

Article

Not peer-reviewed version

Designing Imidazolium-Mediated Polymer Electrolytes for Lithium-Ion Batteries Using Machine-Learning Approaches: An Insight into Ionene Materials

Ghazal Piroozi and [Irshad Kammakakam](#) *

Posted Date: 27 June 2025

doi: [10.20944/preprints202506.2274.v1](https://doi.org/10.20944/preprints202506.2274.v1)

Keywords: Poly (Ionic Liquids); Polymer electrolytes; Lithium-ion Batteries; Machine learning; CatBoost model



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Designing Imidazolium-Mediated Polymer Electrolytes for Lithium-Ion Batteries Using Machine-Learning Approaches: An Insight into Ionene Materials

Ghazal Piroozi and Irshad Kammakakam *

School of Sciences and Humanities, Department of Chemistry, Nazarbayev University;
Astana 010000, Kazakhstan

* Correspondence: irshad.kammakakam@nu.edu.kz

Abstract

Over the past few decades, lithium-ion batteries (LIBs) have gained significant attention due to their inherent potential for environmental sustainability and unparalleled energy storage efficiency. To enhance the performance of lithium-ion batteries, electrolytes have garnered considerable attention as a key component of these batteries. Meanwhile, polymer electrolytes have gained popularity in several fields due to their ability to adapt to various battery geometries, enhanced safety features, greater thermal stability, and effectiveness in reducing dendrite growth on the anode. In general, polymer electrolytes are composed of polymer matrices and lithium salts, mainly categorized as solid polymer electrolytes (SPEs) and gel polymer electrolytes (GPEs), which provide higher energy densities while maintaining structural integrity and safety. Despite many advantages, offering relatively lower ionic conductivity as compared to liquid electrolytes, polymer electrolytes are limited to advanced applications. This limitation has led to recent studies revolving around the development of poly (ionic liquids) (PILs), particularly imidazolium-mediated polymer backbones as novel electrolyte materials, which can increase the conductivity with fine-tuning structural benefits, while maintaining the advantages of both solid and gel electrolytes. There have been various structural conformations explored in the design of multiple PILs, and the accurate measurement of conductivity is typically performed in laboratories, which can be both costly and time-consuming. Therefore, in this study, we aimed to develop intelligent models for the accurate estimation of ionic conductivity in exclusive imidazolium polymeric ionic liquids (PILs). For this purpose, a dataset consisting of 120 datapoints, including 8 different polymers, encompassing all the imidazolium-based PILs reported to date, was compiled from the literature. Most importantly, this study foresees the benefits of newly integrated PIL substructures, so-called ionenes, toward the performance of LIB applications. Four machine learning (ML) models of CatBoost, RF, XGBoost, and LightGBM were developed in this study by incorporating chemical structure and temperature as the models' inputs. The results indicated the superior performance of the CatBoost model compared to other models with R^2 , RMSE, and MAE of 0.986, 0.000187, and 0.0000952, respectively. The importance of features in predicting conductivity was investigated using the CatBoost model. The results indicated that temperature plays the leading role, followed by chemical descriptors such as, PIL_BCUT2D_MRLOW, PIL_SMR_VSA6, and PIL_EState_VSA8. Moreover, the best-performing model (CatBoost) was used to predict conductivity for three novel ionenes, paving the way for a new approach to utilizing innovative polymer architecture toward LIB applications.

Keywords: poly (ionic liquids); polymer electrolytes; lithium-ion batteries; machine LEARNING; CatBoost model

1. Introduction

As the world moves towards green energy, storage systems, including hydrogen storage [1,2], pumped hydrogen storage [3], flow batteries[4], and especially lithium-ion batteries[5] are increasingly in demand. A revolution in electronic devices began 30 years ago when Sony successfully commercialized the world's first lithium-ion battery (LIB) [5]. Since then, LIBs have gained significant attention due to their inherent potential for environmental sustainability and unparalleled energy storage efficiency. LIBs application is not limited to portable electronics; they are also used in various energy sectors and devices, including but not limited to hybrid and big electric vehicles, remote-controlled devices, solar energy equipment, medical tools, and more. Their use is also increasing in the aerospace and military industries[5–7]. LIBs are a part of the rechargeable family of batteries, similar to other batteries, and consist of four main components: the anode, cathode, electrolyte, and separator (Figure 1).

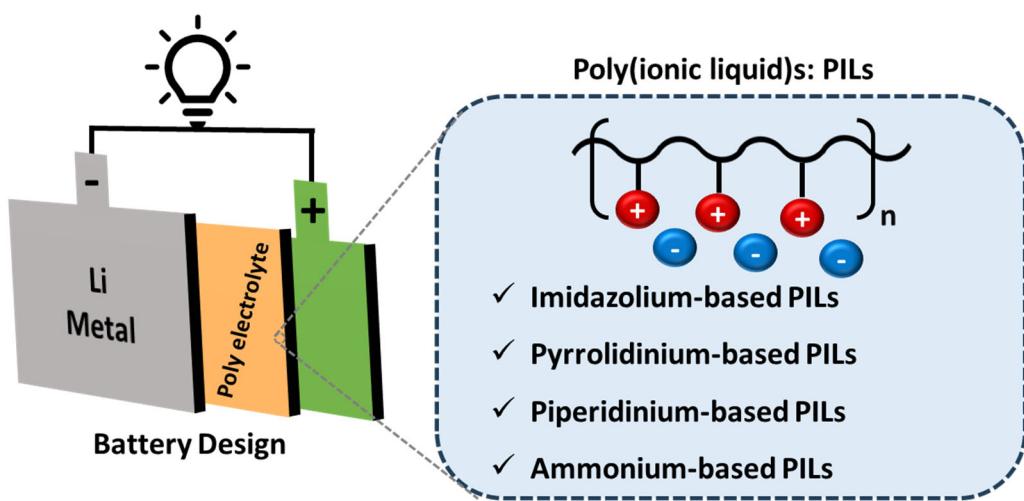


Figure 1. Schematic of a lithium-ion battery illustrating the cathode, anode, electrolyte, and common types of poly(ionic liquid) electrolytes.

To enhance lithium-ion battery performance, electrolytes have garnered significant attention as a key component of batteries. Organic electrolytes consisting of linear and alkyl carbonates are well-known and utilized for their wide operating voltage [1]. Meanwhile, polymer electrolytes have gained popularity in the fields of electrical, aerospace, automotive, and electronics due to their ability to adapt to various battery geometries, improved safety features, low manufacturing costs, higher thermal stability, and effectiveness in reducing dendrite growth on the anode. Polymer electrolytes are composed of polymer matrices and lithium salts that were initially introduced during the 1970s[2–5]. Despite these advantages, polymer electrolytes are highly volatile and flammable, posing significant safety risks. Consequently, research on non-flammable electrolytes with a high lithium-ion transfer number is ongoing to enhance the safety and efficiency of lithium batteries [6]. Solutions being investigated include solid ceramic electrolytes, polymer electrolytes (solid, gel, and composite), aqueous lithium-ion batteries, fluorinated structures, and ionic liquids [3,7,8].

Lithium batteries rely on liquid electrolytes, which have the advantages of high ionic conductivity and superior wetting performance at the electrode surface [9]. Ionic liquids (ILs) are considered emerging potential candidates for replacing carbonate-based electrolytes in the market. ILs are salts with a melting point below 100 °C that have high chemical and thermal stability, as well as very low or zero vapor pressure. Such characteristics enable room-temperature ILs to be ideal options for a broad range of uses, particularly in electrochemical devices like LIBs [9,10]. In general, ILs are a class of molten salts consisting of an array of asymmetric organic cations and organic or inorganic anions. The most widely studied ILs are ammonium-based ILs such as imidazolium, pyrrolidinium, and quaternary ammonium-based ILs [11]. At the same time, polymerized ionic

liquids (PILs) are a subset of polyelectrolytes that share many features of 'free' ILs, including high conductivity, while offering inherent safety and performance issues as full assembled LIBs. Although many PILs have various ammonium-based cations tethered to the polymeric backbones, imidazolium cations are reported to dominate due to their structural integrity and ease of synthesis. Overall, by combining the benefits of polymer electrolytes and ILs, many research studies still have potential chances to discuss and investigate toward high-performance LIBs [12].

On the other hand, the emergence of machine learning (ML) has entered the material science field into a new era. ML has accelerated the process of material discovery, design, and optimization by employing large datasets and advanced algorithms [13]. Most interestingly, ML models can make precise predictions by identifying patterns within the existing datasets while avoiding the pressing challenges of experimentation, including related timeframe and costs. Although ML models are simple mathematical calculations, the complex nature of materials, particularly polymers, presents intricate challenges in developing ML models. Therefore, recent advances in ML models have been widely examined, with more contemporary applications in macromolecular studies, discussing group contributions and chemical structures via SMILES representation. Furthermore, recent ML advancements in the polymer field include accelerated polymer simulations [14], prediction of polymer properties [15,16], adhesion strength prediction[17], and polymer discovery and design [18,19].

Technically, large datasets are necessary to train machine learning (ML) models. Nevertheless, a recent study employed a self-supervised strategy using a graph neural network (GNN) to predict polymer properties solely based on polymer structure data, as reported by Gao et al. [20]. Those results paved the way for tuning to be possible on smaller datasets for specific property prediction tasks, thanks to the pre-trained GNN. In data-scarce scenarios, the results further indicated that the ensemble pre-training approach outperforms other approaches, for electron affinity and ionization potential root mean square error (RMSE) was reduced by 28.39% and 19.09%, respectively. Kazemi-Khasragh et al. [28] employed a transfer strategy approach to focus on the prediction of mechanical and thermal properties of linear polymers. Firstly, the artificial neural network (ANN) algorithm was pre-trained to predict heat capacity at constant pressure (C_p) using 124 data points, then the pre-trained model was fine-tuned to predict specific capacity, shear modulus flexural stress strength, and dynamic viscosity. In addition to transfer learning, researchers in this field have also employed other machine learning models. Babbar et al [21], developed three ML models, namely ANN, convolutional neural network (CNN), and ridge regression (RR), to predict the glass transition temperature (T_g) of polymers. In this study, two types of molecular fingerprints were used as input features: physicochemical and topological fingerprints. The former was extracted from RDKit and used as input for RR and ANN models, while the latter was derived from the SMILES representation using one-hot encoding and used as input for the CNN model. The results highlighted the reasonable performance of RR compared to powerful non-linear models of ANN and CNN. Ascencio-Medina et al [22], by analyzing a dataset of 86 polymers, investigated the dielectric permittivity in polymers. They employed a genetic algorithm to select the most relevant descriptors from a set of 1273 descriptors. Then, by using a gradient boosting regressor (GBR), the dielectric constant was predicted. This model achieved high accuracy with a correlation coefficient (R^2) of 0.938 and 0.822 for the training and test sets, respectively.

