




Cite this: *Phys. Chem. Chem. Phys.*,  
2025, 27, 21719

# Screening of potential candidates for solid electrolyte interphase materials for lithium-ion batteries through a data-driven approach

Sadhana Barman and Utpal Sarkar \*

Material property prediction through machine learning has emerged as a revolutionary approach for diminishing hardships in the design of optimal materials for practical applications. Herein, we used a machine learning approach to refine over 11664 solid electrolyte interphase materials and identified potential candidates in terms of chemical stability at the molecular level, solvation energy and ease of synthesis, thereby obtaining insights for discoveries of new effective optimal interphase materials for lithium-ion batteries. The predicted accuracy of chemical reactivity parameters and solvation energy was in the range of 86.7–91.3% by uncovering atomistic input features. Dipole moments, number of heteroatoms, NHOH count, heavy atom count, number of hydrogen acceptors and donors, several surface area descriptors (PEOE\_VSA1, PEOE\_VSA4, SMR\_VSA6, SMR\_VSA10, EState\_VSA10, VSA\_EState1, VSA\_EState2), kappa index (kappa1), and functional groups (fr\_ketone, fr\_alkyl\_halide, fr\_nitro), etc. have been identified as key factors influencing solvation energy and chemical reactivity, offering critical guidance for screening the materials. These insights enable the strategic selection of SEI materials with chemical stabilities that effectively impact dendrite formation, thereby having the potential to enhance the performance and longevity of electrochemical systems. For the ideal identified candidates that have solid electrolyte interphase-affecting characteristics, the predicted property values perfectly align with the actual values. The predicted solvation energy, chemical hardness, and electrophilicity index are in the ranges of 1.433–5.677 kcal mol<sup>-1</sup>, 10.796–17.530, and 0.270–0.390, respectively, along with a low synthetic accessible score of 1.219–2.260. Non-ideal materials with the predicted solvation energy, chemical hardness, electrophilicity index and synthetic accessibility score are in the ranges of 85.354–300.982 kcal mol<sup>-1</sup>, 1.820–4.005, 2.030–4.823, and 4.002–7.422, respectively, demonstrating the model's robustness for reliable prediction, along with poor solid electrolyte interphase-suppressing characteristics. The most intriguing feature of our work is the molecules containing the elements fluorine, nitrogen and carbon, which define stable SEI candidates, while sulphur, oxygen, nitrogen, and carbon-containing molecules reduce the stable SEI formation capability. This result highlights a robust workflow that can guide the future discovery of materials through property optimization, particularly for dendrite suppression.

Received 17th July 2025,  
Accepted 16th September 2025

DOI: 10.1039/d5cp02726h

[rsc.li/pccp](http://rsc.li/pccp)

## Introduction

The utilization of lithium-ion batteries (LIBs) in electrical systems such as electric vehicles has attracted much attention<sup>1–12</sup> due to their decreased cost, enhanced energy storage capacity, values of volumetric and gravimetric energy density than other storage devices, as well as increased efficiency. The accelerated application of LIBs in automotive fields is because of their rapid production rate and the proliferation of electric automobiles that has been forecasted to lead the future market of energy

devices.<sup>13,14</sup> However, cloudy weather, short life cycles, safety risks of liquid electrolytes that cause leakage, the generation of gas, explosions and limited transportability hinder their application.<sup>14–17</sup> Interfacial dynamics are greatly involved in their electron-ion transfer phenomena for finding optimal electrochemical devices for use in energy storage applications.<sup>14</sup> The solid electrolyte interphase (SEI) is considered a crucial component of modern lithium-ion batteries that exhibits a nanoscale film (10–50 nm thick) and is constructed by the decomposition of electrolyte between the Helmholtz double layers at the anode. During the charging and discharging cycles of the battery, the SEI causes irreversible capacity loss by electrolyte degradation. It can be cured through the limited

Department of Physics, Assam University, Silchar-788011, Assam, India.  
E-mail: [utpalchemiitkgp@yahoo.com](mailto:utpalchemiitkgp@yahoo.com)

formation of the SEI, which conducts Li ions.<sup>18–20</sup> The SEI formation process is a complex one since different lengths and time scales are involved in the processes responsible for SEI formation. The detection of reactive intermediates during the formation of the SEI is still a challenging task through experimental study due to its extreme complexity. To accurately capture this phenomenon in an experiment throughout the length and time scale, which is necessary to develop the dataset for a machine learning study, is very difficult. The other option is simulation, which can mimic the experimental conditions, but due to the involvement of several interactions, along with the consideration of the length and time scale involved, the simulation studies (the molecular properties under experimental conditions) are also limited, and there is a scarcity of data needed for the machine learning approach.

The practical performance of SEI layers in lithium-ion batteries depends on their solvation structure and also on a wide range of electrochemical and physical properties, such as interfacial energy with electrodes, mechanical robustness, lithium-ion diffusivity, and low electronic conductivity. The role of SEI in LIB battery performance is still an interesting topic for the scientific community. However, density functional theory (DFT)<sup>21,22</sup> has appeared as an alternative to analyse these reactive molecules that compensate for the void and provide a better understanding of the reaction chemistry of the SEI. The reactivity parameters like electronegativity, chemical hardness, electrophilicity, *etc.*,<sup>23–28</sup> are used for the chemical characterization of molecules within the DFT framework.

Solvation energy is a critical parameter for battery applications that determines the interaction between solute and solvent molecules. It is a physical property that measures the amount of minimum necessary work required for solvation<sup>29–34</sup> and describes molecular interactions that cover the interactive nature of the solvent in contact with a dissolved molecule, in which the solute and solvent organize. Solvation energy signifies the deformation of free energy that is correlated with the transportation of molecules between the solvent, as well as in an ideal gas at a certain pressure and temperature.<sup>35–39</sup> Forecasting this thermodynamic property has always been a bit challenging, but it has been investigated through *in silico* computational methods for complicated hydration mechanisms.<sup>40–53</sup> This property has extended its applicational domain into various chemical processes of drug delivery systems,<sup>54,55</sup> sustainable synthesis methods,<sup>56</sup> as well as the electrochemical performance of energy storage devices.<sup>57,58</sup> Moreover, its accuracy has been hindered due to a lack of sufficient experimental data despite several breakthroughs. To date, various reliable techniques, *e.g.*, molecular dynamics, quantum chemical simulation, *etc.*, have been used to predict this solvation energy.<sup>59–67</sup> However, some hydration complexity has been explained using the Boltzmann equation, which describes the behaviour of solvent in an isotropic medium.<sup>68–70</sup>

While describing the solvation model, only small molecules are considered at the quantum level during calculations, and complicated examples are underrepresented. Molecular dynamics facilitates a significant understanding of solvation energy, typically

near 4 kcal mol<sup>-1</sup>.<sup>71</sup> Forecasting physical properties provides the possibility for screening for optimal material design that also gives feasibility for synthetic route design, *etc.*<sup>72</sup>

Data-driven approaches, which are considered the fourth paradigm of science, have been proven to be an efficient method for predicting physical properties and are also cost-effective and time-saving.<sup>73–82</sup> These approaches facilitate finding the most relevant features influential for solvation energy. Solvation energy is predicted using a graph neural network, and further transfer learning is implemented on experimentally calculated solution energy datasets.<sup>83,84</sup> However, the range of the target property value found experimentally is usually small and is a limitation of the work since the variation in the property value is also small; consequently, it is relatively easy to predict the target property compared to the case where the variation of the target property is large.

Herein, we have implemented automated machine learning (ML) models to predict the solvation free energy of SEI products through supervised machine learning algorithms. Instead of using calculated features as input, we have developed machine-generated atomistic input features using the Rdkit module,<sup>85</sup> and ML models provide the correlations between these features and solvation energy. We have also narrowed down our entries to find the optimal structures that can be synthesized. We have selected chemical reactivity parameters, namely electronegativity, chemical hardness, and electrophilicity, apart from solvation energy and the synthetic accessibility score, as screening criteria for SEI-forming materials.

## Computational details

### Collection of data

We considered the Li-ion battery electrolyte dataset (LIBE),<sup>86</sup> which includes combined counterparts of principal molecules that contain solvents, salts, and SEI products. Selectively breaking and reforming the bonds of some previously proposed materials involved in SEI formation enables the creation of novel structures for the LIBE dataset. These newly designed products are subjected to recombination and fragmentation during their construction, and are associated with some already known electrolyte components. Since the free energy is included in the dataset, it is expected that it can capture the reactive chemistry of the SEI.<sup>86</sup> The computational validation of reactive organometallic molecules is quite a challenging task. This dataset has been constructed to facilitate the identification of the reaction pathway for the SEI formation mechanism. LIBE includes non-polymeric and non-oligomeric molecules, which can be categorised into solvent molecules, salt molecules, inorganic SEI products, possible dissolved minority species, including gases, lithium ethylene dicarbonate (LEDC) and related derivatives, lithium butylene dicarbonate (LBDC) and related derivatives, lithium ethylene monocarbonate (LEMC) and related derivatives, ethanol and related derivatives, ethylene glycol (EG) and related derivatives, 1,4-butanediol and related derivatives, other molecules related to

LiEC decomposition, and other molecules related to PF6 decomposition relevant to LIBs.<sup>86</sup> The considered structures have been computed with the  $\omega$ B97X-TZVPPD/SMD// $\omega$ B97X-V/def2-TZVPPD/SMD level containing different charges, as well as spin multiplicity. To train our models by implementing machine learning techniques, we have used almost 11 664 datapoints from the LIBE dataset to predict solvation energy. Each SEI formation material present in the dataset has been represented with SMILES strings whose actual solvation energy values span from 1.381 to 323.507 kcal mol<sup>-1</sup>.

### Feature selection

The selection of important features for property prediction through ML is important to interpret the correlation with the targeted property; consequently, it becomes crucial for its reduction, which is done by analysing its Pearson correlation coefficient values. For the prediction of solvation energy, we have extracted about 200 input features from the Rdkit module,<sup>85</sup> and chose important features whose correlation values with solvation energy are greater than 0.10, which reduces the number of input features to a correlated one.

### Machine learning models

We have used multiple algorithms that originate from the desire to explore a diverse range of modelling techniques. Each algorithm has its strengths and weaknesses, and by employing a variety of approaches, we aimed to capture different aspects of the data and potentially uncover subtle patterns that might not be apparent with a single method. Furthermore, utilizing multiple algorithms allows us to assess the robustness of our findings and compare the performance of different models. By evaluating various machine learning techniques, we can identify which algorithms are best suited for our specific dataset, and this comprehensive analysis enables us to make more informed decisions and gain confidence in our results. We intend to leverage ensemble learning techniques by combining the predictions of multiple models. Ensemble methods, such as extreme gradient boosting, extra tree regressor, *etc.*, often yield superior performance by aggregating the outputs of individual models. This collaborative approach enhances prediction accuracy and generalizability, contributing to more reliable results. It is worth mentioning that our analysis involved a comprehensive exploration of various machine learning algorithms. However, this study primarily focuses on presenting the most promising results obtained from this extensive analysis.

### Workflow

The machine learning workflow began with the collection of SEI-forming molecules represented by their SMILES format. We proceeded with feature engineering by extracting several molecular descriptors as input features using cheminformatics tools. After extracting a large set of diverse features, exploratory data analysis (EDA)<sup>87</sup> was implemented to refine the abnormalities from the parent dataset. After getting the refined dataset, correlation studies were carried out to detect the correlation

between input and target features, which resulted in the reduction of input features to highly correlated ones. The dataset is now ready for prediction by splitting it into training (80%) and testing (20%) sets with several machine learning models. Hyperparameter tuning was done to improve the model's performance, which resulted in improved performance in prediction. Finally, multiobjective optimization was done after classifying the whole dataset into ideal and non-ideal sets to select the best candidate molecules from its non-dominated solutions by achieving an optimal balance across multiple desirable properties. The workflow is presented in Fig. 1.

## Results and discussion

Herein, we discuss the step-by-step procedure to identify suitable and unsuitable solid electrolyte interphase materials from a large dataset, based upon the chemical reactivity, solvation energy, and synthetic accessibility (SA) score. To predict the solvation energy along with the chemical reactivity parameters, our workflow of the ML approach considers 11 664 entries in the initial database. Significant atomistic information, in the form of molecular descriptors, representing the physicochemical characteristics of each material, has been decoded, and more than 200 input features that are significantly correlated with the chemical reactivities and solvation energy were considered. We selected (by sorting) highly correlated features following refinement of the dataset for better accuracy by hyperparameter tuning, which resulted in the enhanced performance of the models. Through this workflow of predicting these properties, machine learning models provided us with some highly impactful input features that have prominent contributions to the targeted properties. We separated the materials having solvation energy values less than 50 kcal mol<sup>-1</sup> (6134 entries) and greater than 50 kcal mol<sup>-1</sup> (5530 entries), classifying them as 'desired for SEI' and 'unwanted for SEI', respectively. Finally, we generated the SA score of these materials and implemented 'material screening' by optimizing the chemical hardness, electrophilicity index, solvation energy, and SA score through the Pareto filter method, which gives us the best candidate material.

### Property prediction

The unique properties of each molecule are characterized by its structural, chemical, as well as topological properties, known as molecular descriptors. The appropriate correlation of the solvation energy with the extracted molecular descriptors has been conducted, and the findings are presented in the correlation heatmap plotted in Fig. 2(a).

