



Predicting Thermophysical Properties of Ionic Liquids Using Deep Neural Networks and Random Forests

Elif Acar* and Sadaf Sobhani†
Cornell University, Ithaca, New York, 14853, USA

This study utilizes deep neural networks and random forests to predict the thermophysical properties of ionic liquids (ILs), focusing on viscosity and liquid phase transition temperature (LPTT) at atmospheric pressure. The goal is to identify novel ILs with superior properties for spacecraft thermal control, surpassing current heat transfer fluids. Two curated datasets were employed: 1,139 ILs for viscosity predictions, incorporating temperature and 210 molecular descriptors as features, and 1,458 ILs for LPTT predictions with 210 molecular descriptors. The study explores various training-test split strategies and model configurations to optimize predictive performance. These optimized models were then applied to predict the properties of 328,740 generated IL permutations, targeting low viscosity and LPTT values below 243 K. This work provides a scalable framework for designing high-performance ionic liquids, contributing to the development of advanced thermal management systems for extreme environments.

I. Introduction

Thermal control subsystems are essential for spacecraft survival and performance, especially in extreme environments like lunar and deep space missions, where components must operate within specific temperature ranges. Active thermal control technologies such as single-phase pumped fluid loops are used in various robotic and human spacecraft [1–4]. Pumped fluid loop cooling was used for robotic missions such as Mars Pathfinder (1996), Mars Exploring Rovers (2003), Mars Science Laboratory (2011), and Mars Rover (2020) [1, 5]. Human spacecraft have utilized single-fluid (e.g., Apollo, Gemini) or two-fluid (e.g., Space Shuttle, ISS, Orion) loop configurations [4]. Two-fluid systems minimize crew exposure to hazardous fluids by isolating them in external loops, while enhancing robustness to handle varying heat loads and external thermal conditions. However, these systems require additional hardware to support separate fluid loops, significantly increasing overall system mass [6–9].

Some of the common fluids used in single-phase pumped fluid loops include water, glycol, and ammonia as well as engineered fluids such as Freon-21, HFE-7200, and Fluorinert FC-72 [5]. Some of these fluids are toxic and have either been discontinued or restricted exclusively to the use of external spacecraft fluid loops. Beyond minimizing toxicity, future missions in extreme space environments necessitate enhanced thermal transport and fluidity properties to meet the system requirements. The development of novel fluids for single-phase pumped fluid loops can enhance operational flexibility, reduce system mass, and further mitigate risks to the crew. For instance, fluids with higher specific heat would improve thermal energy transport by carrying more heat per unit flow, while higher thermal conductivity would enhance the convective heat transfer in heat exchangers, boosting system efficiency [10]. Additionally, fluids with lower viscosity would reduce the pressure drop, thereby reducing the pump power required. Previous efforts to enhance the thermophysical properties of heat transfer fluids have explored nano-particle suspensions, which effectively improve thermal conductivity. However, their significantly higher viscosity necessitates larger, bulkier pump systems, resulting in an overall increase in system-level mass [2]. Thus, careful consideration of the systems-level implications is essential in evaluating any candidate fluids.

Tailored ionic liquids (ILs) are promising candidates as heat transfer fluids due to their high thermal stability at elevated temperatures and low vapor pressure [11, 12]. ILs are composed of ions that are positively charged (i.e., cations) and negatively charged (i.e., anions), which are held together by electrostatic forces. These forces and interactions give rise to the aforementioned unique properties of ILs and allow for tailoring of other fluid properties by careful design of the anion-cation pairing. However, there are 10^8 theoretical pairings, thus traditional computational and experimental methods are prohibitively time and cost intensive for identifying tailored ion species combinations.

To address this challenge, machine learning (ML) methods have recently been proposed. The choice of a ML model for IL predictions ultimately depends on the quality and quantity of the available data and the specific property being

*PhD Student, Sibley School of Mechanical and Aerospace Engineering, AIAA Student Member.

