools for battery managements. A good deal of works are designed for battery internal states estimation [3], lifetime prognostics [4][5], faults diagnostics [6], cells equalization [7], charging control [8][9], market arbitrage [10], electric vehicle sizing [11] as well as energy management [12]. Overall, after designing suitable data-driven models, more effective battery management performance could be obtained. However, these works mainly focus on the final manufactured battery performance with the fact that relatively little has been done on technical improvement of their related manufacturing process [13]. As battery manufacturing would play a more direct and significant role in determining battery performance [14], which is also worth being well monitored and analyzed.

In comparison with the field of battery management, fewer existing works have been done so far on designing data-driven solutions to benefit battery manufacturing [15]. For example, according to the cross-industry standard process (CRISP), the linear as well as neural network-based data-driven models are proposed in [16] to predict battery properties and identify the dependency of battery manufacturing chain. Turetskyy et al. [17] adopted the tree-based techniques to analyze variable importance and predict the maximum capacity of manufactured battery. After designing statistical data-driven solution to analyze battery manufacturing fluctuations, their effects on the capacity of manufactured battery are evaluated in [18]. For the aforementioned relevant research works, reasonable data analyses of battery manufacturing can be achieved, but most researches simply adopt the conventional methodologies to predict the properties of battery production. Besides, little has been done so far by designing data-driven solutions, especially using ensemble framework, to in-depth analyze the effects of manufacturing parameters within the key stages such as mixing and coating. It should be known that battery mixing and coating are crucial to determine the qualities of manufactured battery electrode [19]. In order to achieve battery smarter manufacturing and optimize mixing as well as coating stages, it is vital to carry out the efficient sensitivity analyses of the battery electrode properties with respect to its mixing and coating specifications.

Given the aforementioned considerations, a boosting-tree based ensemble machine learning framework is proposed in this paper to conduct sensitivity analysis of key manufacturing parameters and predict battery electrode porosity qualities. The focus of this study is on the effects of mixing and coating key parameters on final battery electrode products. Some key objectives could be summarized as: 1) to quantify importance as well as correlations of four key manufacturing parameters from both mixing and coating via a well-designed ensemble machine learning framework; 2) to classify and predict battery electrode porosity at early production stages via effective data-driven models; 3) to evaluate and compare performance of typical AdaBoost model and two other improved boosting-based tree models (LPBoost and TotalBoost) for battery electrode classification case. All these efforts could help manufacturer to produce more efficient and high performance batteries, further benefitting battery smarter manufacturing for wider sustainable energy applications.

II. Li-ion Battery Production

Li-ion battery production is a complex process that generally contains three parts including battery electrode manufacturing, battery assembly, and battery formation. As illustrated in Fig. 1, electrode production starts from mixing stage, in which the prepared material components will be mixed within a soft blender to generate slurries. After that, the slurries will be coated on the surface of collector foil which is made of copper/aluminium in general. Then the coating product will be dried by the oven. After calendering and cutting steps, battery anode and cathode electrodes will be generated. The whole electrode production process involves electrical, mechanical, as well as chemical operations. All stages within it need to be conducted through using the specific equipment. As two key stages in the battery electrode production, mixing and coating are complicated with many strong-coupled parameters, and these manufacturing parameters would highly affect some battery electrode quality indicators such as porosity. Therefore, a reliable solution that can analyze the sensitivity of interested parameters and predict the manufactured electrode porosity is necessary for improving battery manufacturing.

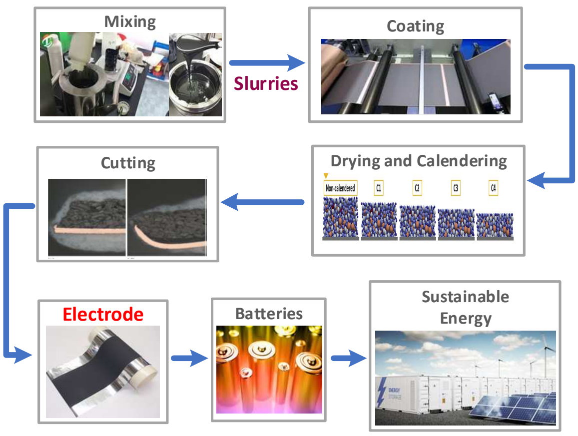

Fig. 1. Typical processes of battery manufacturing particular for electrode production.