Electrical conductivity is another essential property of polymers, which plays a vital role in their ability to transport charge. However, limited research has been conducted in this field for predicting the conductivity of polymers, as summarized in Table 1. Hatakeyama-Sato et al[23], constructed a 104-entry database of lithium-conducting solid polymers, the largest of its kind. The authors employed a transfer-learned graph neural network (GNN) for predicting the conductivity of electrolytes, resulting in a mean absolute error (MAE) of less than 1. The unbiased predictions of the model led to the discovery of superionic conductors with ionic conductivities of about 10⁻³ at room temperature. Li et al [24], incorporated GNNs with quantum calculations to develop an automatic system for identifying potential ionic liquids (ILs) for ionic liquid polymer electrolytes (IPEs). Firstly,

based on the ensemble learning of support vector machine (SVM), random forest (RF), XGBoost, and graph convolutional neural networks (GCNN) the phase of ILs was predicted. After identifying the IL candidates, the datapoints were classified based on conductivity type ($\sigma \geq 5$ and $\sigma < 5$). XGBoost and SVM performed better than other models. According to the results, the median values reported for the groups with $\sigma < 5$ and $\sigma \geq 5$ are 1.8 and 9.1 mS cm⁻¹, respectively. Most recently, Bradford et al [19] constructed a chemistry-informed ML workflow that predicted the conductivity of solid polymers by using chemical structure, temperature, molecular weight (Mw), and salt concentration. The results were primarily developed using the novel approach of ChemArr, which was benchmarked against two other machine learning models, Chempop and XGBoost. Among the developed models, XGBoost exhibited weak performance, while ChemArr outperformed other models by showing low MAE and a high Spearman rank correlation coefficient.

Taking into account the importance of designing polymerized ionic liquids (PILs) as polyelectrolytes in LIB applications, in this study, we have exclusively considered the structural and electrical performance of imidazolium-based poly (IL)s via machine learning (ML) approaches. As such, four impressive ML models, namely: CatBoost, RF, XGBoost, and LightGBM, have been meticulously selected to predict ionic conductivity of imidazolium-based PILs both in the form of solid polymer electrolytes (SPE) as well as gel polymer electrolytes (GPE). Regarding this purpose, input features, including chemical structure and temperature, were gathered from the literature. The models were trained and tested on the dataset that was collected. Afterwards, the importance of the input features on conductivity prediction was investigated by the best-performing model. After training and testing the models, the best-performing model was used to predict conductivity for a new set of data points. Most importantly, we have aimed to anticipate the potential candidacy of ionene materials having conventional polymeric functional groups such as amides and imides. Specifically, we have focused on investigating the electrochemical performance in terms of conductivity data of our recently developed new type of ionic polymer, in which imidazolium cations are tethered within the rigid polyimide (PI) and polyamide (PA) substructures, as depicted in Figure 2 [25–27]. Overall, this study is organized into four sections. First, an introduction to lithium batteries, polymer electrolytes, and ML studies is provided. Then, in the methodology section, a description of each of the developed models is given. Moreover, this section describes the data gathering process. In the Results and Discussion section, the performance of the models will be assessed using various graphical and statistical methods. Additionally, a section is dedicated to the employment of ML models in predicting conductivity for ionenes with robust polymeric backbones. Ultimately, conclusions are drawn toward the insight of the usage of ionene polymeric materials as polymer electrolytes for LIB applications.

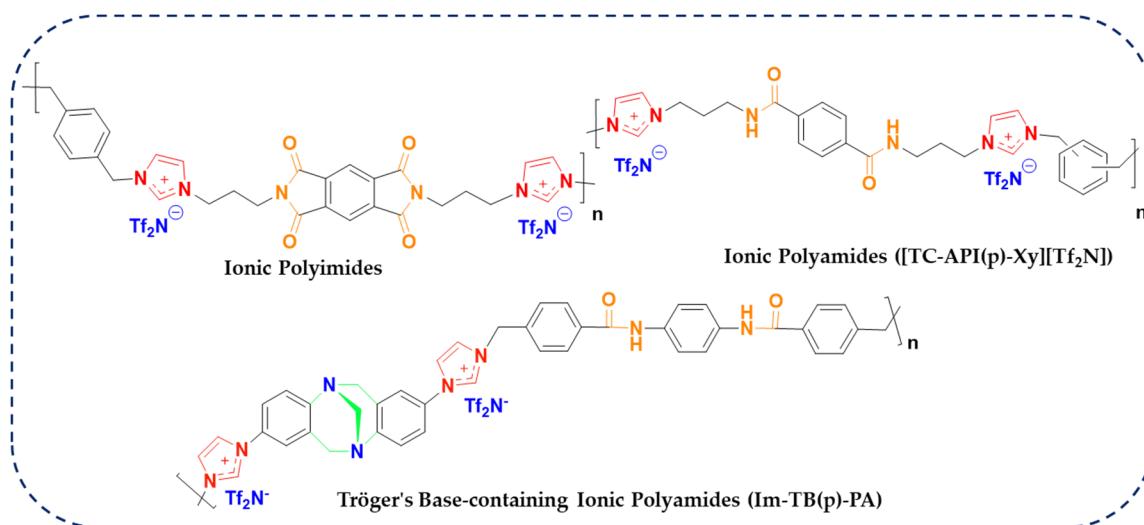


Figure 2. Chemical structures of novel ionene materials having polymeric backbones of amide and imide functionalities used for the ML approach, quantifying the conductivity data in this study.

Table 1. Summary of AI models used for conductivity prediction.

Ref	ML models	Inputs	Evaluation metrics
[28]	NN	Chemical composition, Temperature	NA
[23]	GNN	Chemical structures, Composition ratio, Temperatures	$R^2 = 0.16$
[24]	1-Unsupervised learning 2-Ensemble of SVM, molecular structure fingerprints, RF, XGB, and GCNNE	Molecular structure descriptors, Electronic structural variables, 3D molecular structure fingerprints, RF, XGB, and GCNNE Electrochemical window	$R^2 = 0.82$, $MAE = 1.8$
[29]	DNN	Chemical structure, Temperature, ion exchange capacity	$R_p = 0.951$, $RMSE = 0.014$
[30]	CatBoost, XGBoost, RF*	Chemical structure, Temperature, Ion forms, Polymer main-chain types, anion-conducting moieties.	$RMSE^* = 0.014$, $MAE^* = 0.01$
[19]	ChemArr*, Chemprop, XGBoost	Chemical structure, Temperature, Mw, Spearman $R^* = 0.59$, Salt concentration	$MAE^* = 1$
[31]	RF*, KNN, SVM, Adaboost, GBM	Standard deviation of Li-X ionic bond,	
		Standard deviation of the mean	
		adjacency number of Li atom, Average	
		straight-line path electronegativity,	$MAE^* = 0.237$, $MSE^* = 0.134$
		Average straight-line path width,	
[32]	RF, XGBoost*, LR, KNN, Chemprop	Packing fraction of sublattice, Average atomic volume, Average value of Li-Li bond	
		Chemical structure of polymer, Salt	$R^{2*} = 0.93$, $MAE^* = 0.21$,
		chemical structure, Mw, Molality,	$RMSE^* = 0.31$, $MSE^* = 0.09$
		Temperature	

* Model with the best performance.

2. Materials and Methods

2.1. Model Developments

Four ML model of CatBoost, RF, XGBoost and LightGBM were employed in this study. These models will be briefly discussed in the sections below.

2.1.1. CatBoost

CatBoost as an open-source gradient boosted decision tree (GBDT) method can handle categorical features properly. The main difference between GBDT model and CatBoost is that, instead of preprocessing time, CatBoost deals with categorical features during training time. Prokhorenkova et al[33], introduced target statistics (TS) as an efficient strategy for handling categorical features while losing minimum information. In particular, CatBoost permutes the dataset for each example and calculates an average label value based on the category value placed previously in the permutation. Moreover, CatBoost is different from GBDT in terms of feature combinations. Almost all categorical features should be combined to make a new one. CatBoost considers combinations greedily when building a new split for the tree. For the second and following splits, CatBoost combines all combinations.

preset with all categorical features. Every split in the tree is considered a category with two values and they are added together in combination. Additionally, compared to GBDT, CatBoost performs an unbiased boosting with categorical features. To convert categorical features into numerical values with the TS method, the distribution will vary from the original one. In traditional GBDT methods, the deviation of this distribution will result in a deviation in the solution, which is

an inevitable problem for GBDT. A random permutation of the training data is generated in CatBoost. To improve the robustness of the algorithm, multiple permutations will be used by sampling a random permutation and obtaining its slope. Calculating statistics based on permutations is similar to those calculated for classification features. Different permutations are used to train distinct models, and hence, using multiple permutations will not lead to overfitting[34]. Figure 3 shows a schematic of CatBoost model.