The correlation coefficients, governed by the associated correlation matrix of solvation energy with the top descriptors and reactivity parameters, are distinguished by the different colours in the heat map. The more vibrant colours of the boxes, towards red, signify a more intense correlation, whereas the colour changes towards blue indicate a low correlation between the features. The diagonal boxes featuring the brightest red

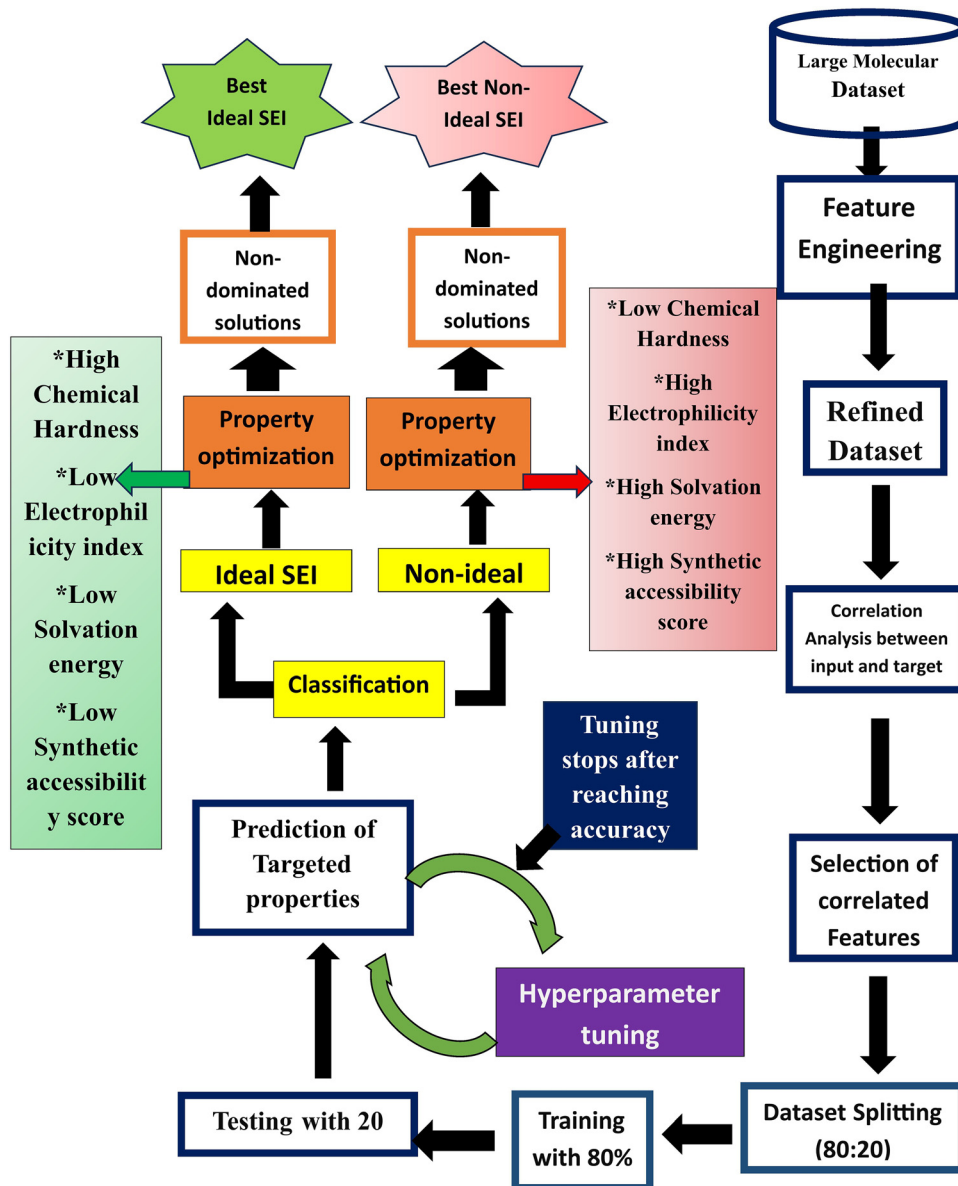


Fig. 1 Workflow for screening ideal and non-ideal SEI materials.

colour symbolize its autocorrelation. Essential molecular descriptors (Table S1) for solvation energy come into focus because of the correlation coefficients obtained through this (correlation heatmap) effective tool. Necessary correlation coefficients for our study are observed to fall within the range of  $-0.34$  to  $0.77$ , related to the relevant descriptors: “electronegativity”, “chemical hardness”, “electrophilicity index”, “topological polar surface area (TPSA)”, “dipole moment”, “partial equalization of orbital electronegativity and surface area contribution of atoms in molecule (PEOE\_VSA1)”, “MOE logP VSA descriptor 2 (SlogP\_VSA2)”, “heavy atom count”, “number of NHs and OHs (NHOH count)”, “number of nitrogen and oxygens (NO count)”, “number of hydrogen bond acceptors (Num H acceptors)”, “number of hydrogen bond donors (Num H donors)”, “number of hetero atoms (Num hetero atoms)”.

Fig. 2(b) portrays the highest ten input features along with chemical reactivity parameters chosen from the Pearson correlation heatmap, Fig. 2(a). The correlation analysis shows that chemical reactivity parameters are nonlinearly correlated with solvation energy, signified by their correlation coefficient values,  $0.14$ ,  $0.24$ , and  $-0.34$ , which ensures a reliable context for screening materials by the multiobjective optimization of these properties. The negative value of the correlation coefficient for chemical hardness ensures the inverse correlation with solvation energy. Three descriptors, “TPSA”, “NO count”, and “Num H acceptors” show the correlation values of  $0.77$ ,  $0.75$ , and  $0.71$ , respectively. In comparison, “dipole moment” and “Num hetero atoms” exhibit lower correlation than the three descriptors above, with correlation values of  $0.69$  and  $0.65$ . “NHOH count”, “Num H donors”, “PEOE\_VSA1”,

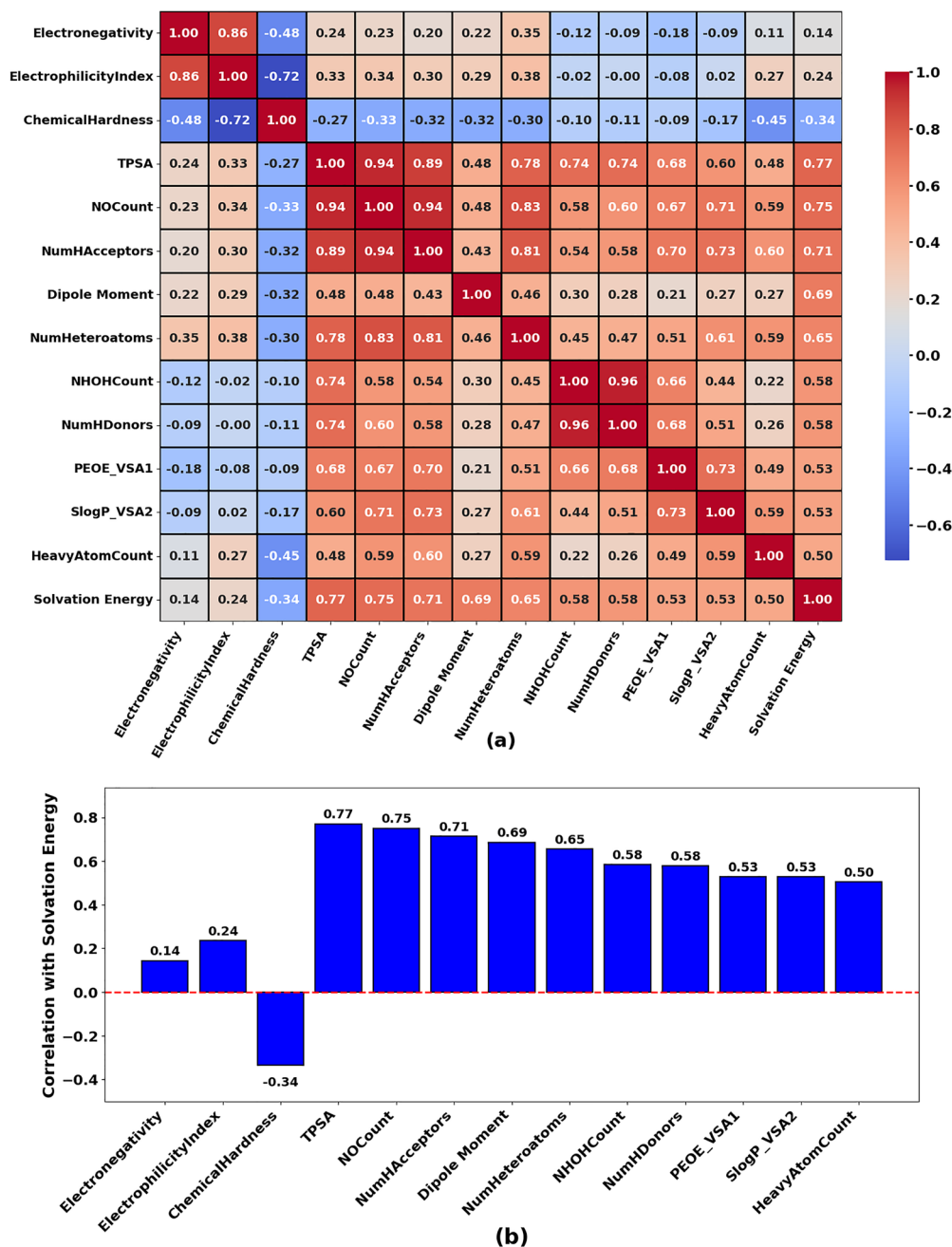


Fig. 2 (a) Correlation heatmap and (b) bar graph of correlation coefficients.

“SlogP\_VSA2”, and “heavy atom count” show correlation with solvation energy in the range between 0.58 and 0.50. In Fig. 3, we visualise the correlation pattern of individual input features with solvation energy; the thirteen scatter plots confirm the nonlinear relationships between the individual features and the target feature (solvation energy). Plots a, b, c, d, g, l, and m of Fig. 3 show the variation of solvation energy with respect to these input features; the remaining plots, e, f, g, h, i, j, and k, demonstrate the discrete relationship pattern between solvation energy and the corresponding input features. These individual pairwise plots help to assess the predictive and optimization relevance between each feature.

Interestingly, an inverse relationship between the solvation energy and chemical hardness (Fig. 3c) indicates that as desolvation increases, the molecule’s reactivity increases, and molecules become more polarised (Fig. 3d, g and l), as indicated by the direct correlation of the solvation energy with dipole moment and surface area descriptors. Such information guides us to find the appropriate parameters for prediction and optimization through its direct and inverse underlying relations. Based on these, we proceeded with the prediction, taking electronegativity, chemical hardness, electrophilicity, and solvation energy as targets and the rest of the features as inputs.

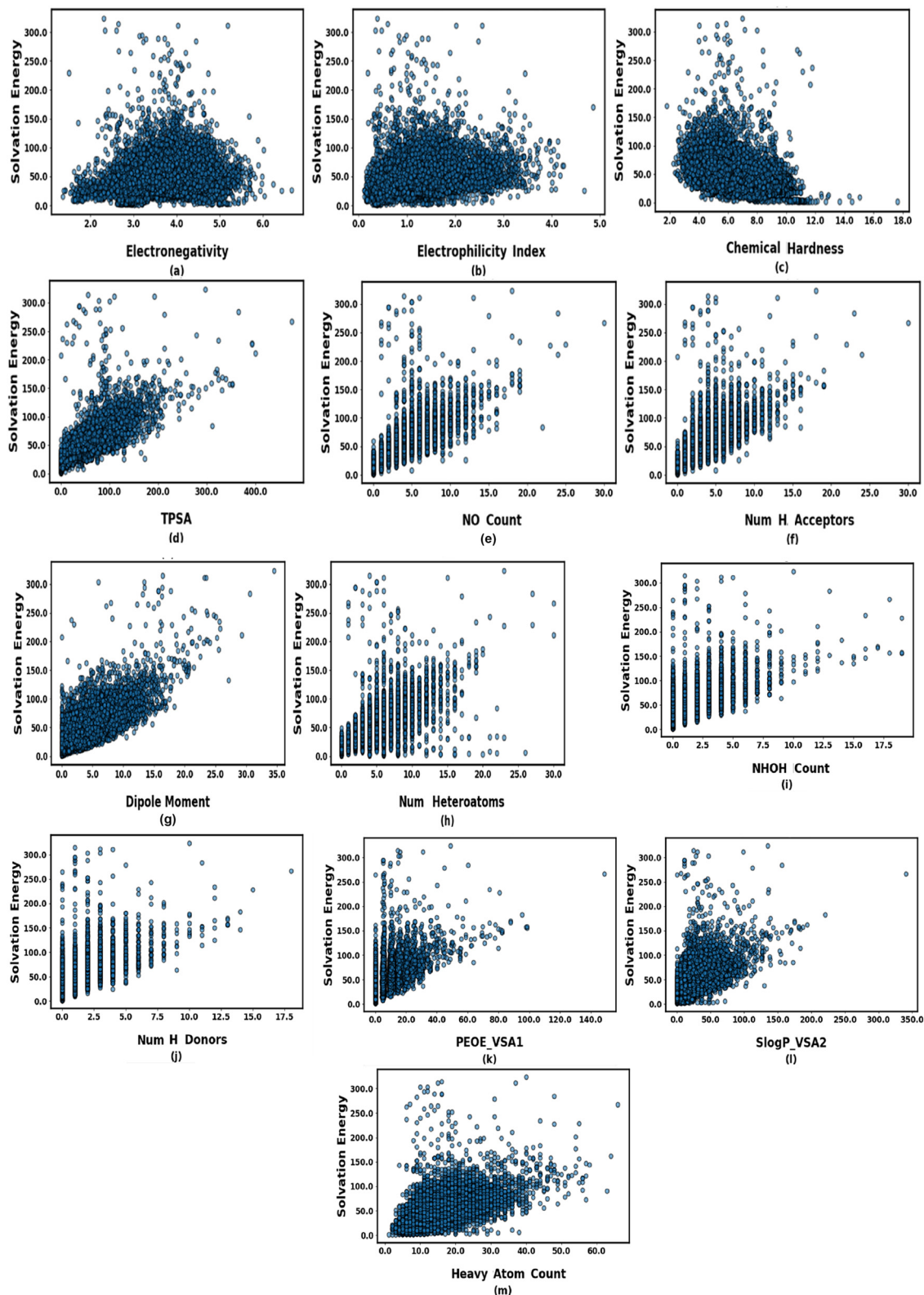


Fig. 3 Scatter plots between input features and solvation energy.

Regression algorithms have been implemented for the accurate prediction of all properties through machine learning models. In our investigation, we use 80% of the dataset for

training, and the remaining 20% of the data for testing. A total of forty-two lazy predict<sup>88</sup> ML models have been tested (Table S2), depending upon the *R*-squared value. Here, we present twenty-two

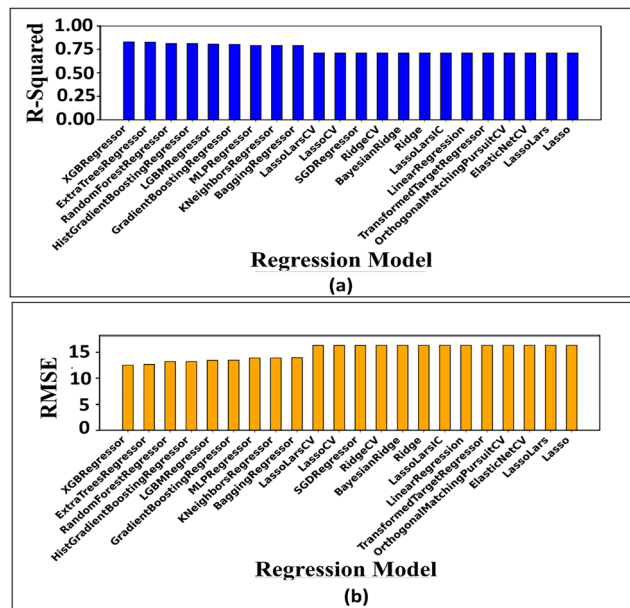


Fig. 4 Comparison of (a)  $R$ -squared and (b) RMSE for twenty-two ML models.