To achieve this, a boosting-based ensemble machine learning framework is designed to classify battery electrode porosity and analyze both importance as well as correlations of some key manufacturing parameters in this study. Specifically, three mixing parameters including the mass content (MC) of active material, solid-to-liquid ratio (StLR), and viscosity (Vis), as well as one coating parameters: comma gap (CG) are investigated. Here StLR means the mass ratio between slurry solid and mass. Vis affects the shear rate of coating stage. CG is the gap between coating comma and coating roll. Without the loss of generality, battery manufacturing experimental dataset from Franco Laboratoire-de-Reactivite-et-Chimie-des-Solides is adopted. Detailed experimental information and data explanation are referred to [20] for readers of interest. To fully investigate the classification performance of designed ensemble machine learning models, battery electrode porosity with the unit of % is labelled with five classes (very low, low, medium, high and very high). Specifically, very low refers to the range of (0, 47.5], low refers to the range of (47.5, 50], medium reflects the range of (50, 52.5], while high and very high refer to the ranges of (52.5, 55], and (55, 70], respectively. After predefining the class labels of electrode porosity, the boosting-based ensemble machine learning framework for both porosity quality classification and parameter sensitivity analyses can be designed.

III. Technologies

In this section, the fundamental of AdaBoost is first given, followed by the descriptions of another two improved boosting techniques including LPBoost and TotalBoost. Then the tree-based ensemble machine learning framework to analyze battery electrode production is designed. To evaluate their performance, some performance indicators are also given.

A. AdaBoost, LPBoost and TotalBoost

Boosting is one of most effective and widely-used solutions to derive ensemble machine learning framework. The key idea of boosting is to sequentially train various weak hypothesis, while the training dataset's distribution would be also changed dynamically based on the performance of previous trained weak learner.

Adaptive boosting (AdaBoost) is a typical and effective boosting solution for real classification applications [21]. Let training dataset *TD* includes *K* observations as: $TD=\{(x_1, y_1), (x_2, y_2), \ldots, (x_K, y_K)\}$. Here x_k ($k=1:K$) reflects the input vector of the interested manufacturing parameters, y_k ($k=1:K$) stands for the preset classification labels with a total number of *C*, $L(x)$ means a weak learner that would output a classification result related to x, then the detailed process to establish AdaBoost-based ensemble machine learning model for classification is summarized in Workflow 1.

Workflow 1: Process to establish AdaBoost-based ensemble classification model

1. Initialize the training observations' weights as:

$$w_k = 1/K, k = 1, 2, ..., K.$$

2. Suppose J is the number of all weak learners $L^{(j)}(x)$. For $j=1$ to J:

   a) Fit $L^{(j)}(x)$ to *TD* based on the weight w_k.

b) Compute the error $er^{(j)}$ as:

$$er^{(j)} = \sum_{k=1}^{K} w_k \cdot I(L^{(j)}(x_k) \neq y_k) \Big/ \sum_{k=1}^{K} w_k$$

where *I(.)* means a zero-one judgement with a rule as:

$$I(L^{(j)}(x_k) \neq y_k) = 1$$

c). Compute the update factor $f^{(j)}$ of weights as:

$$f^{(j)} = \log[(1 - er^{(j)}) / er^{(j)}] + \log(K-1)$$

d). For $k=1,2\ldots,K$, update weight w_k as:

$$w_k \leftarrow w_k \cdot \exp\left[f^{(j)} \cdot I(L^{(j)}(x_k) \neq y_k) \right]$$

e). Renormalize w_k.

3. Output the predicted class result $\tilde{y}(x)$ as:

$$\tilde{y}(x) = \arg\max_{c} \sum_{j=1}^{J} f^{(j)} \cdot I(L^{(j)}(x_k) = y_c)$$

where y_c is the preset classification label ($c=1:C$). *I(.)* here represents another zero-one judgement with a rule of $I(L^{(j)}(x_k) = y_c) = 1$, $\arg\max_{c}$ would output the class that has the largest counted number from results of all trained weak learners.