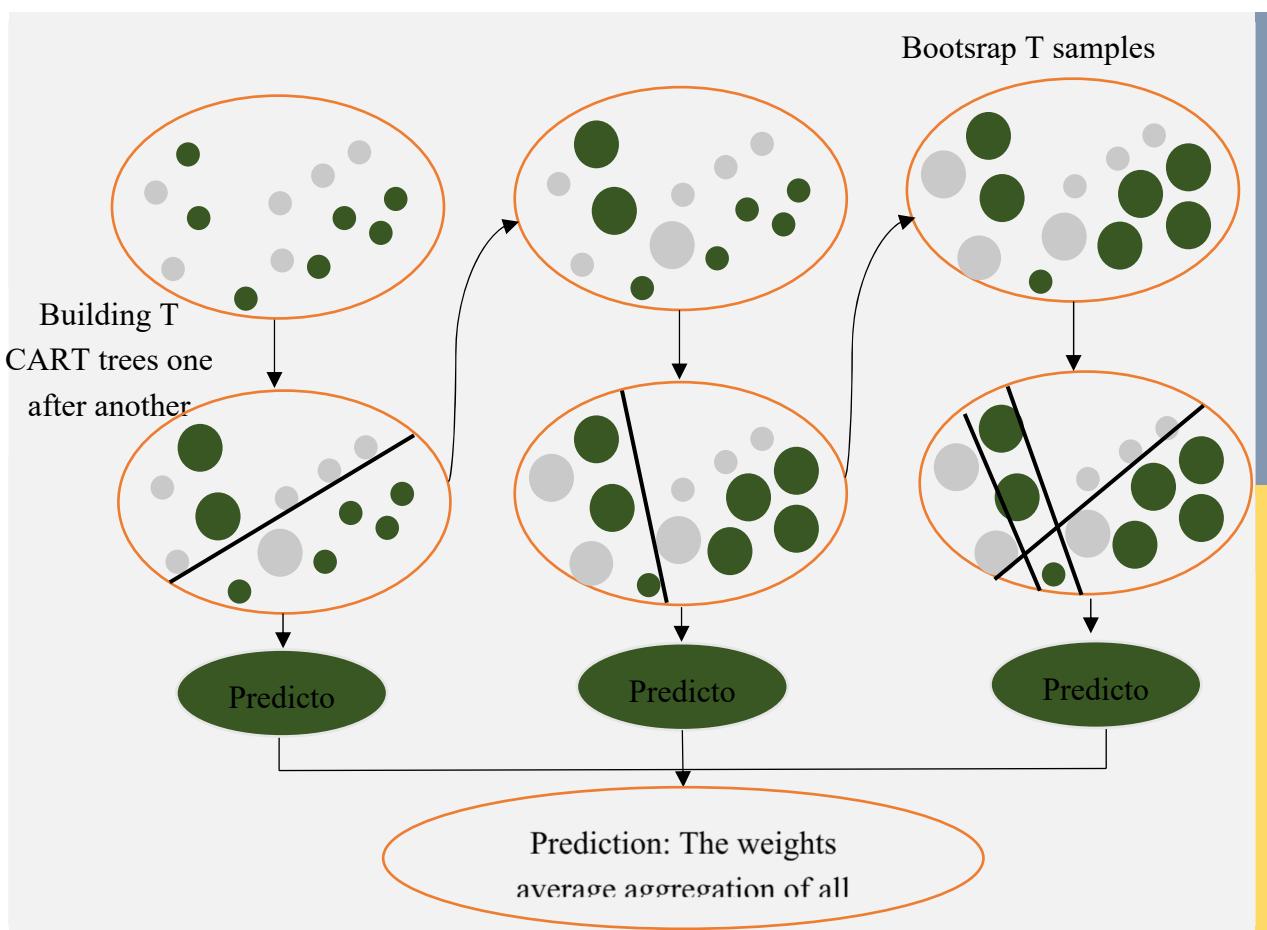


Figure 3. Schematic of the CatBoost algorithm.

2.1.2. Random Forest

Breiman first developed RF [35], the initial goal of RF algorithm development was to solve unsupervised regression and classification problems. This technique involves building multiple independent decision trees, also referred to as ensemble trees, training them based on the desired dataset, and then predicting the target parameter. In this algorithm, bootstrap resampling is used to prevent overfitting, a resampling method that relies on replacement. A bootstrap set is created by replacing several samples with repeated samples from the initial data. The RF algorithm then builds each tree using a bootstrap set. Therefore, since the trees were constructed on varied datasets, their predictions would be different. The next step is to aggregate all the trees, and the final prediction is obtained by averaging each tree's predictions. With the RF model, the degree of importance of each feature and the proximity of samples in pairs can be determined[36,37].

2.1.3. XGBoost

XGBoost is a popular boosting tree algorithm based on a decision tree, also known as a classification and regression tree (CART)[38]. CART divides the dataset into two subsets at each level according to the boundary of a variable for regression tasks, until it reaches the maximum tree depth

specified by users. Searching for the best solutions is done by algorithm for a range of variables to minimize the cost function. Then the prediction is the average of the target value of all samples in a subset. CART trees can be prone to overfitting without proper regularization. One strategy for this is called ensemble packing of a group of estimators, in other words multiple CART models. XGBoost continues to add and train new trees to accommodate the remaining errors from the last iteration. Then a predicted value is assigned to each sample by summing all the scores of the corresponding leaves together. The advantage of XGBoost in performance is its reliable objective function for tree creation[39,40].

2.1.4. LightGBM

LightGBM is a new GBDT algorithm that was first released by Microsoft by Ke et al [41], in 2017 that has been employed in a variety of data mining tasks such as regression, classification, and sorting. The LightGBM algorithm incorporates several novel techniques, including gradient-based one-side sampling (GOSS), exclusive feature bundling (EFB), and a depth-constrained histogram and leaf-based growth strategy. Light GBM grows the tree vertically, while other algorithms, such as XGBoost and GBDT, grow trees horizontally. The mechanism of GOSS involves retaining all large gradient samples while performing random sampling on small gradient samples, based on their proportion. The main idea of EFB is to divide the features into a smaller number of unique mutual bundles.

2.1.5. Evaluation Metrics

The performance of the conductivity prediction models was evaluated by the following metrics: mean absolute error (MAE), root mean square error (RMSE), and determination coefficient (R²). Commonly, these metrics are used in regression tasks, the higher the R² value and the lower the MAE and RMSE value, the better would be the accuracy of the models.

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

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3)$$

2.2. Data Gathering and Model Developments

In this study, SPEs and GPEs mediated with imidazolium-based polymerized ionic liquids were gathered from the literature. For developing robust ML models, a total of 120 datapoints were collected from the literature using WebPlot-Digitizer[42,43], [44–49]. The collected dataset contains SMILES, temperature, and conductivity (Table 2).

Table 2. Collected data from the literature for conductivity prediction.

Author	PIL-name	Number of data	Temperature (K)	Conductivity
[42]	P(EtVIm-TFSI) (NR)	21	298.2-353.3	6.4E ⁻⁶ -4.4E ⁻⁴
[43]	VEIm-TFSI	40	303.1-373.2	9.83E ⁻⁹ -2.41E ⁻⁴
[44]	P-20	7	297.8-352.7	2.9E ⁻⁴ -1.2E ⁻³
[45]	PIL-QSE	7	285.1-358.2	7.1E ⁻⁴ -3.7E ⁻³
[46]	MIm-TFSI + Li-TFSI/EMIm-TFSI	9	301.5-363.2	1.4E ⁻⁵ -6.8E ⁻³
[49]	PVIMTFSI-co-PPEGMA/LiTFSI	6	333-357.8	3.8E ⁻³ -6.6E ⁻³
[47]	HPILSE	23	252.9-353.2	4E ⁻⁵ -5.3E ⁻³
[48]	PIL-GPE	7	298.1-353.2	1.2E ⁻³ -5.3E ⁻³

Figure 4, displays the distribution of ionic conductivity across three temperature ranges 25-65 °C, 65-85 °C, and >85 °C. According to the density plot in Figure 3 (a), in the temperature range of 25 to 65 °C, variability in ionic conductivity is shown, which is visible in the graph with a wide and multi-modal distribution in the plot. Also, the curve shows a long tail toward low conductivity values, with a pronounced peak around 10-5 S/cm. The curve for 65-85 °C group is narrower and more concentrated than 25-65 °C group. In this curve fewer samples fell into the low-conductivity regime. The > 85 °C group shows a very sharp and narrow peak.

By using the RDKit library [50], the molecular descriptors were generated from SMILES representations. Open-source cheminformatics toolkit RDKit, allows the calculation of a wide range of chemical descriptors. Physicochemical and structural features in RDKit are calculated from SMILES strings and then used as quantitative descriptors for ML analysis. Molecular properties including topological polar surface area, number of H₂ bond, and molecular weight captured by molecular descriptors, allow for a better understanding of structure-property relationships. Then ML

models including: CatBoost, RF, XGBoost, and LightGBM were employed in python environment to predict conductivity.

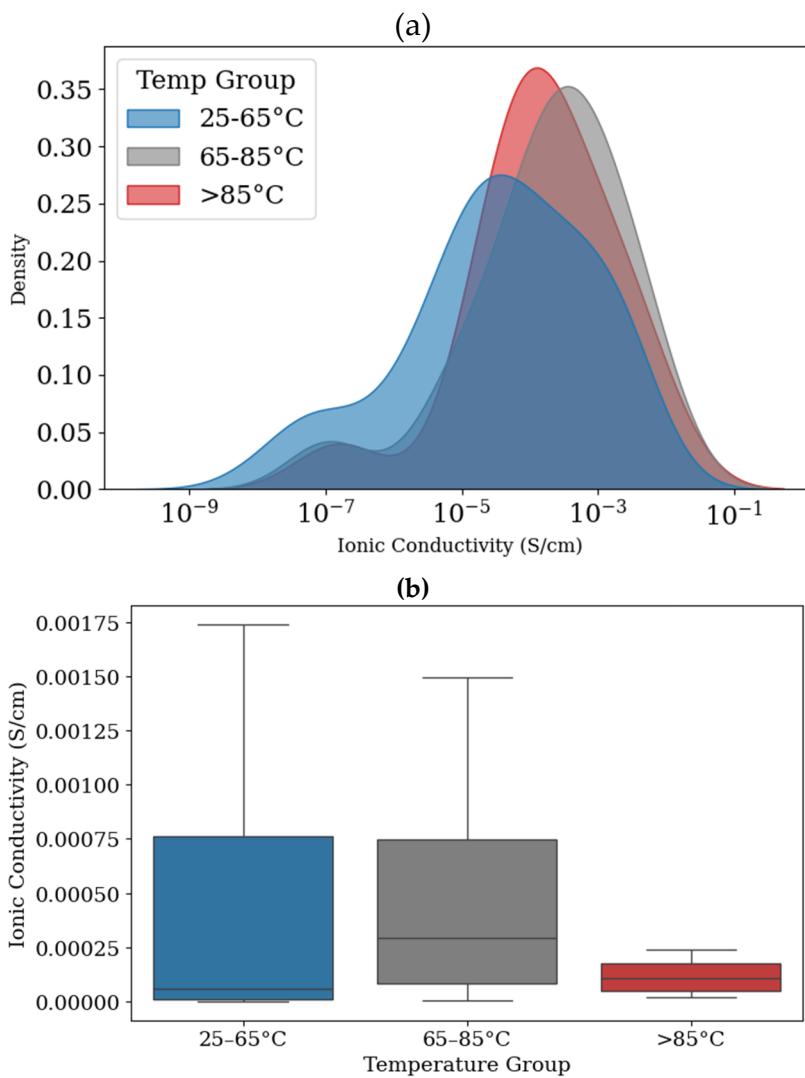


Figure 4. (a) density plot of ionic conductivity in three temperature ranges (25-65 °C, 65-85 °C, and >85 °C) (b) box plot of ionic conductivity across three temperature ranges.