ML models (Fig. 4) for predicting solvation energy, through which performance has been scrutinized.

Model performance has been examined based on its high coefficient of determination ( $R$ -squared) and low root mean squared error (RMSE) values. As depicted in Fig. 4, the performances of the top twenty-two models are demonstrated in Table 1.

Fig. 4a and b indicate that the highest performing ML model is the XGB regressor model, with an  $R$ -squared value of 0.83 and the lowest RMSE value of 12.53. Extra tree regressor also showed superior performance: it achieved an  $R$ -squared value of 0.83 and an RMSE value of 12.68 (Table 1). The four models,

Table 1 Outcomes of different ML regressor models

Serial number	Model	$R$ -squared	RMSE
1	XGB regressor	0.83	12.53
2	Extra trees regressor	0.83	12.68
3	Random forest regressor	0.81	13.20
4	Hist gradient boosting regressor	0.81	13.21
5	LGBM regressor	0.81	13.44
6	Gradient boosting regressor	0.80	13.46
7	MLP regressor	0.79	13.92
8	K neighbors regressor	0.79	13.92
9	Bagging regressor	0.79	13.97
10	Lasso lars CV	0.71	16.34
11	Lasso CV	0.71	16.34
12	SGD regressor	0.71	16.34
13	Ridge CV	0.71	16.35
14	Bayesian ridge	0.71	16.35
15	Ridge	0.71	16.35
16	Lasso lars IC	0.71	16.35
17	Linear regression	0.71	16.35
18	Transformed target regressor	0.71	16.35
19	Orthogonal matching pursuit CV	0.71	16.35
20	Elastic net CV	0.71	16.35
21	Lasso lars	0.71	16.37
22	Lasso	0.71	16.37

*i.e.*, random forest regressor, hist gradient boosting regressor, LGBM regressor, and gradient boosting regressor, appear to capture more diverse datapoints with consistently high  $R$ -squared values with only a slight difference from the highest performing extra trees regressor model and XGB model, which reflects strong predictive power. The  $R$ -squared values of the four models lie in the range of 0.80–0.81, with RMSE values ranging from 13.46–13.20. However, for models with serial numbers 7 to 22 (Table 1), a decline in the  $R$ -squared values was observed. This slight decline indicates the effectiveness of the first six models compared to the rest of the ML models (MLP regressor, K neighbors regressor, bagging regressor, Lasso lars CV, Lasso CV, SGD regressor, ridge CV, Bayesian ridge, ridge, Lasso lars IC, linear regression, transformed target regressor, orthogonal matching pursuit, elastic net CV, Lasso lars, Lasso) for which  $R$ -squared and RMSE cover the ranges 0.71–0.79 and 16.37–13.92, respectively. It manifests that these models capture almost the same data pattern among the input features and the target properties; this trend also demonstrates the reduced predictive power.

Overall, the extra trees regressor and XGB regressor models performed as the top-performing models among twenty-two ML models for solvation energy. Taking this into account, we proceeded to predict the electronegativity, chemical hardness, electrophilicity index, along with solvation energy, using these two highest performing models. Here, we have shown the prediction plot for the XGB regressor and extra trees regressor models (Fig. 5). The accumulation of datapoints, as shown in Fig. 5 near the diagonal line, signifies favourable correspondence between the predicted and actual values for all predicted properties. Fig. 6 shows the variation of residuals plot and the distribution of these residuals in a histogram plot (for XGB regressor), which clarifies the presence of some datapoints that have very few residual errors seen a little way off from the best fitting line.

The predicted chemical reactivity and solvation energy with the best performing XGB regressor model and extra trees regressor model have achieved  $R$ -squared values in the range 0.867–0.913, mean absolute error (MAE) values in the range 0.152–4.849, and RMSE values in the 0.236–10.047 range, as shown in Fig. 5. The distribution of the residuals of these predicted properties is displayed in Fig. 6. The better alignment of the tested datapoints towards the best fitting line can be interpreted as its enhanced predictive power. This prediction enhanced our area of search for optimal SEI products with better predictability of all targeted properties.

The highly contributing input features in predicting all properties for the dataset were detected according to their correlation coefficients, as shown in Fig. 7. The values signify a better correlation between input and output and dictate the accuracy in prediction performance.

The best important features (using XGB regressor) were arranged according to their importance value, as shown in Fig. 7, where some specific surface area descriptors, functional groups, are seen to have the highest importance in predicting these properties. “Functional group nitro (fr\_nitro)” was

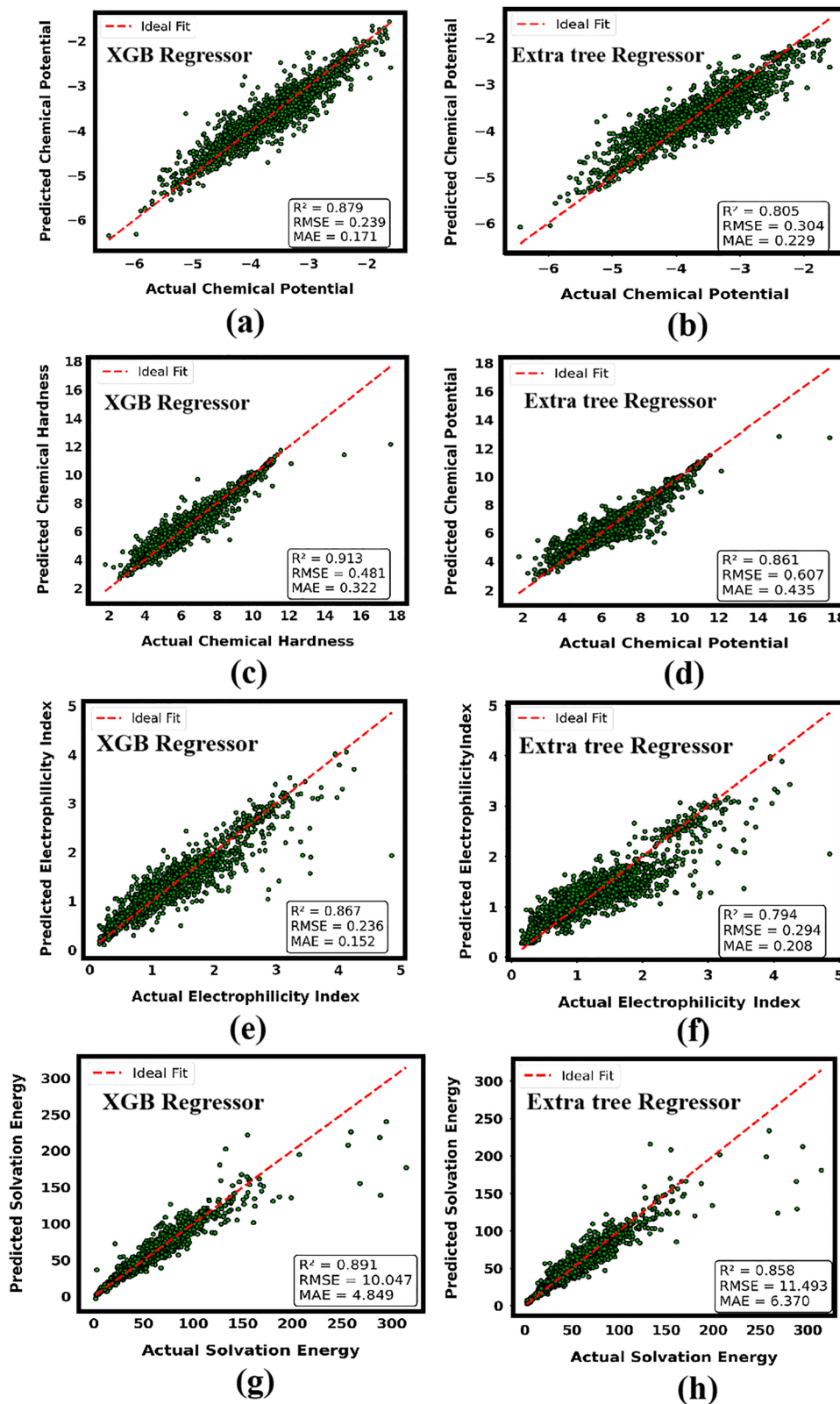


Fig. 5 Prediction of solvation energy through XGB and extra trees regressor.

observed to possess the most important feature for the electro-negativity and electrophilicity index, with importance values of

0.3242 and 0.6010, respectively. “Ring count” and “dipole moment” are the highest importance features for chemical

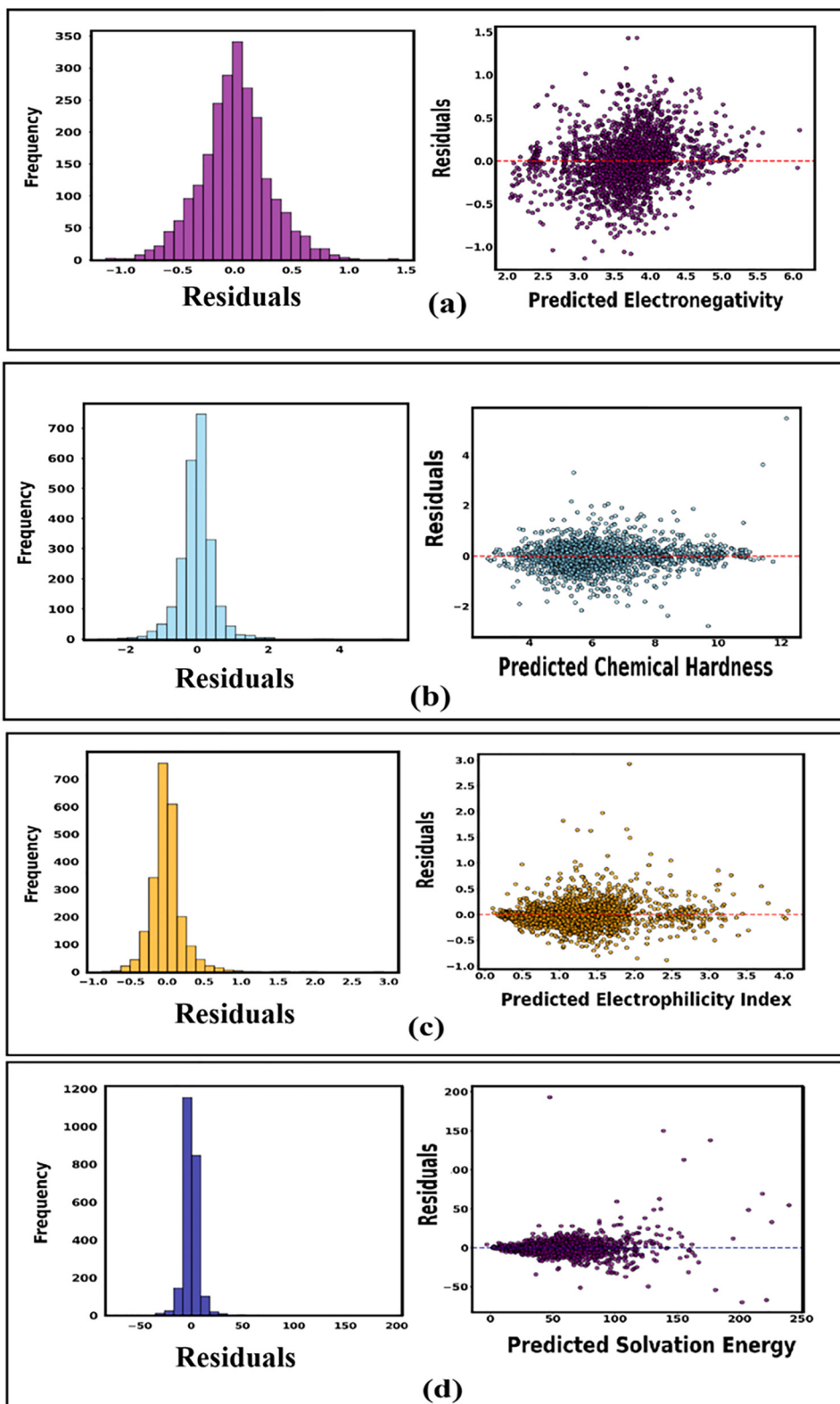


Fig. 6 Residuals and histogram plots for all predicted properties through the XGB regressor.

hardness and solvation energy, with importance values of 0.7318 and 0.2589, respectively. Surface area descriptor “MaxEState-Index”, SMR\_VSA (SMR\_VSA6, SMR\_VSA10), and VSA\_EState

(VSA\_EState1, VSA\_EState2) series contribute to the prediction of both the electronegativity and electrophilicity indexes. “Functional group ketone (fr\_ketone)” and the “kappa1” index

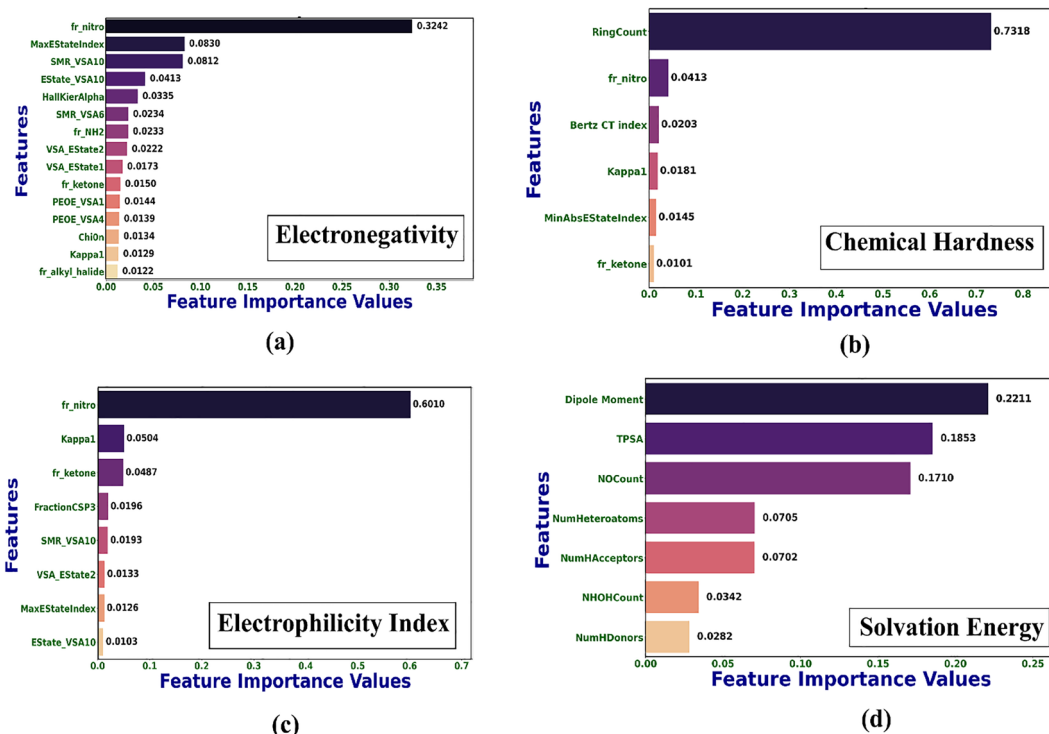


Fig. 7 Feature importance plots for all the predicted properties through the XGB regressor.

participate as the important features for all reactivity parameters, with importance values of 0.0150 and 0.0129 for electronegativity, 0.0101 and 0.0181 for chemical hardness, and 0.0487 and 0.0504 for electrophilicity index. Two other important functional groups for electronegativity prediction chosen by XGB regressor are “functional group of NH<sub>2</sub> (`fr_NH2`)” and “functional group of alkyl halide (`fr_alkyl_halide`)”.