Apart from AdaBoost, another two improved boosting-based solutions including the LPBoost and TotalBoost are also utilized in this study. Specifically, LPBoost uses the weighted linear combination of learners, so that a weak learner can be added in each iteration with the adjustment of previous weak learners' weights [22]. TotalBoost realizes the classification by maximizing the minimal margin [23]. More detailed information of these two boosting solutions can be found in [24]. It should be known that both LPBoost and TotalBoost have the same general workflow as AdaBoost, but these two improved solutions are self-terminating and produce ensembles with small weights [24].

B. Framework for analyzing battery electrode production

In this study, to effectively analyze the sensitivity of mixing and coating parameters of interest and well classify the qualities of manufactured battery electrode porosity, a novel boosting-based ensemble machine learning framework with a model structure shown in Fig. 2 is proposed. To be specific, three mixing parameters (MC, StLR, and Vis) and one coating parameter (CG) of interest are utilized as the inputs to the model, while the relevant manufactured battery electrode porosity is used as the output. The detailed framework through using the boosting tree-based technique to carry out the sensitivity analysis and classify the qualities of manufactured electrode porosity can be summarized with four key parts as follows:

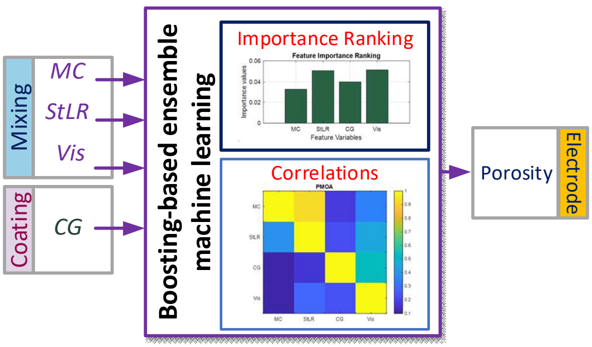

Fig. 2. Model structure of boosting-based ensemble machine learning to classify battery electrode porosity and carry out sensitivity analysis.

Part 1: Data preprocess: In this part, the raw battery electrode manufacturing data will be preprocessed to remove its outliers and set the classification labels. Following the predefined rules in Section II, five classification labels including very low, low, medium, high, and very high are set to reflect the qualities of battery electrode porosity.

Part 2: Ensemble machine learning model construction: to establish effective boosting-based model, the hyper-parameters of AdaBoost, LPBoost, and TotalBoost methods need to be determined. It should be known that for all these three methods, decision tree is usually adopted as their weak learner. Two main hyper-parameters require to be preset: the number of ensembled decision tree (*N*) and their learning rates (*r*). In theory, large *N* leads to the improved accuracy of classification, but too many trees will also cause the overfit issue and increase the computational effort. In order to determine a suitable *N*, an iteration way through comparing learner weights via various numbers of utilized weak learners is adopted. For learning rate, it would reflect a decay rate of each learner's weight, further affecting the performance of each decision tree. As suggested by [18], *r* could be set as 0.1 for the general classification applications. After setting these two hyper-parameters, all boosting-based models can be well trained with the process as illustrated in Workflow 1.

Part 3: Importance and correlation analyses: to quantify the importance of interested manufacturing parameters, Gini index which represents the impurity change due to the splits of each parameter is utilized. In the tree-based classification, impurity could stand for how well a potential split is for decision tree's nodes. The larger Gini index value a manufacturing parameter can obtain, the more important effect this parameter can give. Besides, to carry out the correlation analysis of each manufacturing parameter pair, predictive-measure-of-association (PMOA) value is utilized. Supposing two parameters of interest are P_a and P_b, the PMOA value to reflect their correlation is calculated by:

$$PMOA_{a,b} = \frac{\min(OB_l, OB_r) - 1 + OBl_{a,b} + OBr_{a,b}}{\min(OB_l, OB_r)} \quad (1)$$

where *l* and *r* represent the left child and right child of nodes; OB_l and OB_r mean the observation proportions of $P_a < y$ and $P_a \geq y$, respectively; $OBl_{a,b}$ is the observation proportion under condition of $P_a<y$ and $P_b<z$, while $OBr_{a,b}$ reflects the observation proportion under condition of $P_a \geq y$ and $P_b \geq z$. The solution of obtaining PMOA is to investigate all the potential splits with the best case that is obtained during decision tree's training stage. In this context, PMOA has the ability to quantify the similarity between different rules for splitting observation. After using Eq. (1), the PMOA values of all parameter pairs can be obtained and shown by a 4×4 heat map.