Based on SMILES representations, RDKit generated 434 molecular descriptors which captured electronic, structural, and physicochemical properties of the molecules. In order to eliminate the effect of highly correlated features, Pearson's correlation coefficient was used to calculate the correlation between all pairs of descriptors. After the elimination, a total of 43 features remained. It is worth noting that the PIL prefix has been added to the descriptors. The heatmap plot in Figure 5 shows pairwise correlations between two descriptors, where each cell represents the strength of the relationship between the two features (descriptors). In this plot, each row or column depicts a pairwise linear relationship between features. In each row and column, a different feature is represented, and the colors differentiate their strength and direction based on Pearson's correlation coefficient. When the value is close to +1, it is described in dark red, which shows a strong positive linear correlation between the two features, which means that by increasing one, the other will increase too. A value close to -1, represented in dark blue, indicates that when one feature increases, the other decreases. However, values close to 0 shows little or no linear relationship. In this heatmap plot as it is evidence, features such as PIL_PEOE_VSA, PIL, SMR_VSA and PIL_EState have strong correlations, since they belong to similar molecular descriptor families. There are some features that have relatively medium to high correlations with the target parameter, these features are namely:

PIL_MaxAbsEStateIndex, PIL_PEOE_VSA1, PIL_PEOE_VSA2, PIL_SMR_VSA6, and PIL_fr_methoxy.

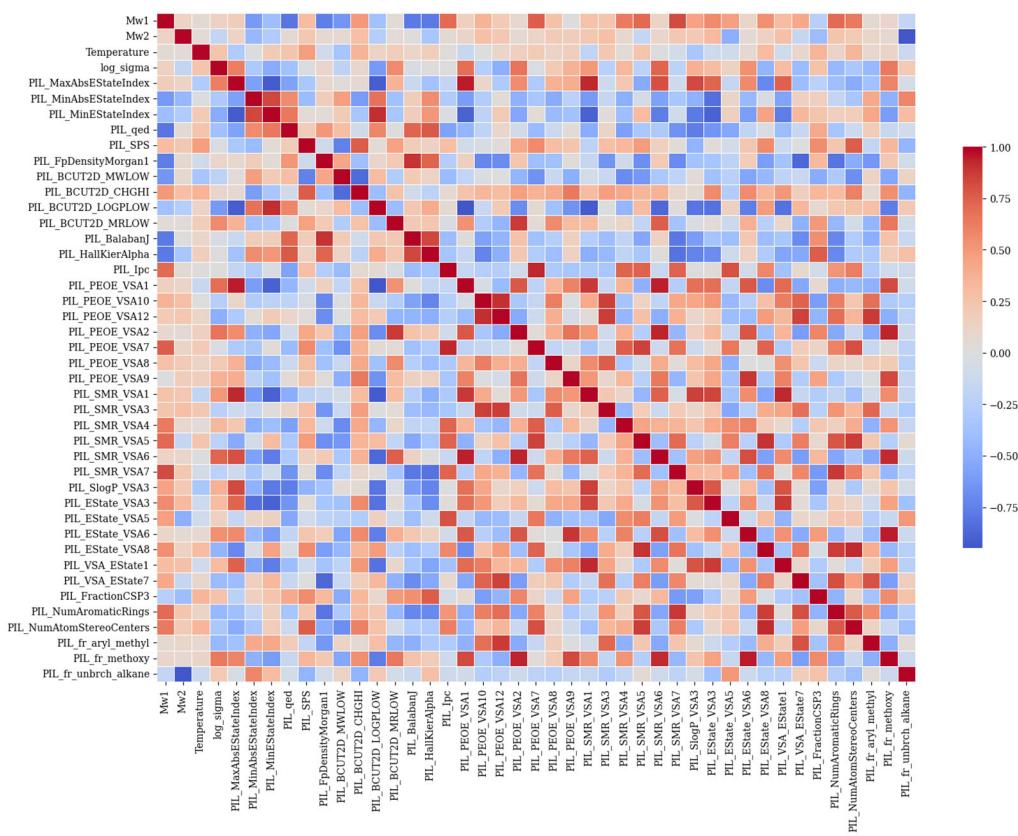


Figure 5. Heatmap of Pearson correlation coefficients between molecular descriptors generated by RDKit.

The data were standardized to make sure the data are transformed into a common scale before feeding them into ML models. On the training data, GridSearchCV and K-fold cross-validation were utilized to optimize model hyperparameters, and to ensure that the random state was set for reproducible cross-validation splits. Tuning of the model hyperparameters was done by using GridSearchCV to find the best parameters for each model. Following the selection of hyperparameters, different algorithms were compared using five-fold cross validation. The data was divided into train and test splits, 80% for train and 20% for test. Each test set included polymers that were not present in the train set. Four independent models were trained by using different random initial values, on 80% of the data. The mean R^2 , the mean RMSE, and the mean MAE were all evaluated as scoring criteria. To ensure that each polymer appeared only once in a test set, this process was repeated five times.

3. Results

R^2 , RMSE, and MAE for the employed models are presented in Table 3. CatBoost performed better than other models, followed closely by RF. The performance of XGBoost is slightly worse than that of RF, while LightGBM had the worst performance. Figure 6, compares the scatter plots of the actual versus predicted conductivity values for different employed ML models: CatBoost, RF, XGBoost, and LightGBM. In this plot dashed black line corresponds to the ideal prediction scenario where predicted values perfectly match the actual values ($y = x$). Hence, the closer the points are to the $y = x$ line the better the models are. For CatBoost model Figure 6 (a), training data points are closely aligned with diagonal line, testing data points also align with the diagonal line, but they are slightly more spread than training data which indicates some error prediction for testing data.

Therefore, CatBoost performs well showing high accuracy for both training and testing datapoints, however the slight spread of testing datapoint suggest reliable generalization with small overfitting. Figure 6 (b), the blue (training) points are aligned with the unity ($y = x$) line, suggesting good fit of the model over blue datapoints. The performance of the model over test datapoints is more scattered comparing to CatBoost, which indicates higher prediction errors over unseen data. According to Figure 6 (c), the training datapoints are no longer align well with $y = x$ line, the deviation from this line is further evidence for XGBoost model (comparing to two previous models) which suggests not very good performance of this model over training data. For testing data this dispersion is even more, especially for higher values of conductivity, revealing that predictions might be less accurate for higher values of conductivity. Among all of the models, LightGBM performs the least accurate model, Figure 6 (d), with higher deviations for training and testing datapoints. This could be a sign for difficulty of the model in capturing complex relationships for both training and testing data points. The heatmaps in Figure 6, shows the prediction errors of the test data across the four ML models: CatBoost, RF, XGBoost, and LightGBM. This plot highlights the area were the models show systematic errors. As can be seen in Figure 6 (a), the heatmap is mainly concentrated along the diagonal line, which interprets as predicted values.

Table 3. Models' performance metrics.

	CatBoost			RF			XGBoost			LightGBM		
	Train	Test	All	Train	Test	All	Train	Test	All	Train	Test	All
R ²	0.994	0.949	0.986	0.976	0.97	0.975	0.962	0.905	0.952	0.878	0.911	0.884
RMSE	1.2E ⁻⁴	3.35E ⁻⁴	1.87E ⁻⁴	2.55E ⁻⁴	2.57E ⁻⁴	2.56E ⁻⁴	3.2E ⁻⁴	4.5E ⁻⁴	3.55E ⁻⁴	5.81E ⁻⁴	4.41E ⁻⁴	5.56E ⁻⁴
MAE	7.33E ⁻⁵	1.83E ⁻⁴	9.52E ⁻⁵	9.5E ⁻⁵	1.26E ⁻⁴	1E ⁻⁴	2E ⁻⁴	2.54E ⁻⁴	2.14E ⁻⁴	3.54E ⁻⁴	3.28E ⁻⁴	3.4E ⁻⁴

are close to actual values. This plot also shows a narrow spread, meaning that the model is making relatively few incorrect predictions with only few outliers. There are some minor deviations for higher conductivity values, but in general the overall distribution is narrow. The systematic bias of CatBoost is minimal across different conductivity ranges. The heatmap of RF, Figure 7 (b), is more spread especially at medium to high conductivity values comparing to CatBoost model. According to this plot, the model tends to underpredict high conductivity values. Similar to CatBoost and RF, the heatmap of XGBoost Figure 7 (c), shows a high density along the diagonal, however a few smaller clusters appear farther from the diagonal line, this could be due to subsets of data where the model performs poorly in predicting the values correctly, also the error region is wider than previous models, which shows deviations for higher conductivity values. The heatmap plot of LightGBM model Figure 7 (d), shows good predictive performance but it is more spread at higher conductivity values. Possibly due to underfitting, LightGBM and XGBoost make larger errors in higher conductivity regions. Comparing to CatBoost and RF, LightGBM and XGBoost seem to struggle more with high-range values.

