Capture the high-efficiency non-fullerene ternary organic solar cells formula by machine-learning-assisted energy-level alignment optimization

Highlights
- ML assists in analyzing energy-level alignment of non-fullerene ternary blends
- Random forest approach provides the best predictive capability
- The effective global optimization scheme in material selection is provided

Authors
Tianyu Hao, Shifeng Leng, Yankang Yang, ..., Yecheng Zou, Yongming Zhang, Feng Liu

Correspondence
fengliu82@sjtu.edu.cn

In brief
We built the dataset based on non-fullerene ternary organic solar cells (OSCs), and used machine-learning (ML) technology to evaluate the energy-level alignment, predict devices performance, and capture the high-performance non-fullerene ternary combinations. Important understanding regarding the relationship between molecular energy-level alignment and device performance in non-fullerene ternary OSCs was quantified by the scale of Gini importance of random forest model, guiding an effective global optimization scheme and direction of primary efforts.
Article

Capture the high-efficiency non-fullerene ternary organic solar cells formula by machine-learning-assisted energy-level alignment optimization

Tianyu Hao,1,3 Shifeng Leng,1,3 Yankang Yang,1 Wenkai Zhong,1 Ming Zhang,1 Lei Zhu,1 Jingnan Song,1 Jinqiu Xu,1 Guanqing Zhou,1 Yecheng Zou,2 Yongming Zhang,1,2 and Feng Liu1,2,4,*

1School of Chemistry and Chemical Engineering, Frontiers Science Center for Transformative Molecules, In-situ Center for Physical Science, and Center of Hydrogen Science Shanghai Jiao Tong University, Shanghai 200240, P.R. China
2State Key Laboratory of Fluorinated Functional Membrane Materials and Dongyue Future Hydrogen Energy Materials Company, Zibo City, Shandong Province 256401, P.R. China
3These authors contributed equally
4Lead contact
*Correspondence: fengliu82@sjtu.edu.cn
https://doi.org/10.1016/j.patter.2021.100333

SUMMARY

Appropriate energy-level alignment in non-fullerene ternary organic solar cells (OSCs) can enhance the power conversion efficiencies (PCEs), due to the simultaneous improvement in charge generation/transportation and reduction in voltage loss. Seven machine-learning (ML) algorithms were used to build the regression and classification models based on energy-level parameters to predict PCE and capture high-performance material combinations, and random forest showed the best predictive capability. Furthermore, two sets of verification experiments were designed to compare the experimental and predicted results. The outcome elucidated that a deep lowest unoccupied molecular orbital (LUMO) of the non-fullerene acceptors can slightly reduce the open-circuit voltage ($V_{OC}$) but significantly improve short-circuit current density ($J_{SC}$), and, to a certain extent, the $V_{OC}$ could be optimized by the slightly up-shifted LUMO of the third component in non-fullerene ternary OSCs. Consequently, random forest can provide an effective global optimization scheme and capture multi-component combinations for high-efficiency ternary OSCs.

INTRODUCTION

In recent years, organic solar cells (OSCs) have attracted much attention as an effective way of using solar energy due to their features of low cost, flexibility, and large-scale fabrication.1,2 Solution-processed bulk heterojunction (BHJ) OSCs composed of electron-donor (D) and -acceptor (A) combinations, based on $\pi$-conjugated polymers and $\pi$-conjugated small molecules, play a leading role in realizing high power conversion efficiencies (PCEs). In the past few years, ternary BHJ OSCs featuring...
multiple donor or acceptor materials in the active layer have emerged as a promising structure to simultaneously improve all solar cell parameters with enhanced light-harvesting capability and reduced energy loss compared with traditional binary OSCs. For ternary OSCs, device efficiency strongly depends on the energy-level alignment of each component and also the film morphology that governs the charge generation, transport, and extraction. Thus, fine-tuning the energy-level alignment with favorable morphology can concurrently enhance the short-circuit current density \( J_{SC} \) and open-circuit voltage \( V_{OC} \) of resultant solar cells. The PCEs of ternary OSCs have been steadily boosted in the past decades, through an array of research covering material design, morphology control, device engineering, among other topics. However, an efficient material combination remains challenging to obtain due to a great variety of material systems, which could only be explored through incessant trial and error experiments. The approach is inefficient and time consuming, which suppresses the material development and device efficiency improvement, and high-performance compound mode based on developed materials might have been missed. There has been a large amount of data generated in related activities that focus on different aspects in efficiency optimization, which, however, can hardly be utilized under a global perspective to summarize the mechanism and predict the future trend. The data-science-based machine-learning (ML) technique provides a new avenue that can assist in the analysis of massive experimental data in an unbiased manner to construct analytical models with predictive capabilities. For example, Schmidt and co-workers used a series of ML approaches to screen complex molecules to improve the device efficiency of OSCs. Lee et al. suggested that the PCEs of fullerene-based ternary OSCs were predictive on the basis of random forest model. Min et al. demonstrated that the random forest and boosted regression tree models can be employed to find D:A pairs for high-performance binary non-fullerene OSCs. In the OSCs field, a general preference for fullerene derivatives as electron acceptors has been superseded by non-fullerene acceptors (NFAs) since 2015. NFAs, which have been proved to be the next generation of high-efficiency materials for the construction of ternary OSCs, have a much stronger light absorption coefficient and highly adjustable energy levels of frontier molecular orbitals in comparison with fullerene derivatives. Thus, many varieties of ternary blend formula can be constructed to further improve the performance of OSCs. As substantial reports have shown, NFA-based ternary blends are promising, but they need to seek out the fundamental mechanism that gives rise to performance. The complicated charge dynamics of NFA-based ternary blends are undetermined owing to the complex chemical composition and film morphology, which entails the limitation of further development of ternary OSCs. To overcome these limitations, ML can be a theoretical tool to research the dataset containing the energy levels of three components and device performance parameters based on published articles about NFA-based ternary OSCs. According to correlation matrices, it was found that the deep lowest unoccupied molecular orbital (LUMO) of NFA material (the major content phase) would slightly reduce \( V_{OC} \) but significantly improve \( J_{SC} \) of devices, and the device \( V_{OC} \) could be improved by reducing differences between the LUMO of the major acceptor phase and the third component (the minor content phase), under the principle of balancing light harvest and energy loss. Different ML algorithms were performed to build regression and classification models to validate the statistical results of correlation matrices. In the regression task, we comprehensively evaluate the predictive capability of ML models with four metrics: coefficient of determination \( R^2 \), root-mean-square error \( \text{RMSE} \), mean absolute error \( \text{MAE} \), and mean absolute percentage error \( \text{MAPE} \). In addition, accuracy, recall rate, and area under curve (AUC) of the receiver operating characteristic (ROC) curve were used to assess the capability to capture high-performance devices of the different classification models. The results show that the random forest algorithm is effective to evaluate the energy-level alignment of ternary non-fullerene blends and guide material selection for obtaining high-performance ternary OSCs. These results were then extended into verification experiments in which a series of ternary blends of different comparative groups were made, yielding results in good accordance with the ML formulation.