The “TPSA” contribution was observed, with a value of 0.1853 for solvation energy. Feature importance values of “NO count”, “number H acceptor”, “number hetero atoms”, “number atom count”, “number H donors”, “NHOH count”, “SlogP\_VSA2”, and “PEOE\_VSA1” for solvation energy were 0.1710, 0.0705, 0.0702, 0.0342, and 0.0282, respectively.

Surface area descriptors are directly correlated with the interactions that influence the chemical reactivity parameters and solvation energy from our study. The surface chemistry of the electrolyte impacts the kinetic stability of the electrolyte; to incorporate it, we considered correlated surface area descriptors as input features, apart from the already mentioned reactivity parameters of the individual molecule. This guides the understanding of the electrical (electronic insulation capability) properties, as well as the reactivity of the SEI, which, to some extent, helps with comprehending the electrode–electrolyte structure property relationship. This comprehension may lead to the basis of interphase formation.

The “TPSA” quantifies the polarity, as well as the potential hydrogen bonding capacity of a molecule, depending on the distribution of polar atoms in the molecule. Dipole–dipole interactions between a polar solvent and solute greatly affect

the solvation energy value, resulting in a more stable structure due to the higher solvation energy. Dipole moments also quantify how polar the molecule is. Polarization plays a critical role in regulating the formation and stability of the solid electrolyte interphase (SEI) in lithium-ion batteries. Under the influence of an internal electric field, materials with high dielectric constants—such as certain separators or electrolyte components—undergo strong electron displacement polarization. This polarization modifies the local electrostatic environment at the electrode–electrolyte interface, influencing the distribution and mobility of lithium ions and coordinating species. As a result, polarization can alter the solvation structure of Li<sup>+</sup>, favouring the inclusion of anions in the solvation sheath, which leads to the formation of anion-derived SEI components such as LiF. These inorganic-rich SEI layers are typically more uniform, mechanically stable, and ionically conductive. Moreover, the enhanced polarization can suppress side reactions by reducing local electron density near the electrode surface, thereby mitigating the formation of amorphous organic oligomers and promoting a compact, low-resistance interphase. Overall, polarization-induced tuning of the interfacial environment emerges as a powerful mechanism for optimizing SEI chemistry and improving battery performance.<sup>89</sup>

ML models, considered as black boxes used for predictions, have been interpreted by the easily understandable Shapley additive explanations (SHAP)<sup>90</sup> method in our study. Since we are emphasizing the solvation effect, in this method, the contribution of the input feature for predicting solvation energy has been interpreted by assigning each highly important input

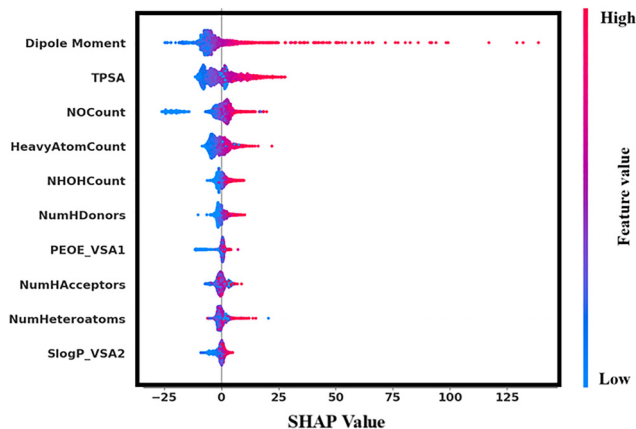


Fig. 8 SHAP plot for solvation energy through the extra trees regressor for refined datasets.

feature a numerical value. These numerical values signify the marginal participation of each feature, which can be chosen by different combinations, as well as contributions of features that influence the result and are constructed on the basis of the cooperative game theory principle.<sup>90</sup> Fig. 8 shows the SHAP plot of input features with predicted solvation energy.

In Fig. 8, the vertical axis represents the arrangement of important features from top to bottom based on their contribution to predicting solvation energy. The horizontal axis represents the impact of the contribution from positive to negative values by the SHAP algorithm. Red datapoints signify how

impactful these datapoints are in the prediction of the target, whereas blue points show datapoints that have a low contribution to the prediction. In our SHAP plot, the datapoints of each feature are distributed from negative to positive values. The red points show how many positive and negative points are participating in the prediction; this can help us understand the contribution of each feature. From our plot, we have seen that most of the positive values, as well as some negative values near 0, are red points that show their influence in prediction, *i.e.*, 'dipole moment', 'TPSA', 'NO count', 'heavy atom count', 'NHOH count', 'PEOE\_VSA1', 'SlogP\_VSA2'. Conversely, a negative impact on the target property also results from some positive values, as shown from the SHAP plot.

Partial dependence plots (PDP)<sup>91</sup> and individual conditional expectation (ICE) plots<sup>92</sup> have been demonstrated in Fig. 9 and 10, respectively, for important features with respect to their predicted solvation energy. The average contribution of input data points has been taken into consideration for the change in the target feature with the help of PDP (Fig. 9). The dependency of input features on the target response has been picturized through these plots, showing us linear and nonlinear relationships between the input and predicted property, and is helpful to have better information about the values of input features that are more influential to the target property. From PDP, it is clear that increasing 'TPSA' and 'dipole moment' values give rise to the linear enhancement of solvation energy. Conversely, the 'NO count', 'Num hetero atoms', 'Num H donor', 'NHOH count', 'Num H acceptors' plots manifest that after a certain

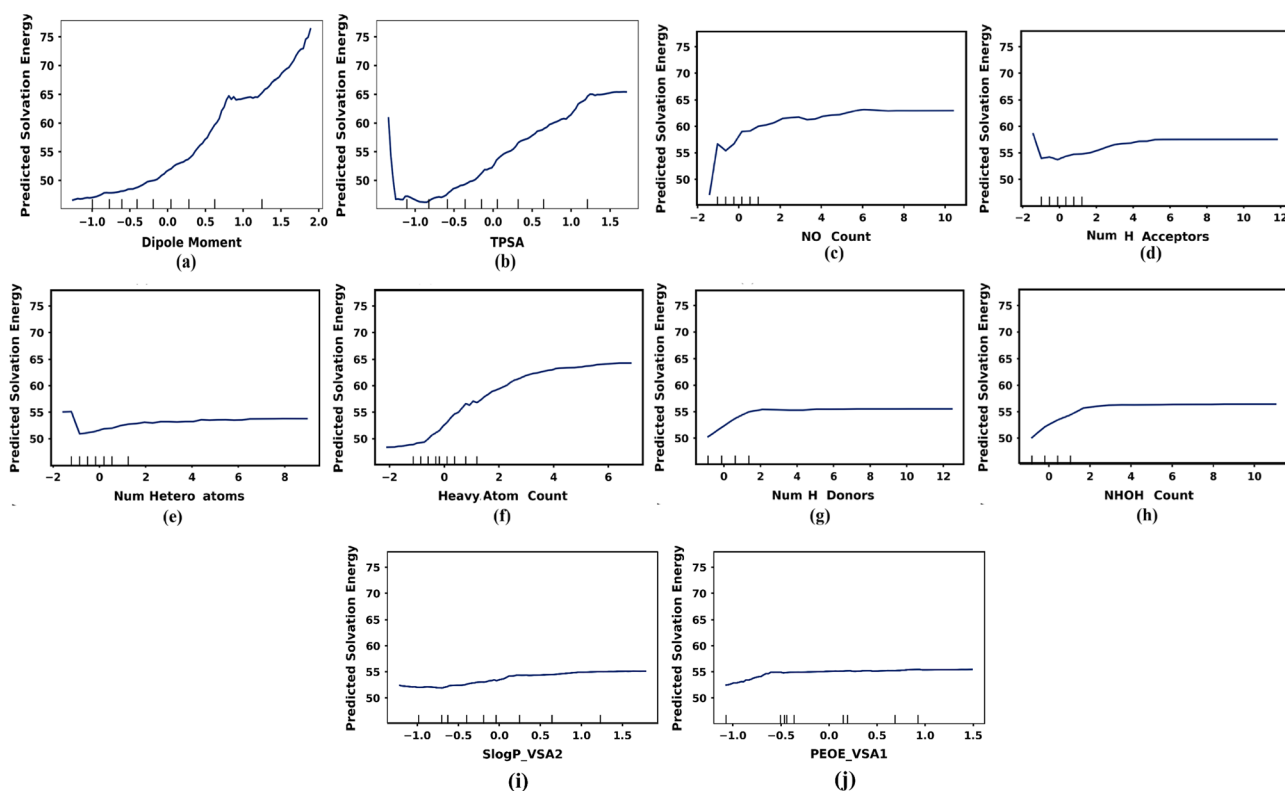


Fig. 9 PDP of solvation energy through the extra trees regressor for refined datasets.

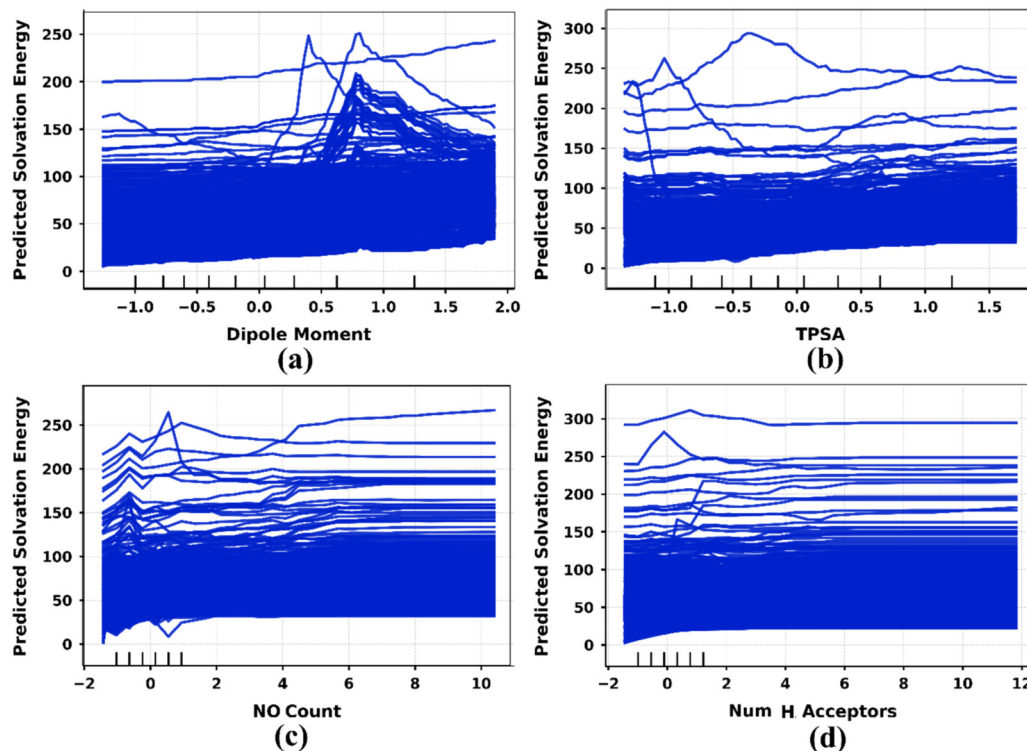


Fig. 10 ICE plots of solvation energy through the extra trees regressor for refined datasets, showing contributions for the top four features.

point, there is no increment of solvation energy with increasing values of these features. The ICE plot provides a better visualization and understanding of the dependency of the target feature on each value of the input features as well as how the target feature has been influenced by each value of the input features in each sample. We have displayed the top four feature plots with respect to solvation energy in Fig. 10, and the remaining features in Fig. S1. For the “TPSA” feature, we have seen (Fig. 9(b)) a linear relation with solvation energy (except for an initial sharp decline for a very short range). For the NO count, the solvation energy value increased rapidly until the NO count reached 0.5; after that, it became saturated with an increase in the NO count value. On the other hand, the third most influential feature, ‘Num H acceptors’, also exhibits non-linear behaviour with solvation energy and becomes saturated when the value is greater than or equal to 6.

The design of new SEIs that can reduce side reactions and enhance battery life and stability is crucial. Efficient SEI materials should be designed in such a way that they become chemically and mechanically stable, emphasize reaction prevention, dissolve in the electrolyte, do not break easily during volume expansion of the electrode, facilitate the interactions of electrons generated from the electrode and electrolyte, and prevent electrolyte reduction reactions that enable the smooth flow of Li-ions from the electrolyte to the electrode. The development of new SEI products can be accelerated through theoretical design of experimental efforts for potential candidates, which reduces the iterative refinement and time for evaluation from concept to practical implementation.<sup>93–96</sup>

Molecular-level information for an individual molecule is encoded in reactivity parameters like HOMO, LUMO, electronegativity, chemical hardness, and electrophilicity, which dictate the chemical stability and reactivity of the molecule. The electrode’s performance depends on the electrolyte’s solvation structure, which characterizes the alkali metal’s desolvation ability.<sup>97</sup> The solvation free energy of the electrolyte molecules, which is evaluated using the implicit solvent approach (polarizable continuum model, PCM, proposed by Tomasi and co-workers<sup>98</sup>), is included in the data set as an input feature. For electrochemical stability, the electrolyte should have the following: (a) good ionic conductivity and electronic insulating properties, which facilitate ion transport and minimal self-discharge; (b) a wide electrochemical window to prevent electrolyte degradation in the range of the working potential; and (c) chemical inertness with respect to the cell separator, electrode substrate, *etc.*<sup>99</sup> A multi-component system encompasses salt solvent additives that result in a large number of interactions, which introduces additional complexities.