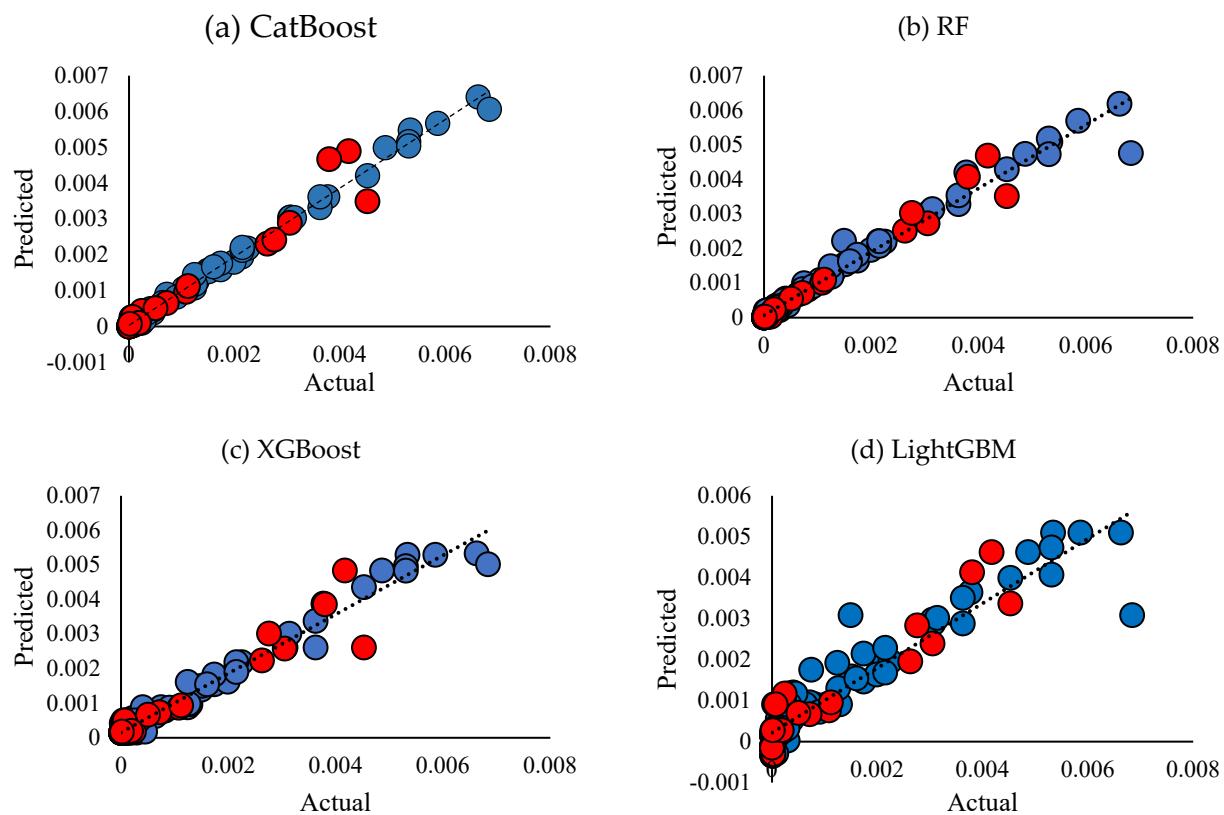


Figure 6. Cross plots of actual vs predicted ionic conductivity with respect to training and testing subsets, (a) CatBoost, (b) RF, (c) LightGBM, (d) XGBoost.

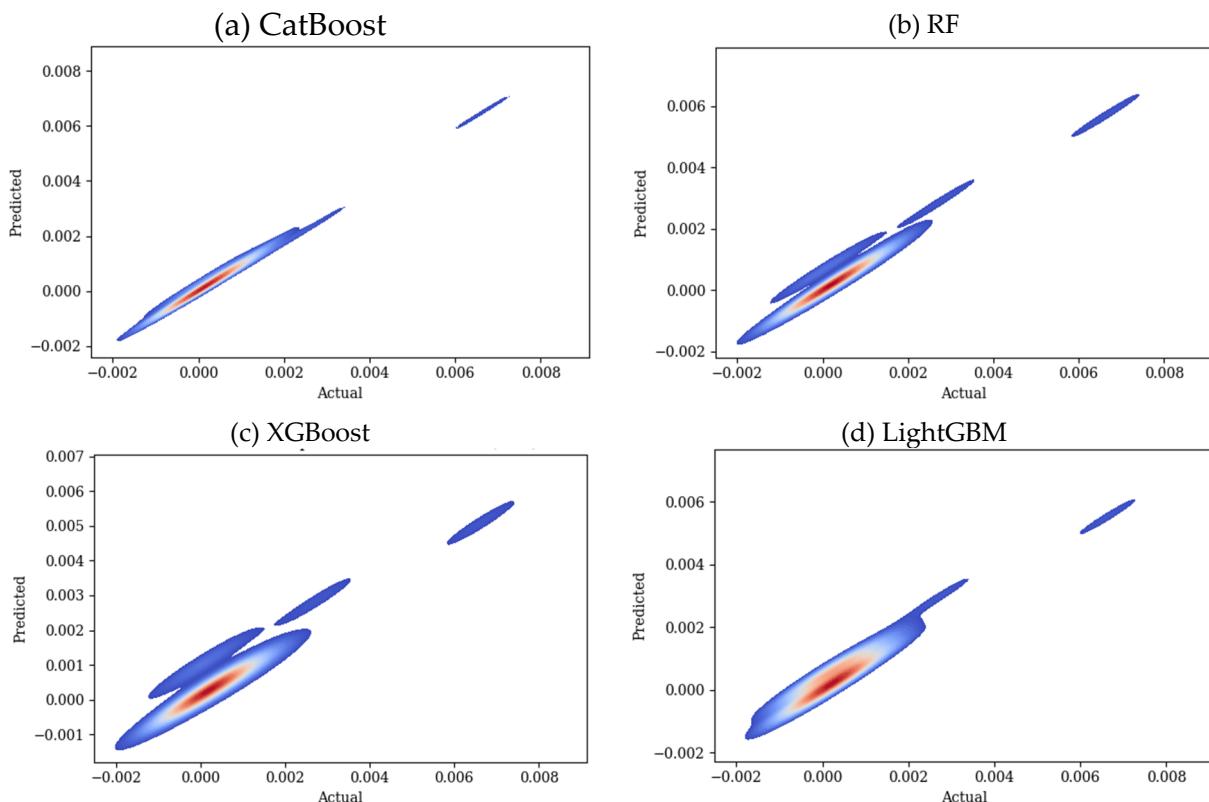
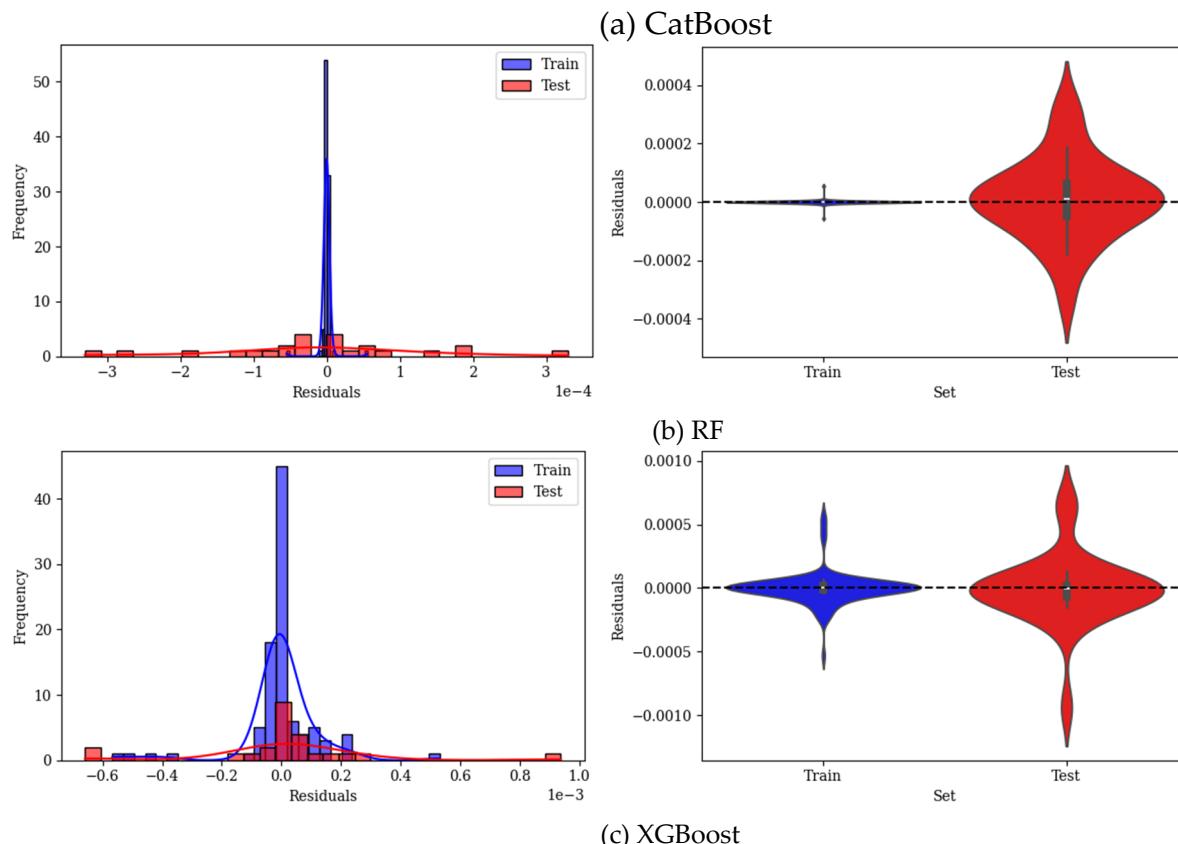


Figure 7. Heatmaps of prediction error for test data across four ML models: (a) CatBoost, (b) RF, (c) XGBoost, (d) LightGBM.

The residual histograms and violin plots in Figure 8, provide analysis of the errors for each of the employed ML models. In residual plots the centered and symmetric distribution indicates good performance of the models with no overprediction or underprediction. A narrower spread, skewness or long tails show higher accuracy and systematic bias, respectively. The residual histogram plot of the CatBoost model is narrow and strongly centered around zero, as shown in Figure 8 (a), which highlights the small errors of this model, also the spread of this model is symmetric with very few residuals. The violin plot of train residual is highly concentrated and packed around zero, however, the residual of test data shows higher dispersion. Moreover, train residuals are tight which shows good fit of the model over training data. Overall, this model exhibits minimal overfitting and good generalization with slightly higher error in the test data. Figure 8 (b), the residual plot of RF shows wider spread as well as longer tail on both sides. Additionally, this model shows more residuals far from zero, which interprets as higher variance in errors. Violin plots of this model shows wider spread for test residuals than train residuals. In general train residuals are highly packed, and test datapoints show higher variance. The residual histogram of XGBoost, Figure 8 (c), has an overall wider spread than previous models, there are some clear deviations in small residuals. Moreover, the plot is shifted towards small residual (right-skewed). The residual spread of this model as can be seen in violin plots is wider than CatBoost and RF for both train and test sets. LightGBM residuals, Figure 8 (d), it has a normal distribution and is centered around zero while train data are more dispersed than test data. Violin plot of train and test is also similar to CatBoost and RF, while it is more dispersed than those models.