RESULTS AND DISCUSSION

Dataset

The dataset in this study contains 157 data points over 1,500 available experimental parameters, including energy levels of LUMO and the highest occupied molecular orbital (HOMO), \( J_{SC} \), \( V_{OC} \), fill factor (FF), and PCE, covering a major publication body of non-fullerene ternary OSCs from 2015 to 2020 (see Data S1 for details on the search). The overall PCE distribution of ternary OSCs devices is shown in Figure 1, which is in the range of 6.6%–17.12%, and the average and standard deviations are 12.53% and 2.76%, respectively. A double peak distribution pattern (one at 11% and one at 16%) marks the steady-state optimization of conventional material blends and the upgrading of new material systems. The statistical result implies that there is high potential to further improve OSC performance by designing better ternary combinations using emerging materials.

Correlation matrices

To evaluate the correlations between the molecular energy levels and device parameters, Pearson correlation coefficients \( r \) were calculated. Figure 2 summaries correlation matrices of the \( r \) values for the HOMO and LUMO levels with respect to devices \( V_{OC} \), \( J_{SC} \), and PCE. As shown in Figure 2A, the \( V_{OC} \) is correlated with the LUMO of acceptor (LUMO1, major phase) and the third component (LUMO2, minor phase) with \( r \) values of 0.32 and \(-0.23 \), respectively. The result suggests that LUMO1 and LUMO2 have opposite effects on \( V_{OC} \) in ternary OSCs. Specifically, a shallow LUMO1 is beneficial to obtain a higher \( V_{OC} \). The \( V_{OC} \) can also be further improved when the slightly upshifted LUMO2 approaches the shallow LUMO1. In addition, the \( V_{OC} \) is correlated with the HOMO of donor \( (r = -0.12) \) and...
the correlations of other indicators with \( V_{\text{OC}} \) are negligible, indicating that the HOMO of the donor also contributes to the origin of \( V_{\text{OC}} \) in ternary non-fullerene OSCs. As shown in Figure 2B, \( J_{\text{SC}} \) is strongly correlated with LUMO1 \((r = -0.71)\), and the correlations between \( J_{\text{SC}} \) and other parameters are relatively weak. This result indicates that \( J_{\text{SC}} \) is dominated by major phase LUMO1, and a deeper LUMO1 can make a noticeable reduction in thin film band gap to improve \( J_{\text{SC}} \). The PCE of an OSC device is defined by the equation

\[
PCE = \frac{(V_{\text{OC}} \times J_{\text{SC}} \times FF)}{P_{\text{i}}} \]  

where \( P_{\text{i}} \) is the incident light power. Thus, the effect of LUMO1 is translated in a product manner to PCE showing high correlation with \( r \) of \(-0.71\) (Figure 2C). Meanwhile, a relatively weak correlation between PCE and LUMO2 of the third component \((r = -0.24)\) is recorded, which is consistent with the relativity between \( V_{\text{OC}} \) and LUMO2. While the major phase LUMO1 needs to fit the bandgap and \( J_{\text{SC}} \) optimization, the minor phase LUMO2 turns out to be a useful handle to optimize \( V_{\text{OC}} \) in ternary OSCs, resulting in a useful deconvolution of parameter space toward target optimization. Parametric analysis based on the correlation matrix finds out the scale of correlation importance to guide the direction where primary efforts should be devoted. The correlations between energy levels and device parameters represent a schematic of the non-fullerene ternary OSCs’ working mechanism. As proposed in Figure 2D, the deep LUMO is beneficial for the photon absorption to generate high \( J_{\text{SC}} \), D:A:T ternary blends following a cascading energy-level alignment are expected to accelerate the dissociation process of generated excitons at multiple interfaces (D-T, T-A, D-A) aided by terraced electron and hole transfer processes, to minimize the thermoresolution as indicated by charge transfer in molecular coordinates.36–38 According to Marcus theory, the energy offset of charge-transfer state \( E_{\text{CT}} \) with respect to the first excited singlet state \( S_1 \) determines the barrier for hole-electron separation.

**Regression model**

All models were written in Python, and NumPy was used to prepare and manipulate the dataset. Afterward, seven ML algorithms accessible from Scikit-Learn were used to build predictive models. 80% of data (125 data points) were randomly input for models building and training, and the remaining 20% of data (32 data points) were selected as a test set to validate capability of regression models to predict PCEs. ShuffleSplit cross-validation technique was used to train models, and details can be found in the regression model building section. Moreover, four evaluation metrics, R2, RMSE, MAE, and MAPE, were chosen to measure the models’ forecasting performance comprehensively.