The SEI layer, determined by the solvate, plays a significant role in battery cyclability. The stability of the electrode is dependent on the solvation structure of the electrolytes. The cathode, anode, and electrolyte composition also regulate the formation of the SEI layer, which should be mechanically strong and flexible to cope with the volume change (expansion/contraction) during charging/discharging.<sup>100</sup> To minimize the loss of capacity and ions, the SEI should be made of stable, insoluble, compact compounds to maintain high capacity, since the physical properties of the electrolyte depend on the

decomposed SEI components.<sup>101</sup> Higher ion diffusivity provides additional stability and conductivity, whereas lower ion diffusivity is associated with enhanced resistivity, *i.e.*, further blocking of ion movements, which results in capacity fading and low longevity/short life time of the battery, thereby controlling the battery performance.<sup>102</sup> The low electronic conductivity of the SEI layer reduces the battery performance because low electronic conductivity originates from high internal resistance, lowering of the charge–discharge rate that results in low coulombic efficiency,<sup>103</sup> and SEI heterogeneities that lead to reaction heterogeneity, uneven ion distribution, and dendrite growth;<sup>104</sup> however, high interfacial energy assists in stable ion deposition and suppresses dendrite formation.<sup>105</sup>

Understanding the solvation behaviour is hard for lithium-based materials as it is correlated with conductivity, stability, and reactivity. Therefore, to systematically analyse these effects on SEI, we proceeded to classify the data to enable more targeted findings by grouping properties in high and low solvation energies. In our investigation, the classification involved defining two distinct classes based on solvation energy: ideal SEI (having a solvation energy value lower than 50 kcal mol<sup>-1</sup>) and non-ideal SEI (having a solvation energy above 50 kcal mol<sup>-1</sup>), divided into two datasets containing 6134 (non-ideal SEI) and 5530 (ideal SEI) candidates, with potential implications for the SEI layer; the distribution is shown in Fig. 10. We focused on both practicability (in terms of accessibility score) and performance (in terms of solvation energy, chemical hardness, electronegativity, and electrophilicity index) to screen potential candidates. Specifically, to highlight the interpretability, we did not screen the molecules solely based on solvation energy but also by chemical hardness, electrophilicity, and synthetic accessibility score. This multiple selection criterion, which includes chemical reactivity parameters, reflects the relevance to electrochemical stability at the molecular level and the applicability of our investigation. It is worth noting that molecules having a high electrophilicity index and low chemical hardness refer to highly reactive species, whereas molecules with low electrophilicity index and high chemical hardness refer to relatively inert or chemically stable species, which are necessary for an ideal SEI.

Along with the above parameters, solvation energies that possess lower values are prone to preventing excess SEI layer formation (ideal SEI), and those with higher solvation energy values have non-ideal characteristics to form the SEI. Lower solvation energy-possessing SEI products will be useful in Li-ion battery materials, which are widely recognised for their capability to facilitate the formation of a stable SEI. Conversely, a high solvation energy value may offer ample opportunities for in-depth mechanistic investigations towards excess dendrite formation that degrades battery performance. Therefore, to design the SEI, the above properties should be characterized through structure–property relationships. In Fig. 10, we have displayed the distribution of solvation energies, indicating the frequency of occurrence in our dataset. Most of the values were found between 20 to 50 kcal mol<sup>-1</sup> (Fig. 10(a)), whereas most higher values were found between 20 to 25 kcal mol<sup>-1</sup>. However, for the second class of materials, *i.e.*, the solvation energy greater than 50 kcal mol<sup>-1</sup>, material distribution is skewed towards higher solvation energies (50–100 kcal mol<sup>-1</sup>), as depicted in Fig. 11(b).

Synthetic accessibility analysis<sup>106</sup> provides significant insight into the facile synthesis of the theoretically computed structure, and analysis for its potential synthetic route. It provides an idea for filtering the structures that can be efficiently synthesized in the laboratory, based on the synthetic accessibility value. The synthetic accessibility value indicates the complexity of synthesis. A lower SA score indicates great feasibility of the structures that can be synthesized experimentally (SA score = fragment score – complexity penalty). The calculation of the SA score is mainly based on the ‘synthesis data’ from one million molecules stored in PubChem.<sup>107</sup> Fig. 12 shows a 3D plot of the structure index, the SA score with respect to their solvation energy. Candidates with favourable solvation may tend to exhibit high structural complexity along with unfavourable chemical reactivity, which means that if the possibility to synthesize the molecule is low and it is non-reactive, or the possibility to synthesize the molecule is high but it is reactive, then even if the solvation energy is favourable, it will not be treated as a good material. In such cases, the significance comes from detecting the candidates not only with low solvation energy but also with

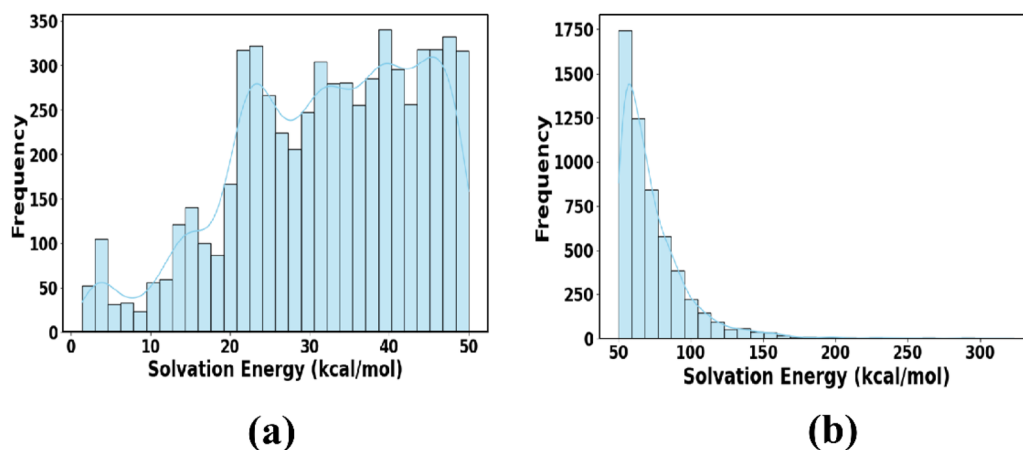


Fig. 11 Distribution of solvation energies of the refined dataset: (a) less than 50 kcal mol<sup>-1</sup> and (b) greater than 50 kcal mol<sup>-1</sup>.

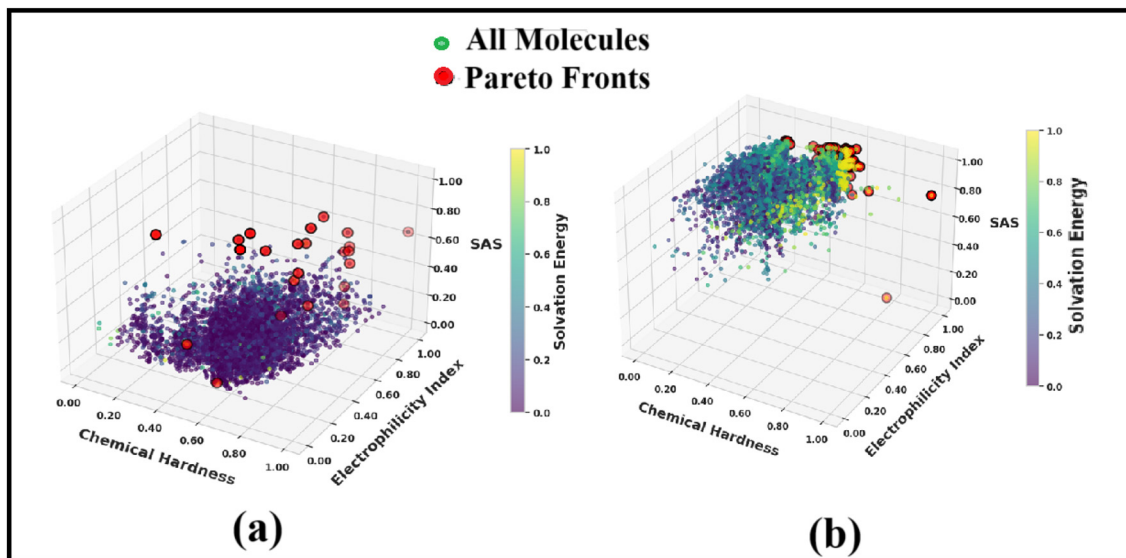


Fig. 12 3D plots of structure index: SA score, chemical hardness, and electrophilicity index with (a) high solvation energy and (b) low solvation energy.

low complexity and suitable chemical reactivity. The optimal solutions chosen are those molecules that are prone to desolvation and are easy to synthesize, less reactive, not very electrophilic, and result in very stable molecules with a high probability of not forming dendrites when they come in contact with electrodes. The non-ideal solutions have very low desolvation probability, along with low reactivity and stability. This methodology ensures a rigorous, computationally efficient approach that balances prediction accuracy with chemical relevance, ultimately guiding the selection of molecules with enhanced screening of ideal and non-ideal SEI molecules.

Our method uses features based on the molecular structures to predict the property values. This way, the predictions are directly connected to the structures. We used these predicted values to find the best molecules through Pareto optimization. This ensures a clear link between the molecular structures and the computed values. We used molecular features (as inputs) to accurately predict property values. These predictions reflect the molecular characteristics, allowing us to connect our results back to the molecular level. The Pareto optimization then helps to identify molecules with the best combined properties, bridging the numerical data and molecular-level conclusions.

The SA scores of our considered structures for low solvation energy (below  $50 \text{ kcal mol}^{-1}$ ) lie between 1.00 and 2.089, and for high solvation energies (above  $50 \text{ kcal mol}^{-1}$ ), the SA score lies in the range of 3.994–7.465. Moreover, we are searching for optimal structures that can be easily synthesized, bearing both low<sup>108,109</sup> and high solvation energies, and based on the reactivity parameters; however, screening from 6134 and 5530 entries is challenging. Therefore, to obtain optimal structures that possess low solvation energies with low SA scores, are less reactive, not very electrophilic, and become very stable molecules, and reactive molecules with high solvation energies and high SA scores, we implemented multi-objective optimization to filter these structures through the Pareto filter method<sup>110</sup>

(Fig. 12). We optimized the solvation energy with reactivity parameters along with their respective SA scores and finally got nine optimal values (Table 2) (low solvation energy with low SA score, inert molecules) and ten optimal values (Table 3) (high solvation energy with high SA score, reactive molecules). Fig. 12 shows the 3D scatter plots of all the datapoints having all optimized properties; the multi-coloured dots in the plot refer to all the structures, and the red marks are the Pareto optimal solutions, which have been considered as the potential structures. Among these potential candidates, as shown in Fig. 13, we selected a total of nineteen (nine ideal SEI and ten non-ideal SEI) possible candidates. Tables 2 and 3 display the predicted and actual values of optimized properties of two important input features for solvation energy, namely the dipole moment and heavy atom account. As observed from both tables, the predicted values of chemical hardness, electrophilicity index, SA score, and solvation energy are closely aligned with the actual values, demonstrating our model's robustness, thus supporting our proceeding with the optimization of these predicted properties to find the best candidates.

Let us consider molecule 'a' of Table 2 that exhibits the predicted solvation energy of  $1.595 \text{ kcal mol}^{-1}$ , predicted chemical hardness of 17.530, predicted electrophilicity index of 0.299, and predicted SA score of 2.260, suggesting that the ML model captures the underlying relation, effectively having their actual values as follows: solvation energy – 1.590, electrophilicity index 0.290, chemical hardness 17.639, and SAS 2.089. However, other structures (Table 2) exhibit solvation energy values of 1.927, 1.844, 5.677, 1.832, 1.433, 1.750, 2.094, 2.133, 2.492  $\text{kcal mol}^{-1}$ , which are very close to their actual values. For one of the reactivity parameters, *i.e.*, chemical hardness, the predicted values fall in the range of 10.796–11.971, which is very close to the actual value range of 10.792–11.986. Similarly, the predicted electrophilicity index from molecule 'b' to molecule 'j' has values of 0.390, 0.342, 0.270, 0.389, 0.379, 0.374, 0.359,

Table 2 Top nine ideal Pareto optimal solutions with their respective predicted properties

Molecules	Chemical hardness	Predicted chemical hardness	Electrophilicity index	Predicted electrophilicity index	Solvation energy	Predicted solvation energy	SAS	Predicted SAS	Dipole moment	Heavy atom count
a	17.639	17.530	0.290	0.299	1.590	1.595	2.089	2.260	0.005	5
b	11.243	11.249	0.392	0.390	1.925	1.927	1.000	1.236	0.024	5
c	10.912	10.910	0.342	0.342	1.841	1.844	1.000	1.268	0.000	6
d	11.469	11.470	0.266	0.270	5.690	5.677	1.014	1.219	0.021	3
e	10.848	10.861	0.390	0.389	1.799	1.832	1.000	1.290	0.015	7
f	11.986	11.971	0.381	0.379	1.423	1.433	1.755	1.786	0.055	3
g	11.784	11.771	0.374	0.374	1.757	1.750	1.606	1.773	0.000	4
h	11.426	11.426	0.358	0.359	2.092	2.094	1.209	1.621	0.000	6
i	10.792	10.796	0.332	0.333	2.134	2.133	1.549	1.428	0.077	7

Table 3 Top ten non-ideal Pareto optimal solutions with their respective properties