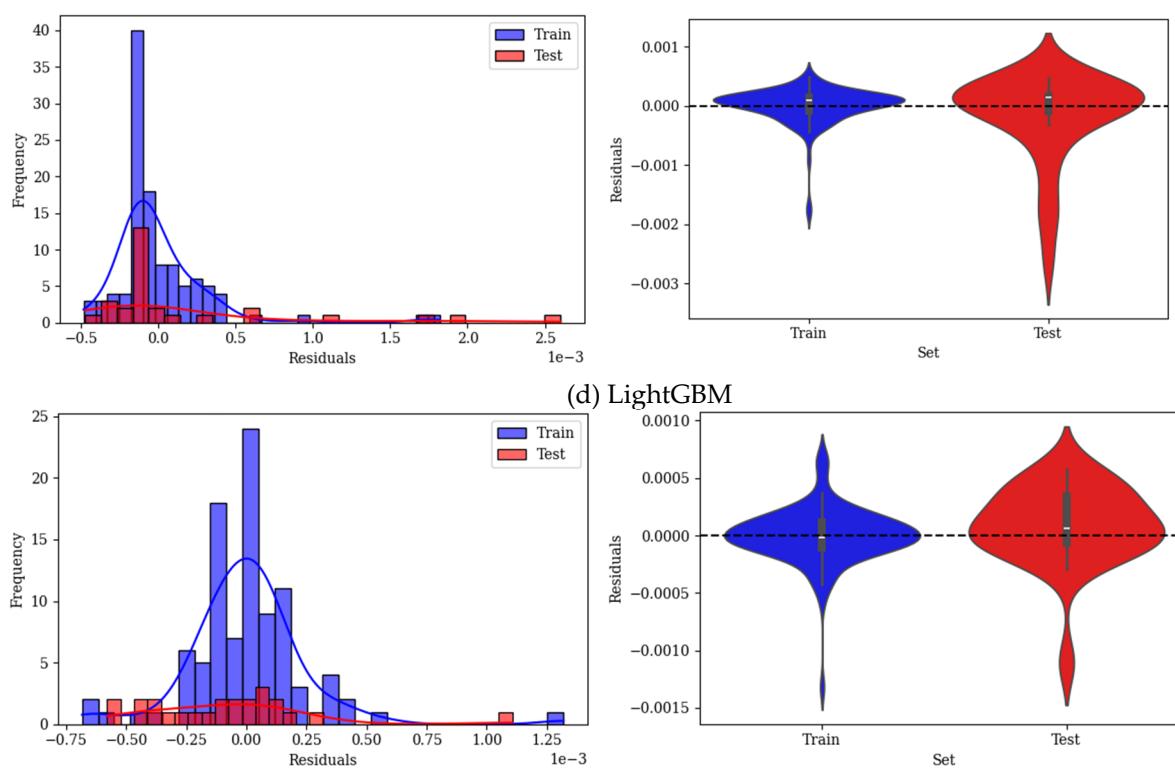


Figure 8. Residual analysis of ML models (a) CatBoost, (b) RF, (c) XGBoost, (d) LightGBM.

3.1. Feature Importance

In this section, the analysis of the features affecting conductivity prediction was examined. For this purpose, the CatBoost model was used, which was introduced as the best model. According to Figure 9, the feature importance plot, temperature is identified as the most important variable in the prediction of conductivity, this finding is in line with physical expectations, typically increasing temperature increases ionic mobility, which leads to higher conductivity, especially in electrolytes and polymer systems. Among the derived molecular descriptors obtained by RDKit library, PIL_BCUT2D_MRLOW showed the second most influential feature. This parameter demonstrates a topological descriptor which is mainly based on molecular refractivity and atomic partial charges. The next important feature is, PIL_SMR_VSA6, which is a descriptor that shows an approximate value of the Van der Waals surface area of atoms which have specific SMR (molar refractivity) values. In general, these two high ranked features reflect electron distribution characteristics of molecules, size, and shape. Another important feature in conductivity prediction is PIL_EState_VSA8, which includes electro topological state indices with the contributions of surface area. PIL_qed is the other influential descriptor, where this feature is based on properties like polarity and molecular weight. PIL_fr_unbrch_alkane, which is a measure of the unbranched alkane fragments in a molecule, highlights the importance of branching of molecules in predicting the target variable. Electronic properties which are derived from the charge distribution are represented by PIL_BCUT2D_CHGHI. This feature is followed by PIL_PEOE_VSA8, which contains the surface area corresponding to partial atomic charges, calculated via PEOE (partial equalization of orbital electronegativities). As it is evident from the graph, there are also some lower rank descriptors such as PIL_MinAbsEStateIndex, PIL_SlogP_VSA3, and PIL_FractionCSP3. The mentioned features show small impact on the model's performance, since they reflect small details about topology of the molecules and their electronic structure. Descriptors with the least influence are placed at the bottom of the plot, which show almost zero or no importance.

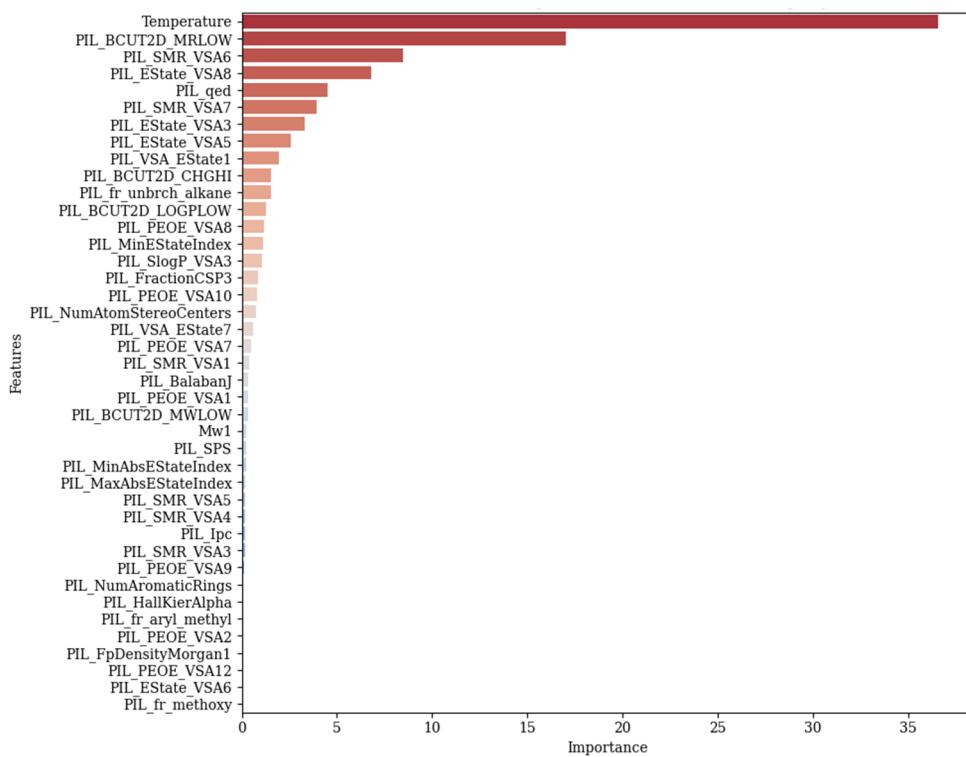


Figure 9. Feature importance plot for predicting conductivity employing CatBoost model.

3.2. Predicting Conductivity for Ionenes

Figure 10 shows predicted conductivity values of three ionenes, namely: TC-API(p)-Xy, Troger's base (Im-TB(p)-PA), and Ionic-polyimide, at random temperatures: 298, 308, 318. This plot shows how increasing temperature can increase ionic conductivity. As it is evident in this plot, each ionene linear line is decreasing by increasing $1000/T$, this can be interpreted as conductivity rises with temperature increases, which indicates that the model has learned the underlying thermally activated transport behavior. The ionic conduction in these materials is governed by Arrhenius kinetics, which implies that ion mobility increases with thermal energy. It is apparent that all the selected ionenes exhibited outstanding electrical properties with significant ionic conductivity (approaching 10^{-3} S cm $^{-1}$). Across the entire temperature range, Troger's base (Im-TB(p)-PA) shows the highest conductivity values. Following this, ionene TC-API(p)-Xy and ionic polyimide exhibit relatively lower conductivity values. These nearly parallel lines suggest similar activation energies for ion transport in all three materials. Vertical separation between these three lines indicates that a difference in their intrinsic conductance exists due to factors such as structural features, ion mobility, and carrier concentration. According to these results, Troger's base may offer the best performance for applications such as electrolytes and ion-exchange membranes, which require high ionic conductivity, thereby paving the way for a new strategy in designing imidazolium ionenes with rigid backbones. In addition to capturing the overall temperature dependence of these ionenes, the ML models also captured their relative ranking, likely based on learned structural and electronic descriptors. In this regard, the model can help predict conductivity trends for new candidates of ionenes.

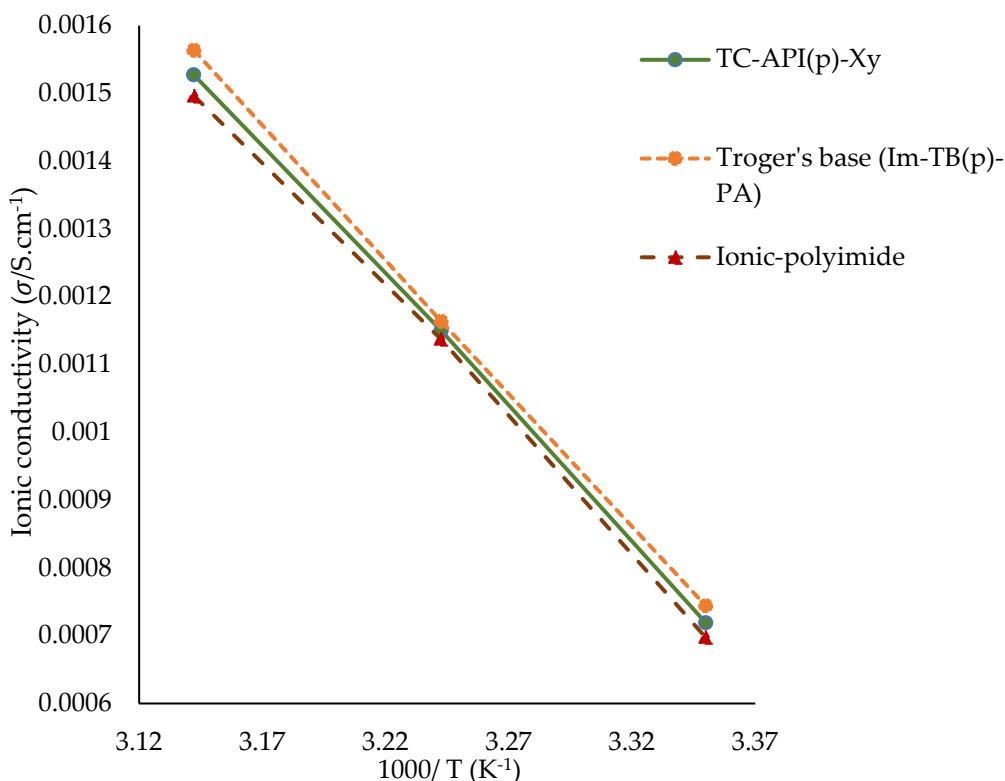


Figure 10. Predicted conductivity values for three ionenes using the best performing model (CatBoost).