In multivariate regression, the overview of the seven algorithms performance, including random forest, XGBoost, K-nearest neighbors (KNNs), decision tree, support vector machine (SVM), ridge regression, and multi-layer perceptron (MLP), are shown in Table 1. Decision tree-based models that are based on true/false responses, including random forest, XGBoost, and decision tree, significantly outperform the other four algorithms. Even a single regression tree (decision tree algorithm) to predict PCE can get a better result \( (R^2 = 0.62, \text{RMSE} = 1.63, \text{MAE} = 1.23, \text{MAPE} = 0.10) \) than other algorithms. However, a single decision tree is not sufficient for the prediction of device parameters as it tends to overfit and provide poorly generalized performance, whereas the random forest and XGBoost algorithms repeatedly fit multiple regression trees to improve models predictive performance.39 As summarized in Table 1, random forest gives the best results compared with the other six ML algorithms, with maximum \( R^2 \), and minimum \( \text{RMSE}, \text{MAE}, \) and \( \text{MAPE} \) on the cross-validation mean score \( (R^2 = 0.75, \text{RMSE} = 1.39, \text{MAE} = 1.04, \text{MAPE} = 0.09) \) and the test set score \( (R^2 = 0.73, \text{RMSE} = 1.44, \text{MAE} = 1.08, \text{MAPE} = 0.09) \), respectively. As described above, the random forest model can provide a favorable outcome to predict the devices’ PCEs. Owing to the highest \( R^2 \) and the lowest \( \text{RMSE}, \text{MAE}, \) and \( \text{MAPE} \), the model based on random forest algorithm was
selected for further investigation. The regression model results for both training and test datasets are depicted in Figure 3A. As shown in the figure, the predictive PCEs of the random forest model are quite close to true PCEs as expected without severe overfitting or underfitting. Moreover, the random forest algorithm offers relative importance of each feature, which is defined rigorously by the sum of impurity distance over all nodes of all trees in the model, to get an intuitive interpretation about how a descriptor affects the target variables. 46–48

Figure 3B shows the Gini importance computed for random forest. As can be seen, the LUMO1 is an outstanding factor in determining the device PCE of non-fullerene ternary system compared with other parameters, which is consistent with the correlation matrixes analysis. The HOMO of donor displays secondary importance in different energy levels, since the HOMO level of donor interacts with LUMO1 to affect the active layer optical bandgap and built-in potential, which partially determine the device $J_{SC}$ and $V_{OC}$ respectively. Moreover,
Shapley additive explanations (SHAP) analysis was also used to find out the feature importance, which got similar results to Gini importance (Figure S3, supplemental information). It should be clarified that the low feature importance of the energy levels of the third minor component do not mean the associated descriptors are irrelevant to the PCE. According to correlation matrixes, the LUMO2 of the third component plays a role in improving device $V_{OC}$ but does not severely reduce device $J_{SC}$. Recent reports have also shown that the device PCE is more profoundly correlated with the $J_{SC}$ in comparison with the $V_{OC}$, which is consistent with the low Gini importance of the third component energy levels. The relatively poor correlation between the device PCE and $V_{OC}$ is recorded mainly because in a fixed binary blend base, or even across different binary blend compositions, the change in $V_{OC}$ is small, usually a 5%–10% value change. Consequently, the main outcomes afforded by the random forest regression model are the high feature importance of LUMO1 and HOMO, and relatively low importance of LUMO2 and HOMO2, and the regression model works well to predict PCEs of NFA-based ternary OSCs. Combining with correlation matrixes analyses, it is indicated that a proper energy-level alignment is quite important to promote device efficiency, and tuning the LUMO1 is indeed an effective tactic to improve $J_{SC}$ significantly with reduction in $V_{OC}$ that can be offset moderately by choosing appropriate LUMO2 without obvious current loss.

Classification model
As for classification algorithms, the output of classifiers can give a more intuitive conclusion as to whether the non-fullerene ternary blends are able to produce high-performance devices. In the previous study, the original dataset was divided into low and high performance with similar data points based on the devices’ PCEs to obtain an unbiased classification dataset. By contrast, our models were established to capture a few high-performance NFA-based ternary OSCs based on the biased classification process. As shown in Figure 4A, the original dataset was separated into two categories unbalanced by using a high PCE benchmark of 16%, labeled class1 and class2. Low-performance devices (class1, PCE < 16%) occupied ~80% of the dataset, while high-performance devices (class2, PCE ≥ 16%) occupied only nearly ~20%.

Table 1. Performance of regression ML models evaluated by R2, RMSE, MAE, and MAPE metrics

| ML techniques       | Cross-validated results | Test set       |
|---------------------|-------------------------|----------------|
|                     | R2          | RMSE         | MAE           | MAPE         | R2          | RMSE         | MAE           | MAPE         |
| Random forest       | 0.75 ± 0.09 | 1.39 ± 0.24  | 1.04 ± 0.18   | 0.09 ± 0.018 | 0.73        | 1.44         | 1.08          | 0.09         |
| XGBoost             | 0.70 ± 0.09 | 1.48 ± 0.20  | 1.09 ± 0.16   | 0.09 ± 0.016 | 0.66        | 1.56         | 1.16          | 0.09         |
| KNN                 | 0.62 ± 0.11 | 1.65 ± 0.23  | 1.29 ± 0.21   | 0.11 ± 0.021 | 0.59        | 1.71         | 1.33          | 0.11         |
| Decision tree       | 0.62 ± 0.16 | 1.67 ± 0.28  | 1.28 ± 0.22   | 0.12 ± 0.021 | 0.62        | 1.63         | 1.23          | 0.10         |
| SVM                 | 0.62 ± 0.11 | 1.67 ± 0.23  | 1.29 ± 0.19   | 0.11 ± 0.019 | 0.57        | 1.76         | 1.51          | 0.12         |
| Ridge regression    | 0.53 ± 0.08 | 1.86 ± 0.20  | 1.54 ± 0.21   | 0.14 ± 0.023 | 0.25        | 2.35         | 1.82          | 0.15         |
| MLP                 | 0.62 ± 0.11 | 1.67 ± 0.24  | 1.33 ± 0.20   | 0.11 ± 0.020 | 0.16        | 2.62         | 2.11          | 0.16         |

Figure 3. The results of regression random forest model
(A) ML model predicted versus experimental true PCE for both training and testing sets. The red and blue points are training and testing data, respectively, and the blue line is provided as a reference for fitting results.
(B) Gini importance of descriptors for random forest model.
Furthermore, the same seven algorithms were applied to build classification ML models. The performance of classification models can be evaluated by computing the number of correctly recognized class 2 examples (true positives [TPs]), the number of correctly recognized examples that do not belong to class 2 (true-negatives [TNs]), and examples that either were incorrectly assigned to the class 2 (false-positives [FPs]) or that were not recognized as class 2 examples (false-negatives [FNs]). As shown in Figure 4B, a confusion matrix in the case of a binary classification is created with these four values. With confusion matrix, many performance evaluation metrics can be obtained, including accuracy, error rate, and recall. In this work, the accuracy, recall rate, and AUC of ROC curve were chosen as the evaluation criteria to assess the seven algorithms, and all the preliminary results are listed in Table 2. The training accuracies of the seven algorithms produce scores in the range of 0.81–0.99, where MLP and random forest model render the lowest and highest accuracy, respectively. Besides the MLP model, the other six classifiers can still achieve high accuracies in the range of 0.87–0.97 in the test set. However, such high accuracies do not suggest each algorithm is successful for classification. Instead, only using accuracy as the evaluation criterion is not able to assess ML models’ performance. We here split the original dataset into low- and high-performance ternary material combinations unbalanced deliberately. For this kind of unbalanced dataset with binary target value, simple fitting statistics cannot give a comprehensive and accurate evaluation of the model, since the accuracy is dependent on the threshold used to divide raw dataset.