C. Performance indicators

In this study, to quantify and investigate the classification performance, confusion matrix (CM) is adopted as a key performance indicator. Supposing positive stands for an interested class while negative relates to other classes, four basic elements including the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) could be obtained. Afterwards, the precision rate $Pra(C_i)$ as well as recall rate $Rra(C_i)$ of interested class can be obtained as:

$$\begin{cases} Pra(C_i) = TP(C_i)/\left[TP(C_i)+FP(C_i)\right] \\ Rra(C_i) = TP(C_i)/\left[TP(C_i)+FN(C_i)\right] \end{cases} \quad (2)$$

Then a popular performance indicator to illustrate the accuracy of classification results called micro F1 score ($microF1$) can be calculated as:

$$microF1 = \left(TP_{all} + TN_{all}\right)/N_{ob} \quad (3)$$

where TP_{all} and TN_{all} represent all correct classifications, N_{ob} means the total amount of observations.

Besides, the receiver operating characteristic (ROC) curve as well as its area under curve (AUC) value are also adopted to evaluate the performance of classification results in this study. It should be known that ROC curve is a statistical plot to reflect the diagnostic ability of a classification model under the case of varying its discrimination threshold. The AUC could give the degree or measure of separability of the classes.

IV. Results and Discussions

In this section, through using all steps within the parts from subsection III-*B*, the experimental tests by designing proper tree-based ensemble machine learning models can be carried out to quantify the importance and correlations of four input manufacturing parameters of interest, while the battery electrode porosity qualities will be also classified. In this study, all mentioned boosting-based models (AdaBoost, TotalBoost, and LPBoost) are evaluated based on the five fold cross-validation.

A. Boosting-based model trainning

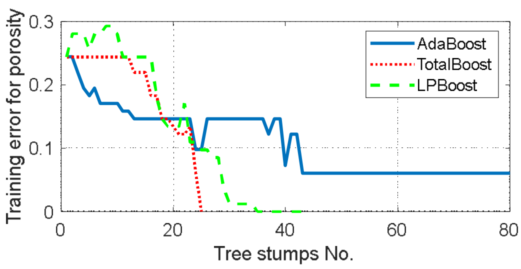

Fig. 3. Training error via the number of tree stumps.

First, to evaluate if the training process of all three boosting-based models would converge and to avoid training overfitting, a training case through using tree stumps with only 1 maximum split as weak learner is carried out. As illustrated in Fig. 3, the training errors of both TotalBoost and LPBoost cases could converge to 0 after using over 20 tree stumps, while the error of AdaBoost case can also converge to around 0.06 after using 40 tree stumps. In light of this, all these three ensemble machine learning models are able to achieve reliable convergence results for classifying the qualities of battery electrode porosity in this study.

Then, to determine the hyper-parameter N of both LPBoost and TotalBoost cases, their learner weights after compacting the corresponding weak learners (decision tree) are shown in Fig. 4. Obviously, both LPBoost as well as TotalBoost models present clear decrease trajectories via the increase of number of ensembled trees. Here the weights of LPBoost become negligible after using 32 decision trees, while the weights of TotalBoost become negligible after using 12 decision trees, indicating that a satisfactory convergence of model training can be achieved. Therefore, the N of LPBoost model and TotalBoost model are set as 32 and 12 for the battery electrode porosity classification, respectively.

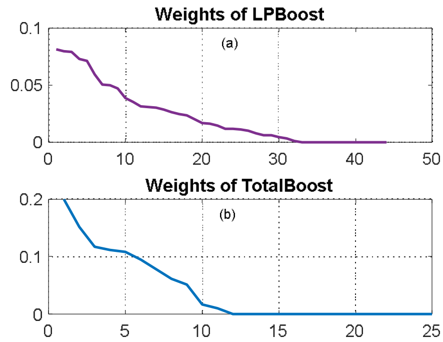

Fig. 4. Weights of LPBoost and TotalBoost via the number of ensembled trees.