4. Conclusions

In this study four ML models were developed for the prediction of ionic conductivity. The models were trained and tested on the data gathered from the literature. The developed models were CatBoost, RF, XGBoost, and LightGBM. There was a significant difference in the performance of the models for predicting conductivity values. Across both training and testing datasets, CatBoost proved to be the most accurate and reliable model. Scatter plots, heatmaps, and residual plots demonstrated the superior performance of this model. According to the mentioned plots, the CatBoost model showed consistent predictions with the actual values. Following closely, RF showed excellent performance for unseen data, although it exhibited higher prediction errors at medium and higher conductivity levels. LightGBM and XGBoost performed poorly, with LightGBM having difficulty in capturing temperature-dependent trends. Ultimately, the best model (CatBoost) was used to analyze the effect of input parameters on conductivity prediction. Based on the feature importance diagram, temperature was identified as the feature with the greatest impact on conductivity. Molecular descriptors extracted from SMILES were ranked in the following order of importance. Descriptors such as PIL_BCUT2D_MRLOW, PIL_SMR_VSA6, and PIL_EState_VSA8. Moreover, the best-performing model (CatBoost) was used to predict conductivity for three ionenes. Interestingly, all the selected ionenes exhibited outstanding electrical properties, with significant ionic conductivity (approaching 10⁻³ S cm⁻¹). Notably, ionene with a rigid and contorted structure of Troger's base (Im-TB(p)-PA) exhibited the highest conductivity. Additionally, the model captured the overall temperature dependency of the employed ionenes. In general, CatBoost stands out as the most robust and reliable model for conductivity prediction, making it well-suited for ionic-mediated polymers with similar applications in the future.

Author Contributions: Conceptualization, G.P. and I.K.; Methodology, Software, and Visualization, G.P. and I.K.; Writing – Original Draft Preparation, G.P. and I.K.; Writing – review & editing, I.K.; Funding Acquisition, I.K.

Acknowledgment: Support for this work provided by Nazarbayev University under the Collaborative Research Project Grant (Grant No: 111024CRP2017) is gratefully acknowledged.

References

1. K. Kohzadvand, M. M. Kouhi, A. Barati, S. Omrani, and M. Ghasemi, "Prediction of interfacial wetting behavior of H₂/mineral/brine; implications for H₂ geo-storage," *J Energy Storage*, vol. 72, Nov. 2023, doi: 10.1016/j.est.2023.108567.
2. G. Piroozi, M. Kouhi, A. S.-J. of E. Storage, and undefined 2025, "Novel intelligent models for prediction of hydrogen diffusion coefficient in brine using experimental and molecular dynamics simulation data: Implications for," *Elsevier G Piroozi, MM Kouhi, A Shafei Journal of Energy Storage, 2025•Elsevier*, Accessed: May 13, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352152X25010102>
3. M. R. N. Vilanova, A. T. Flores, and J. A. P. Balestieri, "Pumped hydro storage plants: a review," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 42, no. 8, pp. 1–14, Aug. 2020, doi: 10.1007/S40430-020-02505-0/TABLES/2.
4. G. L. Soloveichik, "Flow Batteries: Current Status and Trends," *Chem Rev*, vol. 115, no. 20, pp. 11533–11558, Oct. 2015, doi: 10.1021/CR500720T/ASSET/IMAGES/LARGE/CR-2014-00720T_0013.JPG.
5. M. Li, J. Lu, Z. Chen, and K. Amine, "REVIEW 1800561 (1 of 24) 30 Years of Lithium-Ion Batteries," 2018, doi: 10.1002/adma.201800561.
6. T. Kim, W. Song, D. Y. Son, L. K. Ono, and Y. Qi, "Lithium-ion batteries: outlook on present, future, and hybridized technologies," *J Mater Chem A Mater*, vol. 7, no. 7, pp. 2942–2964, Feb. 2019, doi: 10.1039/C8TA10513H.
7. A. Manthiram, "An Outlook on Lithium-Ion Battery Technology," *ACS Cent Sci*, vol. 3, no. 10, pp. 1063–1069, Oct. 2017, doi: 10.1021/ACSCENTSCI.7B00288/ASSET/IMAGES/LARGE/OC-2017-002885_0004.JPG.
8. J. Xia, R. Petibon, D. Xiong, L. Ma, and J. R. Dahn, "Enabling linear alkyl carbonate electrolytes for high voltage Li-ion cells," *J Power Sources*, vol. 328, pp. 124–135, Oct. 2016, doi: 10.1016/J.JPOWSOUR.2016.08.015.
9. L. Long, S. Wang, M. Xiao, and Y. Meng, "Polymer electrolytes for lithium polymer batteries," *J Mater Chem A Mater*, vol. 4, no. 26, pp. 10038–10069, Jun. 2016, doi: 10.1039/C6TA02621D.
10. M. Zhu *et al.*, "Recent advances in gel polymer electrolyte for high-performance lithium batteries," *Journal of Energy Chemistry*, vol. 37, pp. 126–142, Oct. 2019, doi: 10.1016/J.JECHEM.2018.12.013.
11. T. P. Barrera *et al.*, "Next-Generation Aviation Li-Ion Battery Technologies-Enabling Electrified Aircraft," *Electrochemical Society Interface*, vol. 31, no. 3, pp. 69–74, Sep. 2022, doi: 10.1149/2.F10223IF/XML.
12. K. Vishweswariah, A. K. Madikere Raghunatha Reddy, and K. Zaghib, "Beyond Organic Electrolytes: An Analysis of Ionic Liquids for Advanced Lithium Rechargeable Batteries," *Batteries 2024, Vol. 10, Page 436*, vol. 10, no. 12, p. 436, Dec. 2024, doi: 10.3390/BATTERIES10120436.
13. N. Chawla, N. Bharti, and S. Singh, "Recent Advances in Non-Flammable Electrolytes for Safer Lithium-Ion Batteries," *Batteries 2019, Vol. 5, Page 19*, vol. 5, no. 1, p. 19, Feb. 2019, doi: 10.3390/BATTERIES5010019.
14. J. W. Fergus, "Ceramic and polymeric solid electrolytes for lithium-ion batteries," *J Power Sources*, vol. 195, no. 15, pp. 4554–4569, Aug. 2010, doi: 10.1016/J.JPOWSOUR.2010.01.076.
15. E. Peled, D. Golodnitsky, G. Ardel, C. Menachem, D. Bar Tow, and V. Eshkenazy, "The Role of Sei in Lithium and Lithium-Ion Batteries," *MRS Online Proceedings Library (OPL)*, vol. 393, p. 209, 1995, doi: 10.1557/PROC-393-209.
16. K. Vishweswariah, A. M. R. R.- Batteries, and undefined 2024, "Beyond Organic Electrolytes: An Analysis of Ionic Liquids for Advanced Lithium Rechargeable Batteries," *mdpi.comK Vishweswariah, AK Madikere Raghunatha Reddy, K Zaghib Batteries, 2024•mdpi.com*, Accessed: Feb. 26, 2025. [Online]. Available: <https://www.mdpi.com/2313-0105/10/12/436>
17. F. Gebert, M. Longhini, F. Conti, and A. J. Naylor, "An electrochemical evaluation of state-of-the-art non-flammable liquid electrolytes for high-voltage lithium-ion batteries," *J Power Sources*, vol. 556, p. 232412, Feb. 2023, doi: 10.1016/J.JPOWSOUR.2022.232412.
18. T. Stettner, F. C. Walter, and A. Balducci, "Imidazolium-Based Protic Ionic Liquids as Electrolytes for Lithium-Ion Batteries," *Batter Supercaps*, vol. 2, no. 1, pp. 55–59, Jan. 2019, doi: 10.1002/BATT.201800096.