Thus, the recall rate and AUC-ROC are chosen as the other two metrics based on the confusion matrix to evaluate the performance of different models. The recall rate is defined as the number of correctly classified positive examples divided by the number of positive examples in the data, formulated as \( \text{Recall} = \frac{TP}{TP + FN} \), which can easily be obtained from the confusion matrix. ROC curve is a standard way to evaluate the biased model with binary outcomes and is often used to evaluate the predictive ability of ML models. The value of AUC, which can be calculated by integrating the ROC curve, summarizes the relationship between the TP and FP rate of a binary classifier. Several previous works have reported that AUC is preferable to accuracy for evaluating classifiers’ capability to capture the minor samples, making it one of the popular metrics for imbalanced datasets. As for the recall rate and AUC values shown in Table 2, the tree-based classifiers get favorable results expected to be similar to the previous regression models in this work. From Table 2, random forest, XGBoost, and decision tree achieve high recall rate scores of 0.97, 0.92, and 0.90 in the train set, and 0.84, 0.83, and 0.83 in the test set. Although KNN, SVM, and ridge classifier also have accuracies of approximately 0.9 in both train and test set, the low recall rate results are unsatisfactory. Especially for support vector classifier and ridge classifier algorithms, the recall results are only 0.58 and 0.50, which imply that these models are not able to distinguish minor high-performance devices. As for AUC metrics, if the value is in the range of 0.9–1, we consider the model as a perfect classifier. If the value range is 0.8–0.9, that means the model works well. A model with AUC in the range of 0.5–0.8 is normal. The worst possible AUC is 0.5 as this means the model does not have any capacity to discriminate between the positive and negative classes. Low AUC values close to zero mean that the classifier is still working well, and it is just reciprocating the positive and negative class. Random forest algorithm gets the highest AUC of 0.97 and 0.92 in train and test sets, and the other two tree-based algorithms also get admirable AUC values of about 0.9.

In contrast, the KNN, SVM, ridge classifier, and MLP perform normally with AUC values below 0.8 in the test set. Moreover, we repeated the experiments at two other thresholds of 12% and 14% as cutoffs, and the results are shown in Tables S1 and S2. Further, a multi-class classification, labeled class 1 (PCE < 11%), class 2 (11% ≤ PCE < 16%) and class 3 (PCE ≥ 16%), was performed, and the results are shown in Table S3. Similar to the previous classification model, the tree-based algorithms significantly outperform the other algorithms, and random
The analysis based on the random forest model is beneficial to provide practical understanding of the interactions between device performance and energy levels of non-fullerene ternary blend. In addition, the ML model is able to capture high-efficiency NFA-based ternary OSCs. Therefore, in the verification experiment section, the random forest classifier was used to judge and analyze whether and why the material combinations can produce high-efficiency devices.

Verification experiment
In addition to verifying the feasibility of the model, two sets of experiments for tuning LUMO1 and LUMO2 were designed. In the first set, the ternary blend comprising PM6(Poly[2,6-(4,8-bis(5-(2-ethylhexyl)-3-fluoro)thiophen-2-yl)-benzo[1,2-b:4,5-b']dithiophene))-alt-(5,5-(1',3'-di-2-thienyl-5',7'-bis(2-ethylhexyl)benzo[1',2'-c:4',5'-c']dithiophene-4,8-dione)])X-1:ITIC(3,9-bis(2-methylylene-(3-(1,1-dicyanomethylene)-indanone))-5,5,11,11-tetrakis(4-hexyphenyl)-dithieno[2,3-d:2',3'-d']-s-indaceno[1,2-b:5,6-b']dithiophene) was chosen, where X-1 represented three different acceptors derived from Y6(2,2'-(4(Z,2'Z)-(12,13-bis(2-ethylhexyl)-3,9-diundecyloxy)-12,13-dihydro-[1,2,5]thiadiazolo[3,4-е]thieno[2',3':4,5']thieno[2',3:4,5]pyrrolo[3,2-g]thieno[2',3:4,5]thieno[3,2-b]indole-2,10-diyli bis(methanonyli dene)bis(5,6-difluoro-3-oxo-2,3-dihydro-1H-indene-2,1-diylidene)dimalononitri le), and ITIC was used as the third minor component. In the second set of experiments, PM6:Y6:X-2 was selected as the ternary blend, where X-2 represented three different acceptors analogous to ITIC, and Y6 was major acceptor phase. The random forest classification model was used to predict whether the material combinations could produce high-efficiency devices. All of the primary material energy levels, device parameters, and classification ML model results are tabulated in Tables 3 and 4. As shown in Table 3, the Y6-based ternary device shows the lowest JSC of 25.00 mA cm⁻² and highest Voc of 0.868 V due to the shallow LUMO1 of acceptor Y6. As for BTP-BO-4F(2,2'-((2Z,2'Z)-(12,13-bis(2-butyloctyl)-3,9-diundecyloxy)-12,13-dihydro-[1,2,5]thiadiazolo[3,4-e]thieno[2',3':4,5']thieno[2,3:4,5]pyrrolo[3,2-g]thieno[2',3:4,5]thieno[3,2-b]indole-2,10-diyli bis(methanonyli dene)bis(5,6-difluoro-3-oxo-2,3-dihydro-1H-indene-2,1-diylidene)dimalononitri le), and ITIC contributes to the enhancement of light absorption, showing the highest JSC of 25.35 mA cm⁻² but lowest Voc of 0.862 V. In particular, compared with binary reference (PM6: ITIC), the device VC of ternary blends drops by 20%, while JSC gets improved by approximately 80%. This result indicates that LUMO1 has a strong correlation with device JSC to affect device efficiency. According to the ML result, the model successfully captures high-performance material combination PM6:Y6:ITIC (PCE > 16%) and regards the other two combinations as class 1 (PCE < 16%). As shown in Table 4, the addition of the third minor component can optimize Voc effectively, but the JSC edges down slightly. Tuning the up-shifted LUMO2 from −4.15 eV to −4.18 eV improves Voc from 0.848 V to 0.871 V, indicating that an appropriate LUMO2 can help to balance the trade-off between device current and voltage, and the ML model correctly classify PM6:Y6:IT-M(3,9-bis(2-methylene-(3-(1,1-dicyanomethylene)-6,7-methylidene)-indanone))-5,5,11,11-tetrakis(4-hexyphenyl)-dithieno[2,3-d:2',3'-d']-s-indaceno[1,2-b:5,6-b']dithiophene and PM6:Y6:IT-4F(3,9-bis(2-methylene-(3-(1,1-dicyanomethylene)-6,7-difluoro- indanone))-5,5,11,11-tetrakis(4-hexyphenyl)-dithieno[2,3-d:2',3'- d']-s-indaceno[1,2-b:5,6-b']dithiophene) as class 2. However, the