B. Parameter importance and correlations quantification

After presetting the hyper-parameters of all boosting-based models, the sensitivity analyses of manufacturing parameters of interest can be carried out. Through calculating the Gini index values of MC, StLR, CG, and Vis, their importance ranking can be quantified, as illustrated in Fig. 5. Quantitatively, the quantified values of StLR and Vis are higher than those of other parameters, indicating that they are the two most important parameters to determine the battery electrode porosity. In contrast, MC shows the minimum Gini index, which means that it presents the lowest effects on the battery electrode porosity classification.

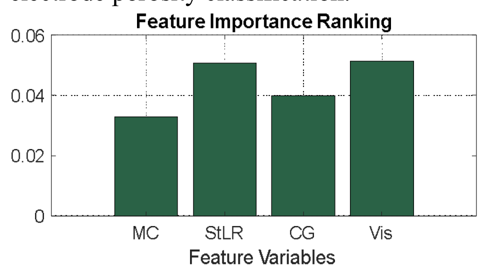

Fig. 5. Importance ranking of battery manufacturing parameters of interest.

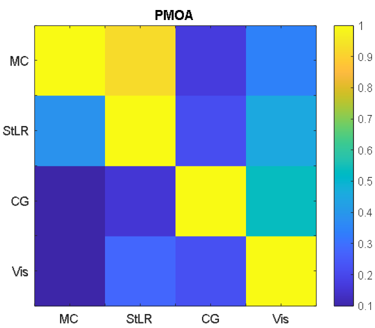

Fig. 6. PMOA-based correlations of interested battery manfuacturing parameter pairs.

To quantify the correlations of each manufacturing parameter pair, the PMOAs of all pairs derived from 4 manufacturing parameters of interest are calculated and illustrated in a heat map matrix, as illustrated in Fig. 6. Obviously, the PMOA of MC and StLR pair gives the largest value around 0.9, indicating that there exists relatively strong correlations between these two manufacturing parameters. This result is expected as the mass ratio between slurry solid and mass actually present strong relations with the active material qualities in theory. For other parameter pairs, their PMOAs are all lower than 0.6, indicating that none strong correlations of these parameter pairs are existed in determining the battery electrode porosity.

C. *Classification performance evaluation*

This subsection details the classification performance of all mentioned boosting-based ensemble machine learning models. After using the well-trained AdaBoost model, LPBoost model, as well as TotalBoost model to classify the qualities of battery electrode porosity, their relevant confusion matrices and *microF1* results are illustrated in Fig. 7 and Table 1, respectively. Quantitatively, AdaBoost presents the worst classification result with 71.2% *microF1*, while TotalBoost model achieves the best result of 74.1% *microF1*, which is 2.1% better than that of LPBoost.

Table 1. Performance indicator of all these three boosting-based ensemble machine learning models

Boosting-based models	microF1 values
AdaBoost	71.2%
LPBoost	72.6%
TotalBoost	74.1%

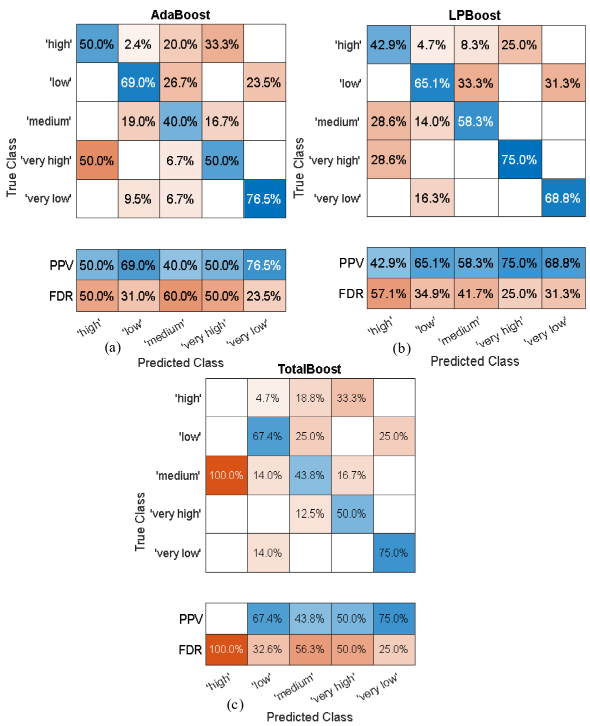

Fig. 7. Confusion matrix of battery electrode porosity classification under different tree-based approaches: a) AdaBoost, b) LPBoost, and c) TotalBoost.