Molecules	Chemical hardness	Predicted chemical hardness	Electrophilicity index	Predicted electrophilicity index	Solvation energy	Predicted solvation energy	SAS	Predicted SAS	Dipole moment	Heavy atom count
a	1.782	1.820	4.856	4.823	170.414	169.177	5.483	5.483	17.396	37
b	3.993	3.995	2.481	2.482	284.010	281.501	6.561	6.562	30.609	48
c	4.016	4.005	2.020	2.030	311.290	300.982	5.408	5.379	23.513	37
d	2.625	2.579	3.101	3.113	132.507	206.703	5.409	4.6985	27.151	26
e	3.177	3.125	3.404	3.387	133.762	121.014	5.812	6.661	10.281	47
f	3.252	3.253	3.443	3.439	229.074	228.776	3.994	4.002	16.311	55
g	3.665	3.671	2.880	2.881	123.888	122.068	7.465	7.422	3.972	50
h	3.198	3.190	3.551	3.558	85.019	85.354	5.692	5.628	2.923	37
i	3.891	3.959	2.072	2.074	136.440	125.233	5.924	6.947	3.374	46
j	3.275	3.274	3.424	3.416	85.228	85.869	5.519	5.533	3.820	36

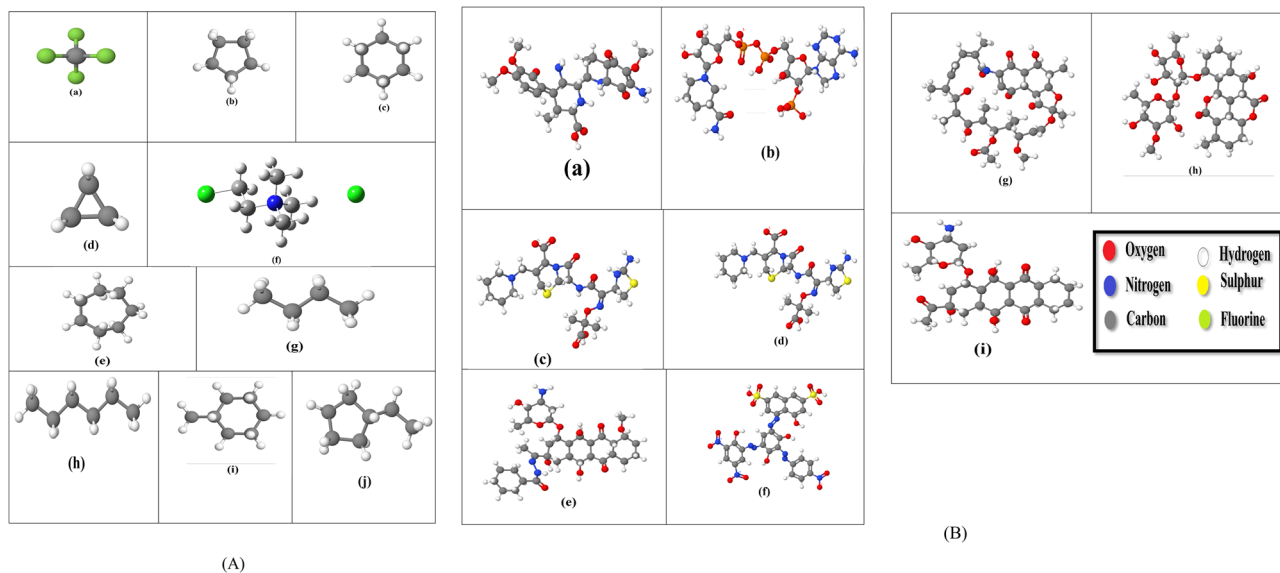


Fig. 13 Recommended (A) ideal SEI candidates and (B) non-ideal SEI candidates.

0.333, 0.389, whose actual values are 0.392, 0.342, 0.266, 0.390, 0.381, 0.374, 0.358, 0.332, 0.388, respectively. The predicted SAS lies in the range of 1.219–2.260, and actual values are in the range of 1.000–2.089. On the other hand, for non-ideal SEI, ten products (Table 3) exhibited higher solvation energy values of 170.414, 284.010, 311.290, 132.507, 133.762, 229.074, 123.888, 85.0189, 136.440, and 85.228 that correspond well with the predicted values of 169.177, 281.501, 300.982, 206.703, 121.014, 228.776,

122.068, 85.354, 125.233, and 85.869, respectively. A comprehensive comparison between the predicted and actual values for chemical hardness, electrophilicity index, and SAS for Table 3 depicts a high degree of agreement across the dataset. For all these properties of ten non-ideal candidates, the prediction closely tracks the actual value with minimal deviation. The significance of the input features associated (Tables 2 and 3) with these optimal solutions provides information about the polarity

(high dipole moment) and non-polarity (low dipole moment) of structures of the molecules. Apart from the dipole moment, the contributions of 'heavy atom count' in solvation energy are also depicted. Hence, the above discussion provides insight for the experimental validation of these (ideal and non-ideal candidates) solvation energy structures, since reasonable results were obtained. The structure possessing the necessary solvation energy and reactivity criteria should be considered on a priority basis, followed by the difficulty in synthesis. The stable SEI-forming optimal candidates, Fig. 13(A), are seen to be enriched with carbon and fluorine, while the non-ideal set of products, Fig. 13(B), are seen to have nitrogen, sulphur, oxygen, and carbon atoms, which are highly recommendable for material selection.

## Conclusion

We have employed machine learning models to forecast solvation energy and chemical reactivity. By evaluating forty-two ML models for solvation energy prediction, we have chosen the top-performing models by extracting two hundred input features. Among these forty-two models, we have chosen the top two performing models, extra tree regressor and XGBoosting, for the enhanced and accurate prediction of solvation energy and chemical reactivity parameters, which revealed the most influential features. We streamlined our findings by classifying the dataset based on solvation energies with low and high values, which are responsible and non-responsible, respectively, for impacting dendrite formation. Furthermore, the generated SA score provides the complexity needed to synthesize the structure. Our research concentrated on those materials that are resistant to excessive dendrite formation, less reactive, and easy to synthesize, as well as those that tend to be poor SEI materials with high complexity and reactive. Finally, our filtered nine stable materials identified by the Pareto optimal method, with a predicted range of solvation energy of 1.433–5.677 kcal mol<sup>-1</sup>, chemical hardness of 10.796–17.530, electrophilicity index of 0.270–0.390, along with a low synthetic accessible score of 1.219–2.260. Non-ideal candidates demonstrated low SEI-forming characteristics with predicted solvation energy, chemical hardness, electrophilicity index, and synthetic accessibility score ranges of 85.354–300.982 kcal mol<sup>-1</sup>, 1.820–4.005, 2.030–4.823, and 4.002–7.422, respectively. The identified ideal structures are very simple systems enriched with carbon and fluorine, and nitrogen, with the number of atoms ranging from 3 to 7. However, non-ideal SEI forming products are rich in nitrogen, sulphur, oxygen, and carbon atoms, whose heavy atom count ranges from 26 to 55. The dipole moment, one of the key properties regarded as influential in accurate solvation energy prediction, spans 0 D to 0.055 D (ideal SEI products) and 2.923 D to 30.609 D (non-ideal products). Hence, our results provide valuable insights, suggesting that materials with similar characteristics within the identified range and specific atom compositions can significantly influence solvation energy, offering a promising strategy to effectively prevent dendrite formation.

## Conflicts of interest

The authors declare no conflict of interest.

## Data availability

Supplementary information (SI) is available. See DOI: <https://doi.org/10.1039/d5cp02726h>.

The dataset is available at the github link '<https://github.com/Sadhana-barman/Solid-electrolyte-interphase-material/tree/main>'.

## Acknowledgements

SB thanks the Department of Science and Technology, New Delhi, for providing her with the DST-INSPIRE Fellowship.

## References

- 1 J. B. Goodenough and K.-S. Park, The Li-Ion Rechargeable Battery: A Perspective, *J. Am. Chem. Soc.*, 2013, **135**, 1167–1176, DOI: [10.1021/ja3091438](https://doi.org/10.1021/ja3091438).
- 2 D. Mohanty, J. Li, R. Born, L. C. Maxey, R. B. Dinwiddie, C. Daniel and D. L. Wood, III, Non-Destructive Evaluation of Slot-Die-Coated Lithium Secondary Battery Electrodes by in-Line Laser Caliper and IR Thermography Methods, *Anal. Methods*, 2014, **6**, 674–683, DOI: [10.1039/c3ay41140k](https://doi.org/10.1039/c3ay41140k).
- 3 D. Mohanty, A. Huq, E. A. Payzant, A. S. Sefat, J. Li, D. P. Abraham, D. L. Wood and C. Daniel, Neutron Diffraction and Magnetic Susceptibility Studies on a High-Voltage Li<sub>1.2</sub>Mn<sub>0.55</sub>Ni<sub>0.15</sub>Co<sub>0.10</sub>O<sub>2</sub> Lithium Ion Battery Cathode: Insight into the Crystal Structure, *Chem. Mater.*, 2013, **25**, 4064–4070, DOI: [10.1021/cm402278q](https://doi.org/10.1021/cm402278q).
- 4 J. Li, B. L. Armstrong, C. Daniel, J. Kiggans and D. L. Wood, Optimization of Multicomponent Aqueous Suspensions of Lithium Iron Phosphate (LiFePO<sub>4</sub>) Nanoparticles and Carbon Black for Lithium-Ion Battery Cathodes, *J. Colloid Interface Sci.*, 2013, **405**, 118–124, DOI: [10.1016/j.jcis.2013.05.030](https://doi.org/10.1016/j.jcis.2013.05.030).
- 5 J. Li, B. L. Armstrong, J. Kiggans, C. Daniel and D. L. Wood, Lithium Ion Cell Performance Enhancement Using Aqueous LiFePO<sub>4</sub> Cathode Dispersions and Polyethyleneimine Dispersant, *J. Electrochem. Soc.*, 2012, **160**, A201–A206, DOI: [10.1149/2.037302jes](https://doi.org/10.1149/2.037302jes).
- 6 S. J. An, J. Li, C. Daniel, D. Mohanty, S. Nagpure and D. L. Wood, The state of understanding of the lithium-ion-battery graphite solid electrolyte interphase (SEI) and its relationship to formation cycling, *Carbon*, 2016, **105**, 52–76, DOI: [10.1002/chin.201628293](https://doi.org/10.1002/chin.201628293).
- 7 C. Daniel, Materials and processing for lithium-ion batteries, *JOM*, 2008, **60**, 43–48, DOI: [10.1007/s11837-008-0116-x](https://doi.org/10.1007/s11837-008-0116-x).
- 8 J. R. Croy, A. Abouimrane and Z. Zhang, Next-Generation Lithium-Ion Batteries: The Promise of near-Term Advancements, *MRS Bull.*, 2014, **39**, 407–415, DOI: [10.1557/mrs.2014.84](https://doi.org/10.1557/mrs.2014.84).