Table 2. Performance of classification ML models evaluated by accuracy, recall rate, and AUC

| ML techniques     | Train set | Test set |
|-------------------|-----------|----------|
|                   | Accuracy  | Recall   | AUC     | Accuracy  | Recall   | AUC     |
| Random forest     | 0.99      | 0.97     | 0.97    | 0.84      | 0.92     |
| XGBoost           | 0.98      | 0.92     | 0.96    | 0.83      | 0.91     |
| KNN               | 0.93      | 0.69     | 0.84    | 0.50      | 0.73     |
| Decision tree     | 0.95      | 0.91     | 0.95    | 0.83      | 0.90     |
| SVM               | 0.90      | 0.58     | 0.78    | 0.90      | 0.50     |
| Ridge classifier  | 0.90      | 0.50     | 0.75    | 0.90      | 0.50     |
| MLP               | 0.81      | 0.08     | 0.54    | 0.78      | 0.00     | 0.50    |
ML model fails to distinguish high-performance combination PM6:Y6:IT-4Cl(3,9-bis(2-methylene-((3-(1,1-dicyanomethylene)-6,7-dichloro)-indanone))-5,5,11,11-tetrakis(4-hexylphenyl)-dithieno[2,3-d:2',3'-d']-s-indaceno[1,2-b:5,6-b']dithiophene), where the LUMO2 of IT-4Cl is deeper than the LUMO1 of Y6. Thus, in order to improve the generalization ability of our ML model, further data collection containing other important features, for example morphology and molecular structure, needs to be accessed based on the current dataset. Despite that, the ML model is still applicable to assist energy-level alignment optimization and capture high-efficiency NFA-based ternary OSCs.

**Conclusion**

In conclusion, the ML models based on the random forest algorithm can provide an effective global optimization scheme in material selection to optimize energy-level alignment of non-fullerene ternary OSCs: (1) the shallow HOMO of donor and the deep LUMO1 of the major acceptor phase to improve device $J_{SC}$; (2) slightly up-shifted LUMO2 of the third minor phase to modulate device $V_{OC}$. In this work, based on correlation matrix analyses, ML regression and classification models were built to predict PCEs of ternary non-fullerene OSCs and capture high-efficiency material combinations by using seven algorithms. With data preprocessing, hyperparameter modulation, and cross-validation technique, the random forest algorithm was found to be very superior to other ML algorithms. In the regression tasks, the random forest approach achieved the highest $R^2$ and the lowest RMSE, MAE, and MAPE when it was used to predict the PCE of non-fullerene ternary OSCs. In addition, the fine-tuned random forest classification model got the highest accuracy, recall rate, and AUC scores, indicating that the model worked well to identify efficient material combinations for producing high-efficiency non-fullerene ternary OSCs. Moreover, the random forest model helped to explain the relationship between the molecular energy-level alignment and device performance, and successfully captured high-efficiency material combinations in our verification experiment. All the results of correlation matrixes, ML models, and verification experiments were in good accordance with each other. Therefore, this work reveals that utilizing advanced ML technology benefits the development of materials and devices.

### Table 3. Material energy levels, photovoltaic parameters, and classification ML model result of PM6:ITIC and PM6:X1:ITIC OSCs where X1 represents BTP-BO-4F, BTP-BO-4Cl and Y6

|                | LUMO1 eV | HOMO1 eV | LUMO2 eV | HOMO2 eV | $J_{SC}$ mA cm$^{-2}$ | $V_{OC}$ V | FF % | PCE % | ML result |
|----------------|----------|----------|----------|----------|-----------------------|-----------|------|-------|-----------|
| None           |          |          | −3.88    | −5.52    | 13.99                 | 1.003     | 62.6 | 8.78  |           |
| BTP-BO-4F      | −4.28    | −5.67    | −3.88    | −5.52    | 25.35                 | 0.862     | 72.5 | 15.8  | class 1   |
| BTP-BO-4Cl     | −4.24    | −5.62    | −3.88    | −5.52    | 25.12                 | 0.867     | 71.9 | 15.6  | class 1   |
| Y6             | −4.20    | −5.56    | −3.88    | −5.52    | 25.00                 | 0.868     | 78.2 | 16.9  | class 2   |

Figure 5. The results of classification random forest model

(A) Normalized confusion matrix of the random forest classifier model. (B) ROC curve and AUC of the random forest classifier model. (C) Gini importance of descriptors for random forest classifier model. (D) One of the decision trees in this random forest classifier that shows how the original data is classified based on true/false responses.
of the photovoltaics field in three ways: (1) validating statistical relationships between the energy levels and devices performance; (2) proving availability of the ML method in non-fullerene ternary OSCs; (3) guiding the direction of primary optimization efforts based on the scale of Pearson correlation coefficients and feature importance. It should be noted that other features, for example morphology and molecular structure, are important and need further data collection, but they are beyond the scope of this paper. There is still great potential to promote the ML model involving other important factors for wider application and higher precision.