Fig. 8 illustrate the ROC curves of battery electrode porosity classification results through using various tree-based approaches. It can be noticed that the AUC values of all approaches are higher than 0.9. Quantitatively, AUC of TotalBoost case presents the largest one with 0.94, which is 2.2% and 3.3% larger than that from LPBoost case and AdaBoost case, respectively. Therefore, our proposed boosting-based ensemble machine learning framework is able to classify the battery electrode porosity with satisfactory AUC values at the early mixing and coating stages, while TotalBoost-based ensemble machine learning model shows more competent performance among these three adopted boosting techniques.

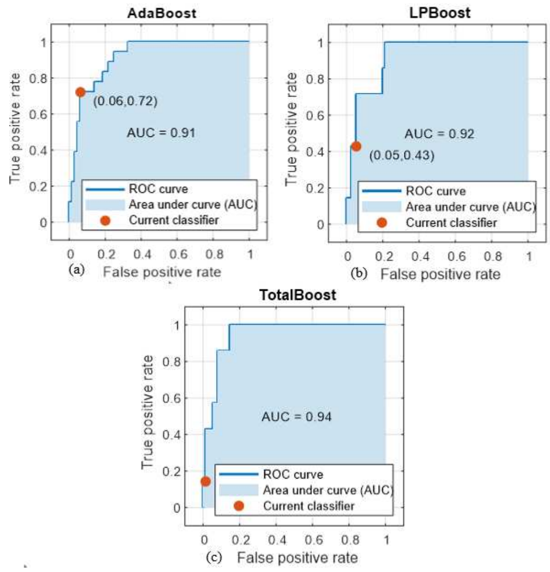

Fig. 8. Receiver operating characteristic curves of battery electrode porosity classification under different tree-based approaches: a) AdaBoost, b) LPBoost, and c) TotalBoost.

Conclusion

Battery manufacturing is crucial for determining the performance of battery and related sustainable energy applications. In this study, an effective boosting tree-based ensemble machine learning framework is designed to carry out effective sensitivity analysis of manufacturing parameters and classify the battery electrode porosity for the first time. Some conclusions can be obtained as follows: 1) StLR and Vis are two important parameters to determine the quality of battery electrode porosity; 2) with the largest PMOA value of 0.9, there is a relatively strong correlation between MC and StLR pair; 3) more other manufacturing parameters should be considered to improve the quality classification of battery electrode porosity. Due to the superiority of sensitivity analysis ability, the designed boosting-based ensemble machine learning framework can be utilized to analyze other manufacturing parameters when the related data are available, further benefitting battery smarter manufacturing and wider applications of battery-based sustainable energy system.

Acknowledgment

This work was supported by the EPSRC under grant EP/R030243/1, and the High Value Manufacturing Catapult project under grant No. 8248 CORE.

REFERENCES

[1] A. Kwade, W. Haselrieder, R. Leithoff, A. Modlinger, F. Dietrich, K. Droeder, Current status and challenges for automotive battery production technologies, Nature Energy 3 (4) (2018) 290-300.

[2] K. Liu, Z. Wei, Z. Yang, and K. Li. Mass load prediction for lithium-ion battery electrode clean production: a machine learning approach. Journal of Cleaner Production 289 (2021): 125159.

[3] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu. State estimation for advanced battery management: Key challenges and future trends. Renewable and Sustainable Energy Reviews 114 (2019): 109334.

[4] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Berecibar, E. N. Maury, J. V. Mierlo, and H. E. Hoster. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. Renewable and Sustainable Energy Reviews 113 (2019): 109254.

[5] K. Liu, Y. Li, X. Hu, M. Lucu, and W. D. Widanage. Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries. IEEE Transactions on Industrial Informatics 16, no. 6 (2019): 3767-3777.

[6] X. Hu, K. Zhang, K. Liu, X. Lin, S. Dey, and S. Onori. Advanced Fault Diagnosis for Lithium-Ion Battery Systems: A Review of Fault Mechanisms, Fault Features, and Diagnosis Procedures. IEEE Industrial Electronics Magazine 14, no. 3 (2020): 65-91.