- 9 S. S. Zhang, Status, Opportunities, and Challenges of Electrochemical Energy Storage, *Front. Energy Res.*, 2013, **1**, 1–6, DOI: [10.3389/fenrg.2013.00008](https://doi.org/10.3389/fenrg.2013.00008).
- 10 M. Armand and J.-M. Tarascon, Building Better Batteries, *Nature*, 2008, **451**, 652–657, DOI: [10.1038/451652a](https://doi.org/10.1038/451652a).
- 11 M. S. Whittingham, Lithium Batteries and Cathode Materials, *Chem. Rev.*, 2004, **104**, 4271–4302, DOI: [10.1021/cr020731c](https://doi.org/10.1021/cr020731c).
- 12 S. J. An, J. Li, C. Daniel, D. Mohanty, S. Nagpure and D. L. Wood, The State of Understanding of the Lithium-Ion-Battery Graphite Solid Electrolyte Interphase (SEI) and Its Relationship to Formation Cycling, *Carbon*, 2016, **105**, 52–76, DOI: [10.1016/j.carbon.2016.04.008](https://doi.org/10.1016/j.carbon.2016.04.008).
- 13 X. Zeng, M. Li, D. A. El-Hady, W. Alshitari, A. S. Al-Bogami, J. Lu and K. Amine, Commercialization of Lithium Battery Technologies for Electric Vehicles, *Adv. Energy Mater.*, 2019, **9**, 1900161, DOI: [10.1002/aenm.201900161](https://doi.org/10.1002/aenm.201900161).
- 14 H. Adenusi, G. A. Chass, S. Passerini, K. V. Tian and G. Chen, Lithium Batteries and the Solid Electrolyte Interphase (SEI)—Progress and Outlook, *Adv. Energy Mater.*, 2023, **13**, 2203307, DOI: [10.1002/aenm.202203307](https://doi.org/10.1002/aenm.202203307).
- 15 L. Lu, X. Han, J. Li, J. Hua and M. Ouyang, A Review on the Key Issues for Lithium-Ion Battery Management in Electric Vehicles, *J. Power Sources*, 2013, **226**, 272–288, DOI: [10.1016/j.jpowsour.2012.10.060](https://doi.org/10.1016/j.jpowsour.2012.10.060).
- 16 J. Wen, Y. Yu and C. Chen, A Review on Lithium-Ion Batteries Safety Issues: Existing Problems and Possible Solutions, *Mater. Express*, 2012, **2**, 197–212, DOI: [10.1166/mex.2012.1075](https://doi.org/10.1166/mex.2012.1075).
- 17 M. J. Lain and E. Kendrick, Understanding the Limitations of Lithium Ion Batteries at High Rates, *J. Power Sources*, 2021, **493**, 229690, DOI: [10.1016/j.jpowsour.2021.229690](https://doi.org/10.1016/j.jpowsour.2021.229690).
- 18 E. Peled, The Electrochemical Behavior of Alkali and Alkaline Earth Metals in Nonaqueous Battery Systems—The Solid Electrolyte Interphase Model, *J. Electrochem. Soc.*, 1979, **126**, 2047–2051, DOI: [10.1149/1.2128859](https://doi.org/10.1149/1.2128859).
- 19 M. M. Thackeray, A. de Kock, M. H. Rossouw, D. Liles, R. Bittihn and D. Hoge, Spinel Electrodes from the Li-Mn-O System for Rechargeable Lithium Battery Applications, *J. Electrochem. Soc.*, 1992, **139**, 363–366, DOI: [10.1149/1.2069222](https://doi.org/10.1149/1.2069222).
- 20 A. K. Padhi, K. S. Nanjundaswamy and J. B. Goodenough, Phospho-olivines as Positive-Electrode Materials for Rechargeable Lithium Batteries, *J. Electrochem. Soc.*, 1997, **144**, 1188–1194, DOI: [10.1149/1.1837571](https://doi.org/10.1149/1.1837571).
- 21 P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.*, 1964, **136**, B864–B871, DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- 22 W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, 1965, **140**, A1133–A1138, DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133).
- 23 R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, Electronegativity: The density functional viewpoint, *J. Chem. Phys.*, 1978, **68**, 3801–3807, DOI: [10.1063/1.436185](https://doi.org/10.1063/1.436185).
- 24 R. G. Parr and R. G. Pearson, Absolute hardness: companion parameter to absolute electronegativity, *J. Am. Chem. Soc.*, 1983, **105**, 7512–7516, DOI: [10.1021/ja00364a005](https://doi.org/10.1021/ja00364a005).
- 25 P. W. Ayers, The physical basis of the hard/soft acid/base Principle, *Faraday Discuss.*, 2007, **135**, 161–190, DOI: [10.1039/B606877D](https://doi.org/10.1039/B606877D).
- 26 R. G. Parr, L. V. Szentpály and S. Liu, Electrophilicity index, *J. Am. Chem. Soc.*, 1999, **121**, 1922–1924, DOI: [10.1021/ja983494x](https://doi.org/10.1021/ja983494x).
- 27 P. K. Chattaraj, U. Sarkar and D. R. Roy, Electrophilicity index, *Chem. Rev.*, 2006, **106**, 2065–2091, DOI: [10.1021/cr040109f](https://doi.org/10.1021/cr040109f).
- 28 T. Koopmans, Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms, *Physica*, 1934, **1**, 104–113, DOI: [10.1016/S0031-8914\(34\)90011-2](https://doi.org/10.1016/S0031-8914(34)90011-2).
- 29 J. Ferraz-Caetano, F. Teixeira and M. N. D. S. Cordeiro, Explainable Supervised Machine Learning Model to Predict Solvation Gibbs Energy, *J. Chem. Inf. Model.*, 2024, **64**, 2250–2262, DOI: [10.1021/acs.jcim.3c00544](https://doi.org/10.1021/acs.jcim.3c00544).
- 30 H. Lim and Y. Jung, Delfos: Deep Learning Model for Prediction of Solvation Free Energies in Generic Organic Solvents, *Chem. Sci.*, 2019, **10**, 8306–8315, DOI: [10.1039/c9sc02452b](https://doi.org/10.1039/c9sc02452b).
- 31 T. N. Borhani, S. García-Muñoz, C. Vanesa Luciani, A. Galindo and C. S. Adjiman, Hybrid QSPR Models for the Prediction of the Free Energy of Solvation of Organic Solute/Solvent Pairs, *Phys. Chem. Chem. Phys.*, 2019, **21**, 13706–13720, DOI: [10.1039/c8cp07562j](https://doi.org/10.1039/c8cp07562j).
- 32 V. Subramanian, E. Ratkova, D. Palmer, O. Engkvist, M. Fedorov and A. Llinas, Multisolvant Models for Solvation Free Energy Predictions Using 3D-RISM Hydration Thermodynamic Descriptors, *J. Chem. Inf. Model.*, 2020, **60**, 2977–2988, DOI: [10.1021/acs.jcim.0c00065](https://doi.org/10.1021/acs.jcim.0c00065).
- 33 H. Lim and Y. Jung, MLSolvA: Solvation Free Energy Prediction from Pairwise Atomistic Interactions by Machine Learning, *J. Cheminf.*, 2021, **13**, 42, DOI: [10.1186/s13321-021-00533-z](https://doi.org/10.1186/s13321-021-00533-z).
- 34 S. C. Kim, X. Gao, S. L. Liao, H. Su, Y. Chen, W. Zhang, L. C. Greenburg, J. A. Pan, X. Zheng, Y. Ye, M. S. Kim, P. Sayavong, A. Brest, J. Qin, Z. Bao and Y. Cui, Solvation-Property Relationship of Lithium-Sulphur Battery Electrolytes, *Nat. Commun.*, 2024, **15**, 1268, DOI: [10.1038/s41467-023-44527-x](https://doi.org/10.1038/s41467-023-44527-x).
- 35 K. Kwak, S. Park and M. D. Fayer, Dynamics around Solutes and Solute–Solvent Complexes in Mixed Solvents, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14221–14226, DOI: [10.1073/pnas.0701710104](https://doi.org/10.1073/pnas.0701710104).
- 36 K. Kwak, D. E. Rosenfeld, J. K. Chung and M. D. Fayer, Solute–Solvent Complex Switching Dynamics of Chloroform between Acetone and Dimethylsulfoxide—Two-Dimensional IR Chemical Exchange Spectroscopy, *J. Phys. Chem. B*, 2008, **112**, 13906–13915, DOI: [10.1021/jp806035w](https://doi.org/10.1021/jp806035w).
- 37 S. Boobier, D. R. J. Hose, A. J. Blacker and B. N. Nguyen, Machine Learning with Physicochemical Relationships: Solubility Prediction in Organic Solvents and Water, *Nat. Commun.*, 2020, **11**, 5753, DOI: [10.1038/s41467-020-19594-z](https://doi.org/10.1038/s41467-020-19594-z).
- 38 G. Xiong, Z. Wu, J. Yi, L. Fu and Z. Yang, *et al.*, ADMETlab 2.0: An Integrated Online Platform for Accurate and

- Comprehensive Predictions of ADMET Properties, *Nucleic Acids Res.*, 2021, **49**, W5–W14, DOI: [10.1093/nar/gkab255](https://doi.org/10.1093/nar/gkab255).
- 39 G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts and D. L. Mobley, Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database, *J. Chem. Eng. Data*, 2017, **62**, 1559–1569, DOI: [10.1021/acs.jced.7b00104](https://doi.org/10.1021/acs.jced.7b00104).
- 40 Y. Basdogan, M. C. Groenenboom, E. Henderson, S. De, S. B. Rempe and J. A. Keith, Machine Learning-Guided Approach for Studying Solvation Environments, *J. Chem. Theory Comput.*, 2019, **16**, 633–642, DOI: [10.1021/acs.jctc.9b00605](https://doi.org/10.1021/acs.jctc.9b00605).
- 41 S.-H. Chong and S. Ham, Atomic Decomposition of the Protein Solvation Free Energy and Its Application to Amyloid-Beta Protein in Water, *J. Chem. Phys.*, 2011, **135**, 204110, DOI: [10.1063/1.3610550](https://doi.org/10.1063/1.3610550).
- 42 S. Chong and S. Ham, Interaction with the Surrounding Water Plays a Key Role in Determining the Aggregation Propensity of Proteins, *Angew. Chem.*, 2014, **53**, 3961–3964, DOI: [10.1002/anie.201309317](https://doi.org/10.1002/anie.201309317).
- 43 C. W. Coley, R. Barzilay, W. H. Green, T. S. Jaakkola and K. F. Jensen, Convolutional Embedding of Attributed Molecular Graphs for Physical Property Prediction, *J. Chem. Inf. Model.*, 2017, **57**, 1757–1772, DOI: [10.1021/acs.jcim.6b00601](https://doi.org/10.1021/acs.jcim.6b00601).
- 44 C. J. Cramer and D. G. Truhlar, A Universal Approach to Solvation Modeling, *Acc. Chem. Res.*, 2008, **41**, 760–768, DOI: [10.1021/ar800019z](https://doi.org/10.1021/ar800019z).
- 45 J. S. Delaney, ESOL: Estimating Aqueous Solubility Directly from Molecular Structure, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1000–1005, DOI: [10.1021/ci034243x](https://doi.org/10.1021/ci034243x).
- 46 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins, *J. Chem. Theory Comput.*, 2015, **12**, 281–296, DOI: [10.1021/acs.jctc.5b00864](https://doi.org/10.1021/acs.jctc.5b00864).
- 47 C. Hille, S. Ringe, M. Deimel, C. Kunkel, W. E. Acree, K. Reuter and H. Oberhofer, Generalized Molecular Solvation in Non-Aqueous Solutions by a Single Parameter Implicit Solvation Scheme, *J. Chem. Phys.*, 2019, **150**, 041710, DOI: [10.1063/1.5050938](https://doi.org/10.1063/1.5050938).
- 48 A. Klamt and M. Diedenhofen, Calculation of Solvation Free Energies with DCOSMO-RS, *J. Phys. Chem. A*, 2015, **119**, 5439–5445, DOI: [10.1021/jp511158y](https://doi.org/10.1021/jp511158y).
- 49 A. Klamt, F. Eckert and W. Arlt, COSMO-RS: An Alternative to Simulation for Calculating Thermodynamic Properties of Liquid Mixtures, *Annu. Rev. Chem. Biomol. Eng.*, 2010, **1**, 101–122, DOI: [10.1146/annurev-chembioeng-073009-100903](https://doi.org/10.1146/annurev-chembioeng-073009-100903).
- 50 A. Klamt and G. Schüürmann, COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805, DOI: [10.1039/p29930000799](https://doi.org/10.1039/p29930000799).
- 51 A. V. Marenich, C. J. Cramer and D. G. Truhlar, Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions, *J. Phys. Chem. B*, 2009, **113**, 6378–6396, DOI: [10.1021/jp810292n](https://doi.org/10.1021/jp810292n).
- 52 A. V. Marenich, C. J. Cramer and D. G. Truhlar, Generalized Born Solvation Model SM12, *J. Chem. Theory Comput.*, 2012, **9**, 609–620, DOI: [10.1021/ct300900e](https://doi.org/10.1021/ct300900e).
- 53 B. Mennucci, Polarizable Continuum Model, *Wiley Interdiscip. Rev.:Comput. Mol. Sci.*, 2012, **2**, 386–404, DOI: [10.1002/wcms.1086](https://doi.org/10.1002/wcms.1086).
- 54 D. L. Mobley and J. P. Guthrie, FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files, *J. Comput.-Aided Mol. Des.*, 2014, **28**, 711–720, DOI: [10.1007/s10822-014-9747-x](https://doi.org/10.1007/s10822-014-9747-x).
- 55 G. L. Perlovich, Thermodynamic Approaches to the Challenges of Solubility in Drug Discovery and Development, *Mol. Pharmaceutics*, 2013, **11**, 1–11, DOI: [10.1021/mp400460r](https://doi.org/10.1021/mp400460r).
- 56 W. L. Jorgensen, Efficient Drug Lead Discovery and Optimization, *Acc. Chem. Res.*, 2009, **42**, 724–733, DOI: [10.1021/ar800236t](https://doi.org/10.1021/ar800236t).
- 57 E. Moine, R. Privat, B. Sirjean and J.-N. Jaubert, Estimation of Solvation Quantities from Experimental Thermodynamic Data: Development of the Comprehensive CompSol Databank for Pure and Mixed Solutes, *J. Phys. Chem. Ref. Data*, 2017, **46**, 043105, DOI: [10.1063/1.5000910](https://doi.org/10.1063/1.5000910).
- 58 L. Cheng, R. S. Assary, X. Qu, A. Jain, S. P. Ong, N. N. Rajput, K. Persson and L. A. Curtiss, Accelerating Electrolyte Discovery for Energy Storage with High-Throughput Screening, *J. Phys. Chem. Lett.*, 2015, **6**, 283–291, DOI: [10.1021/jz502319n](https://doi.org/10.1021/jz502319n).
- 59 J. Kim, S. Ko, C. Noh, H. Kim, S. Lee, D. Kim, H. Park, G. Kwon, G. Son, J. W. Ko, Y. Jung, D. Lee, C. B. Park and K. Kang, Biological Nicotinamide Cofactor as a Redox-Active Motif for Reversible Electrochemical Energy Storage, *Angew. Chem.*, 2019, **58**, 16764–16769, DOI: [10.1002/anie.201906844](https://doi.org/10.1002/anie.201906844).
- 60 D. Bedrov, O. Borodin and J. B. Hooper, Li<sup>+</sup> Transport and Mechanical Properties of Model Solid Electrolyte Interphases (SEI): Insight from Atomistic Molecular Dynamics Simulations, *J. Phys. Chem. C*, 2017, **121**, 16098–16109, DOI: [10.1021/acs.jpcc.7b04247](https://doi.org/10.1021/acs.jpcc.7b04247).
- 61 A. Muralidharan, M. I. Chaudhari, L. R. Pratt and S. B. Rempe, Molecular Dynamics of Lithium Ion Transport in a Model Solid Electrolyte Interphase, *Sci. Rep.*, 2018, **8**, 10736, DOI: [10.1038/s41598-018-28869-x](https://doi.org/10.1038/s41598-018-28869-x).
- 62 T. Hou, G. Yang, N. N. Rajput, J. Self, S.-W. Park, J. Nanda and K. A. Persson, The Influence of FEC on the Solvation Structure and Reduction Reaction of LiPF<sub>6</sub>/EC Electrolytes and Its Implication for Solid Electrolyte Interphase Formation, *Nano Energy*, 2019, **64**, 103881, DOI: [10.1016/j.nanoen.2019.103881](https://doi.org/10.1016/j.nanoen.2019.103881).
- 63 O. Borodin, G. V. Zhuang, P. N. Ross and K. Xu, Molecular Dynamics Simulations and Experimental Study of Lithium