**Experimental Procedures**

**Resource availability**

**Lead contact**
Prof. Feng Liu is the lead contact for this study and can be reached at fengliu82@sjtu.edu.cn.

**Materials availability**
Chloroform (Sigma-Aldrich, ≥ 99%), methanol (Alfa, 99.8%), and 1-chloronaphthalene (CN) (TCI, 97%) were purchased from commercial sources and used without further purification. PM6, Y6, BTP-BO-4F, BTP-BO-4Cl, IT-4F, IT-4Cl, IT-M, and PFN-Br were purchased from Solamer Material Inc.

**Data and code availability**
The dataset is provided in supplemental information. The hyperparameters of all regression and classification tasks used in code are shown in Table S4.

**Regression model building**
The Scikit-Learn is a Python module integrating a range of advanced ML algorithms for medium-scale supervised and unsupervised problems. To build predictive models, seven ML algorithms accessible from Scikit-Learn, including random forest, XGBoost, KNNS, decision tree, SVM, ridge regression, and MLP, were performed. In all ML algorithms, a set of descriptors were used as input features to obtain the output $y$. In this study, the output PCE was effectively determined by a function $f(x_i)$, and the form of the function and optimization procedure of its hyperparameters was strongly dependent on the selected algorithm. First, continuous real-valued features were normalized to $[0, 1]$ by the MinMaxScalar module of the Scikit-Learn package. Then the ShuffleSplit cross-validation technique was used to avoid inconsistency and ensure the predictive accuracy of models, in which the training set was shuffled and split into train set and validation set 50 times to evaluate the performance of each material parameter combination. Further, how the hyperparameters of random forest affect model performance is shown in Figure S4.

**Classification model building**
The original PCEs were separated unbalanced into low-performance devices (class 1, PCE < 16%) and high-performance devices (class 2, PCE > 16%), which occupied ~80% and ~20% of the dataset respectively. Similar to the regression models, the same seven algorithms were applied to build classification ML models. The hyperparameters optimization process of the random forest classification model is shown in Figure S5.

### Table 4. Material energy levels, photovoltaic parameters, and classification ML model result of PM6:Y6 and PM6:Y6:X2 OSCs where X2 represents IT-4Cl, IT-4F and IT-M

| LUMO1 eV | HOMO1 eV | LUMO2 eV | HOMO2 eV | $J_{SC}$ mA cm$^{-2}$ | $V_{OC}$ V | FF | PCE % | ML result |
|----------|----------|----------|----------|----------------------|----------|-----|-------|-----------|
| None     | -4.20    | -5.56    |          | 25.24                | 0.841    | 77.6| 16.5  |           |
| IT-4Cl   | -4.20    | -5.56    | -4.29    | -5.78               | 24.98    | 0.846| 77.2  | class 1   |
| IT-4F    | -4.20    | -5.56    | -4.15    | -5.70               | 24.94    | 0.848| 78.3  | class 2   |
| IT-M     | -4.20    | -5.56    | -4.18    | -5.74               | 24.67    | 0.871| 76.5  | class 2   |

**Model performance metrics**
As for regression tasks, $R^2$, RMSE, MAE, and MAPE were used to evaluate the models listed in Equations (1), (2), (3), and (4).

\[
R^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{(Equation 1)}
\]

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad \text{(Equation 2)}
\]

\[
MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad \text{(Equation 3)}
\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \quad \text{(Equation 4)}
\]

where $y_i$ and $\hat{y}_i$ represent the measured value and the predicted value, and $\bar{y}_i$ and $\bar{\hat{y}}_i$ represent the mean value for the measured and the predicted values, respectively, and $n$ is the total number of the dataset.

In classification tasks, the accuracy, recall rate, and AUC of ROC curve were chosen as the evaluation criteria to assess the seven algorithms. Based on confusion matrix, comprising TPs, TNs, FPs, and FN, accuracy and recall rate were defined as:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{(Equation 5)}
\]

\[
\text{Recall} = \frac{TP}{TP + FP} \quad \text{(Equation 6)}
\]

To plot the ROC curve, a series of thresholds were picked, and the original dataset was subsequently used to train and test classifier models for obtaining all confusion matrices matched with all possible thresholds. After that, the TP rate (TPR) and FP rate (FPR) associated with each threshold was computed. Eventually, FPR and TPR were set as the horizontal axis and vertical axis, and points were drawn whose coordinates were (FPR, TPR) of each threshold. In the ROC space, in this way, the ROC curve was obtained, and AUC could be calculated by integration.

**Fabrication and characterization of devices**
The ternary devices were fabricated with a structure of indium tin oxide (ITO)/PEDOT:PSS/active layer/PFN-Br/Ag. The glass substrates with pre-patterned ITO were cleaned by sonication in detergent, deionized water, acetone, and isopropanol for 15 min successively. After drying, the substrates were treated in an ultraviolet-ozone chamber for 15 min. Then a PEDOT:PSS layer (40 nm) was deposited through spin-coating from a PEDOT:PSS aqueous solution (Baytron P VP Al 4083 from H. C. Starck) at 3,500 rpm and dried subsequently at 150°C for 15 min in air. Then the device was transferred to a nitrogen glove box, where the active blend layer of PM6A1:A2 (A1 is Y6, BTP-BO-4F, or BTP-BO-4Cl, and A2 is IT-M, IT-4F, ITIC, and IT-4Cl) was spin-coated from its chloroform solution (7.5 mg/mL total) with additives (0.5% v/v) 1-CN onto the PEDOT:PSS layer. The thickness of the active layer was about 120 nm for these devices. Then a methanol solution of PFN-Br at a concentration of 0.5 mg/mL was deposited atop the active layer at 3,000 rpm for 30 s. Finally, top Ag electrode was deposited in vacuum onto the cathode buffer layer at a pressure of ~5.0 × 10⁻² Pa.
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.patter.2021.100333.

ACKNOWLEDGMENTS

T.H. and S.L. contributed equally to this work. This work was financially supported by the grant from the National Natural Science Foundation of China (grant no. 51973110, 21734009, 21905102, and 61805138).