19. J. Lu, F. Yan, and J. Texter, "Advanced applications of ionic liquids in polymer science," *Prog Polym Sci*, vol. 34, no. 5, pp. 431–448, May 2009, doi: 10.1016/J.PROGPOLYMSCI.2008.12.001.
20. Y. Liu, T. Zhao, W. Ju, and S. Shi, "Materials discovery and design using machine learning," *Journal of Materomics*, vol. 3, no. 3, pp. 159–177, Sep. 2017, doi: 10.1016/J.JMAT.2017.08.002.
21. D. Yong and J. U. Kim, "Accelerating Langevin Field-Theoretic Simulation of Polymers with Deep Learning," *Macromolecules*, vol. 55, no. 15, pp. 6505–6515, Aug. 2022, doi: 10.1021/ACS.MACROMOL.2C00705/SUPPL_FILE/MA2C00705_SI_002.ZIP.
22. H. Doan Tran *et al.*, "Machine-learning predictions of polymer properties with Polymer Genome," *J Appl Phys*, vol. 128, no. 17, p. 171104, Nov. 2020, doi: 10.1063/5.0023759/1062836.
23. N. Andraju, G. W. Curtzwiler, Y. Ji, E. Kozliak, and P. Ranganathan, "Machine-Learning-Based Predictions of Polymer and Postconsumer Recycled Polymer Properties: A Comprehensive Review," *ACS Appl Mater Interfaces*, vol. 14, no. 38, pp. 42771–42790, Sep. 2022, doi: 10.1021/ACSAMI.2C08301/ASSET/IMAGES/LARGE/AM2C08301_0008.JPG.
24. J. Shi, F. Albreiki, N. Yamil J. Colón, S. Srivastava, and J. K. Whitmer, "Transfer Learning Facilitates the Prediction of Polymer-Surface Adhesion Strength," *J Chem Theory Comput*, vol. 19, no. 14, pp. 4631–4640, Jul. 2023, doi: 10.1021/ACS.JCTC.2C01314/SUPPL_FILE/CT2C01314_SI_001.PDF.
25. B. K. Wheatle, E. F. Fuentes, N. A. Lynd, and V. Ganesan, "Design of Polymer Blend Electrolytes through a Machine Learning Approach," *Macromolecules*, vol. 53, no. 21, pp. 9449–9459, Nov. 2020, doi: 10.1021/ACS.MACROMOL.0C01547/ASSET/IMAGES/MEDIUM/MA0C01547_M038.GIF.
26. G. Bradford *et al.*, "Chemistry-Informed Machine Learning for Polymer Electrolyte Discovery," *ACS Cent Sci*, vol. 9, no. 2, pp. 206–216, Feb. 2023, doi: 10.1021/ACSCENTSCI.2C01123/SUPPL_FILE/OC2C01123_SI_001.PDF.
27. Q. Gao, T. Dukker, A. M. Schweidtmann, and J. M. Weber, "Self-supervised graph neural networks for polymer property prediction †," *Cite this: Mol. Syst. Des. Eng.*, vol. 9, p. 1130, 2024, doi: 10.1039/d4me00088a.
28. E. Kazemi-Khasragh, C. González, and M. Haranczyk, "Toward diverse polymer property prediction using transfer learning," *Comput Mater Sci*, vol. 244, p. 113206, Sep. 2024, doi: 10.1016/J.COMMATSCI.2024.113206.
29. A. Babbar, S. Ragunathan, D. Mitra, A. Dutta, and T. K. Patra, "Explainability and extrapolation of machine learning models for predicting the glass transition temperature of polymers," *Journal of Polymer Science*, vol. 62, no. 6, pp. 1175–1186, Mar. 2024, doi: 10.1002/POL.20230714.
30. E. Ascencio-Medina *et al.*, "Prediction of Dielectric Constant in Series of Polymers by Quantitative Structure-Property Relationship (QSPR)," *Polymers (Basel)*, vol. 16, no. 19, p. 2731, Oct. 2024, doi: 10.3390/POLYM16192731/S1.
31. K. Hatakeyama-Sato, T. Tezuka, M. Umeki, and K. Oyaizu, "AI-Assisted Exploration of Superionic Glass-Type Li⁺ Conductors with Aromatic Structures," *J Am Chem Soc*, vol. 142, no. 7, pp. 3301–3305, Feb. 2020, doi: 10.1021/JACS.9B11442/SUPPL_FILE/JA9B11442_SI_001.PDF.
32. K. Li, J. Wang, Y. Song, and Y. Wang, "Machine learning-guided discovery of ionic polymer electrolytes for lithium metal batteries," *Nat Commun*, vol. 14, no. 1, Dec. 2023, doi: 10.1038/s41467-023-38493-7.
33. S. Ibrahim and M. R. Johan, "Conductivity, Thermal and Neural Network Model Nanocomposite Solid Polymer Electrolyte S LiPF6)," *Int J Electrochem Sci*, vol. 6, no. 11, pp. 5565–5587, Nov. 2011, doi: 10.1016/S1452-3981(23)18428-8.
34. F. H. Zhai *et al.*, "A deep learning protocol for analyzing and predicting ionic conductivity of anion exchange membranes," *J Memb Sci*, vol. 642, p. 119983, Feb. 2022, doi: 10.1016/J.MEMSCI.2021.119983.
35. Y. K. Phua, T. Fujigaya, and K. Kato, "Predicting the anion conductivities and alkaline stabilities of anion conducting membrane polymeric materials: development of explainable machine learning models," *Sci Technol Adv Mater*, vol. 24, no. 1, p. 2261833, 2023, doi: 10.1080/14686996.2023.2261833/ASSET/DD76E15A-EC30-460E-829B-86B9C6F41D97/ASSETS/GRAPHIC/TSTA_A_2261833_F0003_OC.JPG.
36. J. Wang and P. Rajendran, "Conductivity Prediction Method of Solid Electrolyte Materials Based on Pearson Coefficient Method and Ensemble Learning," *Journal of Artificial Intelligence and Technology*, Nov. 2024, doi: 10.37965/JAIT.2024.0551.

37. U. Arora, M. Singh, S. Dabade, and A. Karim, "Analyzing the efficacy of different machine learning models for property prediction of solid polymer electrolytes," Jul. 2024, doi: 10.26434/CHEMRXIV-2024-0T7MW.

38. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," *proceedings.neurips.cc*L Prokhorenkova, G Gusev, A Vorobev, AV Dorogush, A GulinAdvances in neural information processing systems, 2018•*proceedings.neurips.cc*, Accessed: Mar. 11, 2025. [Online]. Available: <https://proceedings.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html>

39. J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," *J Big Data*, vol. 7, no. 1, pp. 1–45, Dec. 2020, doi: 10.1186/S40537-020-00369-8/FIGURES/9.

40. L. Breiman, "Random forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

41. A. Parmar, R. Katariya, and V. Patel, "A Review on Random Forest: An Ensemble Classifier," *Lecture Notes on Data Engineering and Communications Technologies*, vol. 26, pp. 758–763, 2019, doi: 10.1007/978-3-030-03146-6_86.

42. M. Belgiu and L. Drăgu, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, Apr. 2016, doi: 10.1016/J.ISPRSJPRS.2016.01.011.

43. T. Chen, T. He, M. Benesty, V. K. version 0.4-2, and undefined 2015, "Xgboost: extreme gradient boosting," *cran.ms.unimelb.edu.au*T Chen, T He, M Benesty, V Khotilovich, Y Tang, H Cho, K Chen, R Mitchell, I Cano, T ZhouR package version 0.4-2, 2015•*cran.ms.unimelb.edu.au*, 2024, Accessed: Mar. 11, 2025. [Online]. Available: <https://cran.ms.unimelb.edu.au/web/packages/xgboost/vignettes/xgboost.pdf>

44. L. Mackey, J. Bryan, M. M.-N. 2014 W. on High, and undefined 2015, "Weighted classification cascades for optimizing discovery significance in the higgsml challenge," *proceedings.mlr.press*L Mackey, J Bryan, MY MoNIPS 2014 Workshop on High-energy Physics and Machine Learning, 2015•*proceedings.mlr.press*, vol. 42, pp. 129–134, 2015, Accessed: Mar. 11, 2025. [Online]. Available: <http://proceedings.mlr.press/v42/mack14.html>

45. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-August-2016, pp. 785–794, Aug. 2016, doi: 10.1145/2939672.2939785.

46. G. Ke, Q. Meng, T. Finley, ... T. W.-A. in neural, and undefined 2017, "Lightgbm: A highly efficient gradient boosting decision tree," *proceedings.neurips.cc*G Ke, Q Meng, T Finley, T Wang, W Chen, W Ma, Q Ye, TY LiuAdvances in neural information processing systems, 2017•*proceedings.neurips.cc*, Accessed: Mar. 11, 2025. [Online]. Available: <https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html>

47. K. Yin, Z. Zhang, L. Yang, and S. I. Hirano, "An imidazolium-based polymerized ionic liquid via novel synthetic strategy as polymer electrolytes for lithium-ion batteries," *J Power Sources*, vol. 258, pp. 150–154, Jul. 2014, doi: 10.1016/J.JPOWSOUR.2014.02.057.

48. F. Ma et al., "Solid Polymer Electrolyte Based on Polymerized Ionic Liquid for High Performance All-Solid-State Lithium-Ion Batteries," *ACS Sustain Chem Eng*, vol. 7, no. 5, pp. 4675–4683, Mar. 2019, doi: 10.1021/ACSSUSCHEMENG.8B04076/ASSET/IMAGES/LARGE/SC-2018-04076A_0008.jpeg.

49. Y. C. Tseng, S. H. Hsiang, C. H. Tsao, H. Teng, S. S. Hou, and J. S. Jan, "In situ formation of polymer electrolytes using a dicationic imidazolium cross-linker for high-performance lithium-ion batteries," *J Mater Chem A Mater*, vol. 9, no. 9, pp. 5796–5806, Mar. 2021, doi: 10.1039/D0TA09249E.

50. T. Huang, M. C. Long, X. L. Wang, G. Wu, and Y. Z. Wang, "One-step preparation of poly (ionic liquid)-based flexible electrolytes by in-situ polymerization for dendrite-free lithium-ion batteries," *Chemical Engineering Journal*, vol. 375, p. 122062, Nov. 2019, doi: 10.1016/J.CEJ.2019.122062.

51. T. L. Chen, R. Sun, C. Willis, B. F. Morgan, F. L. Beyer, and Y. A. Elabd, "Lithium ion conducting polymerized ionic liquid pentablock terpolymers as solid-state electrolytes," *Polymer (Guildf)*, vol. 161, pp. 128–138, Jan. 2019, doi: 10.1016/J.POLYMER.2018.12.013.

52. D. Zhou et al., "In situ synthesis of hierarchical poly (ionic liquid)-based solid electrolytes for high-safety lithium-ion and sodium-ion batteries," *Nano Energy*, vol. 33, pp. 45–54, Mar. 2017, doi: 10.1016/J.NANOEN.2017.01.027.

53. H. Niu et al., "Preparation of imidazolium based polymerized ionic liquids gel polymer electrolytes for high-performance lithium batteries," *Mater Chem Phys*, vol. 293, p. 126971, Jan. 2023, doi: 10.1016/J.MATCHEMPHYS.2022.126971.
54. S. S. More et al., "Imidazolium Functionalized Copolymer Supported Solvate Ionic Liquid Based Gel Polymer Electrolyte for Lithium Ion Batteries," *ACS Appl Polym Mater*, vol. 27, p. 45, Oct. 2024, doi: 10.1021/ACSAPM.4C02103/ASSET/IMAGES/LARGE/AP4C02103_0005.JPG.
55. A. B. Georgescu et al., "Database, Features, and Machine Learning Model to Identify Thermally Driven Metal-Insulator Transition Compounds," *Chemistry of Materials*, vol. 33, no. 14, pp. 5591–5605, Jul. 2021, doi: 10.1021/ACS.CHEMMATER.1C00905/ASSET/IMAGES/LARGE/CM1C00905_0017.JPG.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.