- Ion Transport in Dilithium Ethylene Dicarboxylate, *J. Phys. Chem. C*, 2013, **117**, 7433–7444, DOI: [10.1021/jp4000494](https://doi.org/10.1021/jp4000494).
- 64 M. Ebrahimi, J. B. Hooper and D. Bedrov, Structural, Mechanical, and Dynamical Properties of Amorphous Li<sub>2</sub>CO<sub>3</sub> from Molecular Dynamics Simulations, *Crystals*, 2018, **8**, 473, DOI: [10.3390/cryst8120473](https://doi.org/10.3390/cryst8120473).
- 65 T. Zhang, X. Zhu and J. Xiong, *et al.*, Electron displacement polarization of high-dielectric constant fiber separators enhances interface stability, *Nat. Commun.*, 2025, **16**, 4867, DOI: [10.1038/s41467-025-60256-9](https://doi.org/10.1038/s41467-025-60256-9).
- 66 P. Ganesh, P. R. C. Kent and D. Jiang, Solid-Electrolyte Interphase Formation and Electrolyte Reduction at Li-Ion Battery Graphite Anodes: Insights from First-Principles Molecular Dynamics, *J. Phys. Chem. C*, 2012, **116**, 24476–24481, DOI: [10.1021/jp3086304](https://doi.org/10.1021/jp3086304).
- 67 M. J. Boyer and G. S. Hwang, Molecular Dynamics Investigation of Reduced Ethylene Carbonate Aggregation at the Onset of Solid Electrolyte Interphase Formation, *Phys. Chem. Chem. Phys.*, 2019, **21**, 22449–22455, DOI: [10.1039/c9cp04316k](https://doi.org/10.1039/c9cp04316k).
- 68 S. Ringe, H. Oberhofer, C. Hille, S. Matera and K. Reuter, Function-Space-Based Solution Scheme for the Size-Modified Poisson–Boltzmann Equation in Full-Potential DFT, *J. Chem. Theory Comput.*, 2016, **12**, 4052–4066, DOI: [10.1021/acs.jctc.6b00435](https://doi.org/10.1021/acs.jctc.6b00435).
- 69 D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley and W. Sherman, Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field, *J. Chem. Theory Comput.*, 2010, **6**, 1509–1519, DOI: [10.1021/ct900587b](https://doi.org/10.1021/ct900587b).
- 70 G. König, F. C. Pickard, Y. Mei and B. R. Brooks, Predicting hydration free energies with a hybrid QM/MM approach: an evaluation of implicit and explicit solvation models in SAMPL4, *J. Comput.-Aided Mol. Des.*, 2014, **28**, 245–257, DOI: [10.1007/s10822-014-9708-4](https://doi.org/10.1007/s10822-014-9708-4).
- 71 M. C. Barrera and M. Jorge, A Polarization-Consistent Model for Alcohols to Predict Solvation Free Energies, *J. Chem. Inf. Model.*, 2020, **60**, 1352–1367, DOI: [10.1021/acs.jcim.9b01005](https://doi.org/10.1021/acs.jcim.9b01005).
- 72 J. Degen, C. Wegscheid-Gerlach, A. Zaliani and M. Rarey, On the Art of Compiling and Using “Drug-Like” Chemical Fragment Spaces, *ChemMedChem*, 2008, **3**, 1503–1507, DOI: [10.1002/cmdc.200800178](https://doi.org/10.1002/cmdc.200800178).
- 73 Y. Liu, B. Guo, X. Zou, Y. Li and S. Shi, Machine Learning Assisted Materials Design and Discovery for Rechargeable Batteries, *Energy Storage Mater.*, 2020, **31**, 434–450, DOI: [10.1016/j.ensm.2020.06.033](https://doi.org/10.1016/j.ensm.2020.06.033).
- 74 C. Ling, A Review of the Recent Progress in Battery Informatics, *npj Comput. Mater.*, 2022, **8**, 33, DOI: [10.1038/s41524-022-00713-x](https://doi.org/10.1038/s41524-022-00713-x).
- 75 D. P. Finegan, I. Squires, A. Dahari, S. Kench, K. L. Jungjohann and S. J. Cooper, Machine-Learning-Driven Advanced Characterization of Battery Electrodes, *ACS Energy Lett.*, 2022, **7**, 4368–4378, DOI: [10.1021/acsenerylett.2c01996](https://doi.org/10.1021/acsenerylett.2c01996).
- 76 X. Chen, X. Liu, X. Shen and Q. Zhang, Applying Machine Learning to Rechargeable Batteries: From the Microscale to the Macroscale, *Angew. Chem.*, 2021, **60**, 24354–24366, DOI: [10.1002/anie.202107369](https://doi.org/10.1002/anie.202107369).
- 77 T. Gao and W. Lu, Machine Learning toward Advanced Energy Storage Devices and Systems, *iScience*, 2021, **24**, 101936, DOI: [10.1016/j.isci.2020.101936](https://doi.org/10.1016/j.isci.2020.101936).
- 78 Y. Liu, O. C. Esan, Z. Pan and L. An, Machine Learning for Advanced Energy Materials, *Energy AI*, 2021, **3**, 100049, DOI: [10.1016/j.egyai.2021.100049](https://doi.org/10.1016/j.egyai.2021.100049).
- 79 J. Mao, J. Miao, Y. Lu and Z. Tong, Machine Learning of Materials Design and State Prediction for Lithium Ion Batteries, *Chin. J. Chem. Eng.*, 2021, **37**, 1–11, DOI: [10.1016/j.cjche.2021.04.009](https://doi.org/10.1016/j.cjche.2021.04.009).
- 80 A. J. Y. Wong, X. Zhou, Y. Lum, Z. Yao, Y. C. Chua, Y. Wen and Z. W. Seh, Battery Materials Discovery and Smart Grid Management Using Machine Learning, *Batteries Supercaps*, 2022, **5**, e202200309, DOI: [10.1002/batt.202200309](https://doi.org/10.1002/batt.202200309).
- 81 X. Feng, Q. Zhang and Z. W. Seh, Toward Automated Computational Discovery of Battery Materials, *Adv. Mater. Technol.*, 2023, **8**, 2200616, DOI: [10.1002/admt.202200616](https://doi.org/10.1002/admt.202200616).
- 82 A. Y. S. Eng, C. B. Soni, Y. Lum, E. Khoo, Z. Yao, S. K. Vineeth, V. Kumar, J. Lu, C. S. Johnson, C. Wolverton and Z. W. Seh, Theory-Guided Experimental Design in Battery Materials Research, *Sci. Adv.*, 2022, **8**, eabm2422, DOI: [10.1126/sciadv.abm2422](https://doi.org/10.1126/sciadv.abm2422).
- 83 F. H. Vermeire and W. H. Green, Transfer Learning for Solvation Free Energies: From Quantum Chemistry to Experiments, *Chem. Eng. J.*, 2021, **418**, 129307, DOI: [10.1016/j.cej.2021.129307](https://doi.org/10.1016/j.cej.2021.129307).
- 84 J. Yu, C. Zhang, Y. Cheng, Y.-F. Yang, Y.-B. She, F. Liu, W. Su and A. Su, SolvBERT for Solvation Free Energy and Solubility Prediction: A Demonstration of an NLP Model for Predicting the Properties of Molecular Complexes, *Digital Discovery*, 2023, **2**, 409–421, DOI: [10.1039/d2dd00107a](https://doi.org/10.1039/d2dd00107a).
- 85 G. Landrum RDKit: Open-source cheminformatics 2022\_09\_4 (Q3 2022) Release – January 16, 2023. <https://www.rdkit.org/> (accessed January 18, 2023).
- 86 E. W. C. Spotte-Smith, S. M. Blau, X. Xie, H. D. Patel, M. Wen, B. Wood, S. Dwaraknath and K. A. Persson, Quantum Chemical Calculations of Lithium-Ion Battery Electrolyte and Interphase Species, *Sci. Data*, 2021, **8**, 203, DOI: [10.1038/s41597-021-00986-9](https://doi.org/10.1038/s41597-021-00986-9).
- 87 C. Chatfield, Exploratory Data Analysis, *Eur. J. Oper. Res.*, 1986, **23**, 5–13, DOI: [10.1016/0377-2217\(86\)90209-2](https://doi.org/10.1016/0377-2217(86)90209-2).
- 88 S. R. Pandala and B. Silva, *Lazy Predict Project: The Python Package Index*, 2022, <https://pypi.org/project/lazypredict/>.
- 89 T. Zhang, X. Zhu, J. Xiong, Z. Xue, Y. Cao, K. C. Gordon, G. Xu and M. Zhu, Electron Displacement Polarization of High-Dielectric Constant Fiber Separators Enhances Interface Stability, *Nat. Commun.*, 2025, **16**, 4867, DOI: [10.1038/s41467-025-60256-9](https://doi.org/10.1038/s41467-025-60256-9).
- 90 L. S. Shapley, A Value for n-Person Games, in *Contributions to the Theory of Games*, ed. H. W. Kuhn and A. W. Tucker, Princeton University Press, Princeton, NJ, 1953, vol. 2, pp. 307–317.
- 91 A. Goldstein, A. Kapelner, J. Bleich and E. Pitkin, Peeking Inside the Black Box: Visualizing Statistical Learning with

- Plots of Individual Conditional Expectation, *J. Comput. Graph. Stat.*, 2015, **24**, 44–65, DOI: [10.1080/10618600.2014.907095](https://doi.org/10.1080/10618600.2014.907095).
- 92 B. M. Greenwell, B. C. Boehmke and A. J. McCarthy, *A Simple and Effective Model-Based Variable Importance Measure*, *arXiv*, 2018, preprint, arXiv:1805.04755, DOI: [10.48550/arXiv.1805.04755](https://doi.org/10.48550/arXiv.1805.04755).
- 93 J.-H. Seok, S. Lee, D.-A. Lim, K. H. Ahn, C. H. Lee, K. Kim and D.-W. Kim, A Single Lithium-Ion Conducting Monomer as a SEI-Forming Additive for Lithium-Ion Batteries, *J. Mater. Chem. A*, 2025, **13**, 13976–13987, DOI: [10.1039/d5ta00347d](https://doi.org/10.1039/d5ta00347d).
- 94 L. Ma, M. S. Kim and L. A. Archer, Stable Artificial Solid Electrolyte Interfaces for Lithium Batteries, *Chem. Mater.*, 2017, **29**, 4181–4189, DOI: [10.1021/acs.chemmater.6b03687](https://doi.org/10.1021/acs.chemmater.6b03687).
- 95 Y. Wu, C. Wang, C. Wang, Y. Zhang, J. Liu, Y. Jin, H. Wang and Q. Zhang, Recent Progress in SEI Engineering for Boosting Li Metal Anodes, *Mater. Horiz.*, 2024, **11**, 388–407, DOI: [10.1039/d3mh01434g](https://doi.org/10.1039/d3mh01434g).
- 96 J. Lee, J. Kim, S. Kim, C. Jo and J. Lee, A Review on Recent Approaches for Designing the SEI Layer on Sodium Metal Anodes, *Mater. Adv.*, 2020, **1**, 3143–3166, DOI: [10.1039/d0ma00695e](https://doi.org/10.1039/d0ma00695e).
- 97 X. Zhang, Y. Liu and J. Wang, Recent Progress in Aqueous Zinc-Ion Battery Materials, *Energy Mater.*, 2021, **1**, 1–20, DOI: [10.20517/energymater.2021.04](https://doi.org/10.20517/energymater.2021.04).
- 98 S. Miertuš, E. Scrocco and J. Tomasi, Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of AB Initio Molecular Potentials for the Prediction of Solvent Effects, *Chem. Phys.*, 1981, **55**, 117–129, DOI: [10.1016/0301-0104\(81\)85090-2](https://doi.org/10.1016/0301-0104(81)85090-2).
- 99 J. Zhang, M. Liu, J. Qi, N. Lei, S. Guo, J. Li, X. Xiao and L. Ouyang, Advanced Mg-Based Materials for Energy Storage: Fundamental, Progresses, Challenges and Perspectives, *Prog. Mater. Sci.*, 2025, **148**, 101381, DOI: [10.1016/j.pmatsci.2024.101381](https://doi.org/10.1016/j.pmatsci.2024.101381).
- 100 K. Xu, Nonaqueous Liquid Electrolytes for Lithium-Based Rechargeable Batteries, *Chem. Rev.*, 2004, **104**, 4303–4418, DOI: [10.1021/cr030203g](https://doi.org/10.1021/cr030203g).
- 101 Y. Ein-Eli, A New Perspective on the Formation and Structure of the Solid Electrolyte Interface at the Graphite Anode of Li-Ion Cells, *Electrochem. Solid-State Lett.*, 1999, **2**, 212, DOI: [10.1149/1.1390787](https://doi.org/10.1149/1.1390787).
- 102 X. Wang, S. Li, W. Zhang, D. Wang, Z. Shen, J. Zheng, H. L. Zhuang, Y. He and Y. Lu, Dual-Salt-Additive Electrolyte Enables High-Voltage Lithium Metal Full Batteries Capable of Fast-Charging Ability, *Nano Energy*, 2021, **89**, 106353, DOI: [10.1016/j.nanoen.2021.106353](https://doi.org/10.1016/j.nanoen.2021.106353).
- 103 K. He, Y. Xiong, C. Zhang, Z. Dou, T. Yi, S. Lin, C. Li and Y. Sun, An Investigation on the Electrochemical and Thermal Characteristics of  $\text{LiMn}_{0.6}\text{Fe}_{0.4}\text{PO}_4/\text{LiNi}_{0.5}\text{Co}_{0.2}\text{Mn}_{0.3}\text{O}_2$  Composite Cathode Materials for Lithium-Ion Batteries in Different Health States, *J. Electrochem. Soc.*, 2023, **170**, 090501, DOI: [10.1149/1945-7111/acf0eb](https://doi.org/10.1149/1945-7111/acf0eb).
- 104 Z. Wen, Y. Kang, Q. Wu, X. Shen, P. Lai, Y. Yang, C. C. Li and J. Zhao, High-Interfacial-Energy Heterostructure Facilitates Large-Sized Lithium Nucleation and Rapid  $\text{Li}^+$  Desolvation Process, *Sci. Bull.*, 2022, **67**, 2531–2540, DOI: [10.1016/j.scib.2022.11.026](https://doi.org/10.1016/j.scib.2022.11.026).
- 105 M. Mao, X. Ji, Q. Wang, Z. Lin, M. Li, T. Liu, C. Wang, Y.-S. Hu, H. Li, X. Huang, L. Chen and L. Suo, Anion-Enrichment Interface Enables High-Voltage Anode-Free Lithium Metal Batteries, *Nat. Commun.*, 2023, **14**, 1082, DOI: [10.1038/s41467-023-36853-x](https://doi.org/10.1038/s41467-023-36853-x).
- 106 P. Ertl and A. Schuffenhauer, Estimation of Synthetic Accessibility Score of Drug-Like Molecules Based on Molecular Complexity and Fragment Contributions, *J. Cheminf.*, 2009, **1**, 8, DOI: [10.1186/1758-2946-1-8](https://doi.org/10.1186/1758-2946-1-8).
- 107 The PubChem Database. <https://pubchem.ncbi.nlm.nih.gov/>.
- 108 Z. Jiang, J. Mo, C. Li, H. Li, Q. Zhang, Z. Zeng, J. Xie and Y. Li, Anion-Regulated Weakly Solvating Electrolytes for High-Voltage Lithium Metal Batteries, *Energy Environ. Mater.*, 2023, **6**, e12440, DOI: [10.1002/eeem2.12440](https://doi.org/10.1002/eeem2.12440).
- 109 L. Li, K. Ren, W. Xie, Q. Yu, S. Wu, H.-W. Li, M. Yao, Z. Jiang and Y. Li, Do Weaker Solvation Effects Mean Better Performance of Electrolytes for Lithium Metal Batteries?, *Chem. Sci.*, 2025, **16**, 7981–7988, DOI: [10.1039/d5sc01495f](https://doi.org/10.1039/d5sc01495f).
- 110 K. Miettinen, *Nonlinear Multiobjective Optimization*, Springer, US, 1998, DOI: [10.1007/978-1-4615-5563-6](https://doi.org/10.1007/978-1-4615-5563-6).