AUTHOR CONTRIBUTIONS

T.H. conceived the idea and carried out ML and data process. S.L. built the dataset and guided the project. Y.Y. contributed to device fabrication. All authors contributed to the writing of the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 2, 2021
Revised: July 10, 2021
Accepted: July 26, 2021
Published: August 18, 2021

REFERENCES

1. Halls, J.J.M., Walsh, C.A., Greenham, N.C., Maseggia, E.A., Friend, R.H., Moratti, S.C., and Holmes, A.B. (1995). Efficient photodiodes from interpenetrating polymer networks. Nature 376, 498–500.
2. Li, G., Zhu, R., and Yang, Y. (2012). Polymer solar cells. Nat. Photon. 6, 153–161.
3. Ameri, T., Khoram, P., Min, J., and Brabec, C.J. (2013). Organic ternary solar cells: a review. 25, 4245–4266.
4. Zhang, J., Zhang, Y., Fang, J., Lu, K., Wang, Z., Ma, W., and Wei, Z. (2015). Conjugated polymer-small molecule alloy leads to high efficient ternary organic solar cells. J. Am. Chem. Soc. 137, 8176–8183.
5. Lu, L., Kelly, M.A., You, W., and Yu, L. (2015). Status and prospects for ternary organic photovoltaics. Nat. Photon. 9, 491–500.
6. Chen, Y.-C., Hsu, C.-Y., Lin, R.Y.-Y., Ho, K.-C., and Lin, J.T. (2013). Materials for the active layer of organic photovoltaics: ternary solar cell approach 6, 20–35.
7. Zhou, Z., Xu, S., Song, J., Jin, Y., Yue, Q., Qian, Y., Liu, F., Zhang, F., and Zhu, X. (2018). High-efficiency small-molecule ternary solar cells with a hierarchical morphology enabled by synergizing fullerene and non-fullerene acceptors. Nat. Energy 3, 952–959.
8. Yan, T., Song, W., Huang, J., Peng, R., Huang, L., and Ge, Z. (2019). 16.67% rigid and 14.06% flexible organic solar cells enabled by ternary heterojunction strategy. Adv. Mater. e1902210.
9. Song, J., Li, C., Zhu, L., Guo, J., Xu, J., Zhang, X., Weng, K., Zhang, K., Min, J., Hao, X., et al. (2019). Ternary organic solar cells with efficiency >16.5% based on two compatible nonfullerene acceptors. Adv. Mater. 31, e1905645.
10. Li, X., Meng, H., Shen, F., Su, D., Hsu, S., Shan, J., Huang, J., and Zhan, C. (2019). Semitransparent fullerene-free polymer solar cell with 44% AVT and 7% efficiency based on a new chlorinated small molecule acceptor. Dyes Pigment. 166, 196–202.
11. Xu, Y., Jiang, H., Lau, T.-K., Zhu, J., Wang, J., Lu, X., Zhan, X., and Lin, Y. (2019). Fused thiophene-benzene-thienophene electron acceptors for organic solar cells. J. Energy Chem. 37, 58–65.
12. Yuan, J., Zhang, Y., Zhou, L., Zhang, G., Yip, H.-L., Lau, T.-K., Lu, X., Zhu, C., Peng, H., Johnson, P.A., et al. (2019). Single-junction organic solar cell with over 15% efficiency using fused-ring acceptor with electron-deficient core. Joule 3, 1140–1151.
13. Li, C., Song, J., Cai, Y., Han, G., Zheng, W., Yi, Y., Ryu, H.S., Woo, H.Y., and Sun, Y. (2020). Heteroatom substitution-induced asymmetric A–D–A type non-fullerene acceptor for efficient organic solar cells. J. Energy Chem. 40, 144–150.
14. Yang, X., and Loos, J. (2007). Toward high-performance polymer solar cells: the importance of morphology control. Macromolecules 40, 1353–1362.
15. Moué, A.J., and Meerholz, K. (2009). Morphology control in solution-processed bulk-heterojunction solar cell mixtures 19, 3028–3036.
16. Liu, Y., Zhao, J., Li, Z., Mu, C., Ma, W., Hu, H., Jiang, K., Lin, H., Ade, H., and Yan, H. (2014). Aggregation and morphology control enables multiple cases of high-efficiency polymer solar cells. Nat. Commun. 5, 5293.
17. Kang, H., Lee, W., Oh, J., Kim, T., Lee, C., and Kim, B.J. (2016). From fullerene–polymer to all-polymer solar cells: the importance of molecular packing, orientation, and morphology control. Acc. Chem. Res. 49, 2424–2434.
18. Xie, Y., Yang, F., Li, Y., Uddin, M.A., Bi, P., Fan, B., Cai, Y., Hao, X., Woo, H.Y., Li, W., et al. (2018). Morphology control enables efficient ternary organic solar cells 30, 1803045.
19. Zhang, M., Zhu, L., Qiu, C.-q., Zhang, Y.-m., and Liu, F.J.A.P.S. (2019). Organic tandem solar cells: a review. Energy Environ. Sci. 12, 347–363.
20. Ameri, T., Dennler, G., Lungenschmied, C., and Brabec, C.J. (2009). Organic tandem solar cells: a follow up review. Energy Environ. Sci. 5, 2390–2413.
21. Liu, Y., Liu, F., Wang, H.-W., Nordlund, D., Sun, Z., Ferdous, S., and Russell, T.P. (2014). Sequential deposition: optimization of solvent swelling for high-performance polymer solar cells. ACS Appl. Mater. Interfaces 7, 653–661.
22. Ameri, T., Li, N., and Brabec, C.J. (2013). Highly efficient organic tandem solar cells: a follow up review. Energy Environ. Sci. 6, 2390–2413.
23. Liu, Y., Liu, F., Wang, H.-W., Nordlund, D., Sun, Z., Ferdous, S., and Russell, T.P. (2014). Sequential deposition: optimization of solvent swelling for high-performance polymer solar cells. ACS Appl. Mater. Interfaces 7, 653–661.
24. Li, J., Gao, S., Wang, X., Zhang, Q., Kan, B., Xia, R., Liu, F., Yang, X., Feng, H., Ni, W., et al. (2017). Solution-processed organic tandem solar cells with power conversion efficiencies >12%. Nat. Photon. 11, 85–90.
25. Pyzer-Knapp, E.O., Simm, G.N., and Aspuru-Guzik, A. (2016). A Bayesian approach to calibrating high-throughput virtual screening results and application to organic photovoltaic materials. Mater. Horiz. 3, 226–233.
26. Raccuglia, P., Elbert, K.C., Adler, P.D., Falk, C., Wenny, M.B., Mollo, A., Zeller, M., Friedler, S.A., Schrier, J., and Norquist, A.J. (2016). Machine-learning-assisted materials discovery using failed experiments. Nature 533, 73–76.
27. Sahu, H., Rao, W., Troisi, A., and Ma, H. (2018). Toward predicting efficiency of organic solar cells via machine learning and improved descriptors. Adv. Energy Mater. 8, 1801032.
28. Padula, D., Simpson, J.D., and Troisi, A. (2019). Combining electronic and structural features in machine learning models to predict organic solar cells properties. Mater. Horiz. 6, 343–349.
29. Sahu, H., Yang, F., Ye, X., Ma, J., Fang, W., and Ma, H. (2019). Designing promising molecules for organic solar cells via machine learning assisted virtual screening. J. Mater. Chem. A. 7, 17480–17488.
30. Jorgensen, P.B., Mesta, M., Shl, S., Garcia Lastra, J.M., Jacobsen, K.W., Thygesen, K.S., and Schmidt, M.N. (2018). Machine learning-based screening of complex molecules for polymer solar cells. J. Chem. Phys. 148, 241735.
31. Lee, M.-H. (2019). Insights from machine learning techniques for predicting the efficiency of fullerene derivatives-based ternary organic solar cells at ternary blend design. Nat. Commun. 9, 1900891.
32. Wu, Y., Guo, J., Sun, R., and Min, J. (2020). Machine learning for accelerating the discovery of high-performance donor/acceptor pairs in non-fullerene organic solar cells. NPJ Comput. Mater. 6, 120.
33. Zhan, L., Li, S., Lau, T.-K., Cui, Y., Lu, X., Shi, M., Li, C.-Z., Li, H., Hou, J., and Chen, H. (2020). Over 17% efficiency ternary organic solar cells enabled by two non-fullerene acceptors working in an alloy-like model. Energy Environ. Sci. 13, 635–645.