[7] M. A. Hannan, M. M. Hoque, S. E. Peng, and M. N. Uddin. Lithium-ion battery charge equalization algorithm for electric vehicle applications. IEEE Transactions on Industry Applications 53, no. 3 (2017): 2541-2549.

[8] K. Liu, C. Zou, K. Li, and T. Wik. Charging pattern optimization for lithium-ion batteries with an electrothermal-aging model. IEEE Transactions on Industrial Informatics 14, no. 12 (2018): 5463-5474.

[9] F. Chang, T. Chen, W. Su, and Q. Alsafasfeh. Control of battery charging based on reinforcement learning and long short-term memory networks. Computers & Electrical Engineering 85 (2020): 106670.

[10] J Cao, D. Harrold, Z. Fan, T. Morstyn, D. Healey, and K. Li. Deep Reinforcement Learning-Based Energy Storage Arbitrage With Accurate Lithium-Ion Battery Degradation Model. IEEE Transactions on Smart Grid 11, no. 5 (2020): 4513-4521.

[11] T. Chen, B. Zhang, H. Pourbabak, A. Kavousi-Fard, and W. Su. Optimal routing and charging of an electric vehicle fleet for high-efficiency dynamic transit systems. IEEE Transactions on Smart Grid 9, no. 4 (2016): 3563-3572.

[12] G. Dong, and Z. Chen. Data-driven energy management in a home microgrid based on Bayesian optimal algorithm. IEEE Transactions on Industrial Informatics 15, no. 2 (2018): 869-877.

[13] M. Duquesnoy, I. Boyano, L. Ganborena, P. Cereijo, E. Ayerbe, and A. Franco. Machine Learning-Based on Assessment of the Impact of the Manufacturing Process on Battery Electrode Heterogeneity. Energy and AI (2021): 100090.

[14] L. Mauler, F. Duffner, and J. Leker. Economies of scale in battery cell manufacturing: The impact of material and process innovations. Applied Energy 286 (2021): 116499.

[15] K. Liu, X. Hu, H. Zhou, L. Tong, D. Widanalage, and J. Marco. Feature analyses and modelling of lithium-ion batteries manufacturing based on random forest classification. IEEE/ASME Transactions on Mechatronics (2021).

[16] J. Schnell, C. Nentwich, F. Endres, A. Kollenda, F. Distel, T. Knoche, and G. Reinhart. Data mining in lithium-ion battery cell production. Journal of Power Sources 413 (2019): 360-366.

[17] A. Turetskyy, S. Thiede, M. Thomitzek, N. V. Drachenfels, T. Pape, and C. Herrmann, Toward Data-Driven Applications in Lithium-Ion Battery Cell Manufacturing. Energy Technology, 1900136, (2019).

[18] L. Hoffmann, J. K. Grathwol, W. Haselrieder, R. Leithoff, T. Jansen, K. Dilger, K. Dröder, A. Kwade, and M. Kurrat. Capacity Distribution of Large Lithium-Ion Battery Pouch Cells in Context with Pilot Production Processes. Energy Technology 8, no. 2 (2020): 1900196.

[19] V. Wenzel, H. Nirschl, and D. Nötzel. Challenges in Lithium‐Ion‐Battery Slurry Preparation and Potential of Modifying Electrode Structures by Different Mixing Processes. Energy Technology 3, no. 7 (2015): 692-698.

[20] R. P. Cunha, T. Lombardo, E. N. Primo, and A. A. Franco. Artificial intelligence investigation of NMC cathode manufacturing parameters interdependencies. Batteries & Supercaps 3, no. 1 (2020): 60-67.

[21] Y. Zhou, T. A. Mazzuchi, and S. Sarkani. M-AdaBoost-A based ensemble system for network intrusion detection. Expert Systems with Applications 162 (2020): 113864.

[22] Y. Li, H. Shi, F. Han, Z. Duan, and H. Liu. Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy. Renewable energy 135 (2019): 540-553.

[23] M. K. Warmuth, J. Liao, and G. Rätsch. Totally corrective boosting algorithms that maximize the margin. In Proceedings of the 23rd international conference on Machine learning, pp. 1001-1008. 2006.

[24] H. Chen, Z. Lin, and C. Tan. Random subspace-based ensemble modeling for near-infrared spectral diagnosis of colorectal cancer. Analytical biochemistry 567 (2019): 38-44.