34. Du, X., Yuan, Y., Zhou, L., Lin, H., Zheng, C., Luo, J., Chen, Z., Tao, S., and Liao, L.-S. (2020). Delayed fluorescence emitter enables near 17% efficiency ternary organic solar cells with enhanced storage stability and reduced recombination. Energy Loss 30, 1909837.

35. Nian, L., Kan, Y., Wang, H., Gao, K., Xing, B., Rong, Q., Wang, R., Wang, J., Liu, F., Chen, J., et al. (2018). Ternary non-fullerene polymer solar cells with 13.51% efficiency and a record-high fill factor of 78.13%. Energy Environ. Sci. 11, 3392–3399.

36. Liu, J., Chen, S., Qian, D., Gautam, B., Yang, G., Zhao, J., Bergqvist, J., Zhang, F., Ma, W., Ade, H., et al. (2016). Fast charge separation in a non-fullerene organic solar cell with a small driving force. Nat. Energy 1, 16089.

37. Lee, J., Lee, S.M., Chen, S., Kumari, T., Kang, S.-H., Cho, Y., and Yang, C. (2018). Organic photovoltaics with multiple donor-acceptor pairs. Adv. Mater. 1804762. https://doi.org/10.1002/adma.201804762.

38. Bi, P., and Hao, X. (2019). Versatile ternary approach for novel organic solar cells: a review. Sol. RRL 3, 1800263.

39. Marcus, R.A. (1993). Electron transfer reactions in chemistry: theory and experiment (Nobel Lecture). Rev. Mod. Phys. 32, 1111–1121.

40. Marcus, R.A. (1989). Relation between charge transfer absorption and fluorescence spectra and the inverted region 93, 3078–3086.

41. Benduhn, J., Tvingstedt, K., Piersimoni, F., Ulbrich, S., Fan, Y., Tropiano, M., McGarry, K.A., Zeika, O., Riede, M.K., Douglas, C.J., et al. (2017). Intrinsic non-radiative voltage losses in fullerene-based organic solar cells. Nat. Energy 2, 17053.

42. Karki, A., Vollbrecht, J., Dixon, A.L., Schopp, N., Schrock, M., Reddy, G.N.M., and Nguyen, T.Q. (2019). Understanding the high performance of over 15% efficiency in single-junction bulk heterojunction organic solar cells. Adv. Mater. 31, e1903868.

43. Azzouzi, M., Yan, J., Kirchhart, T., Liu, K., Wang, J., Wu, H., and Nelson, J. (2018). Nonradiative energy losses in bulk-heterojunction organic photovoltaics. Phys. Rev. X 8, 031055.

44. Xie, Y., Wang, W., Huang, W., Lin, F., Li, T., Liu, S., Zhan, X., Liang, Y., Gao, C., Wu, H., and Cao, Y. (2019). Assessing the energy offset at the electron donor/acceptor interface in organic solar cells through radiative efficiency measurements. Energy Environ. Sci. 12, 3556–3566.

45. Bento, A.P., Gaulton, A., Hersey, A., Bellis, L.J., Chambers, J., Davies, M., et al. (2002). Classification and regression by randomForest. R News 23, 18–22.

46. Shavlik, J.W., Mooney, R.J., and Towell, G.G. (1991). Symbolic and neural learning algorithms: an experimental comparison. Mach. Learn. 6, 111–143.

47. Diaz-Uriarte, R., and Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7, 3.

48. Cao, B., Adutwum, L.A., Olinsky, A.O., Luber, E.J., Olsen, B.C., Mar, A., and Buriak, J.M. (2018). How to optimize materials and devices via design of experiments and machine learning: demonstration using organic photovoltaics. ACS Nano. 12, 7434–7444.

49. Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Inf. Process. Manage. 45, 427–437.

50. Otten, J.D., Karssemeijer, N., Hendriks, J.H., Groenewoud, J.H., Fracheboud, J., Verbeek, A.L., de Koning, H.J., and Holland, R. (2005). Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. J. Natl. Cancer Inst. 97, 748–754.

51. William, H.W. (2009). Google scholar search performance: comparative recall and precision. Portal: Libraries and the Academy 9, 5–24.

52. Jin, H., and Ling, C.X. (2005). Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowledge Data Eng. 17, 299–310.

53. Brzezinski, D., and Stefanowski, J. (2017). Prequential AUC: properties of the area under the ROC curve for data streams with concept drift. Knowledge Inf. Syst. 52, 531–562.

54. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2012). Scikit-learn: machine learning in python. J. Mach. Learn. 12, 2825–2830.

55. Hanley, J.A., and McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve 143, 29–36.