†Assistant Professor, Sibley School of Mechanical and Aerospace Engineering, AIAA Member.

predicted. Among these models, neural networks are widely used due to their ability to handle complex, non-linear, and high-dimensional datasets. Several studies have shown the reliability of neural networks as regression models for IL research. More than 2,000 known ILs are available for model training and validation, with additional ones under ongoing synthesis and characterization. However, some properties of ILs, such as ionic conductivity and viscosity, are more available than others, such as melting point. In fact, given their applicability to energy storage and electrochemical conversion, ILs have primarily been characterized by ionic conductivity, solubility, viscosity, density, and toxicity, followed by some studies on melting point, thermal conductivity, and heat capacity [13]. Datta et al. predicted the ionic conductivity of ILs using deep neural networks (DNN) and investigated the impact of different data-splitting approaches, defined as data-points-split and IL-split [14], achieving coefficients of determination of 0.96 and 0.83, respectively. They also identified key molecular descriptors that are correlated with this property, which can be used in the design of new ILs. Valderrama et al. were among the first to conduct research on predicting the viscosity of ILs using ML methods [15]. They used an artificial neural network (ANN) and achieved satisfactory results, with deviations below 5% for all 31 predictions, despite the limited number of unique ILs and features such as temperature, density, mass of the cation, anion, and mass connectivity index. Padaszyński and Domańska used a feed-forward artificial neural network (FFANN) with temperature, pressure, and group contributions as features to predict the viscosity of 1,484 unique ILs [16]. Their method performed very well overall, achieving a mean squared error of 0.0603 log units and a coefficient of determination of 0.972 on the test set, though lower performance was observed for ILs with considerable data scatter or limited data for certain cations and anions. Following this, Padaszyński compared multiple linear regression (MLR), feed-forward artificial neural network (FFANN), and least-squares support vector machines (LSSVM) methods, using a larger dataset of 2,068 unique ILs and group contributions as features [17]. FFANN and LSSVM performed well on the test set, with coefficients of determination of 0.8637 and 0.8374, but they advised using these models cautiously within validated cation and anion groups to avoid uncertainties in predictions. Acar et al. used deep neural networks (DNN) and convolutional neural networks (CNN) with molecular descriptors to successfully predict viscosity data at room temperature [18]. They found that feature reduction improved results and that CNN slightly outperformed DNN, with coefficients of determination of 0.998 and 0.9869, respectively. Another commonly used ML model is Random Forest (RF), which can be successfully used in both classification and regression problems. Similar to DNN models, RF can handle non-linear relationships, however, it is better suited for smaller datasets. Venkatraman et al. employed RF for both regression and classification to predict the melting point of ILs, and achieved particularly promising results [19]. The regression model had coefficients of determination of 0.67 for training and 0.66 for testing, while the classification model achieved 0.84 training accuracy and 0.81 test accuracy. Venkatraman et al. also utilized RF regression model to predict CO_2 solubility of and demonstrated good performance with a coefficient of determination of 0.85 on a test set including unique ILs not used in training [20]. Furthermore, to predict melting points, Low et al. used a Kernel Ridge Regression (KRR) model and compared its performance to other models, including those by Venkatraman et al., which achieved a coefficient of determination of 0.66. The KRR model showed improved results, achieving a coefficient of determination of 0.76 [20, 21]. Nevertheless, compared to some of the other IL properties, such as viscosity and CO_2 solubility, melting point predictions were less accurate. Combined with the small dataset size, the primary reason for this limitation was recognized as the feature set's inability to effectively define changes in chemical interactions during solid-to-liquid transitions [22]. All these studies share the common goal of predicting the thermophysical properties of ILs using ML models, differing mostly in preprocessing techniques and model types. This research expands upon prior work by transforming the melting point dataset into the liquid phase transition temperature (LPTT) dataset by the inclusion of related equilibrium temperature data, which contains both liquid-crystal and glass transition temperatures. Additionally, the IL-split approach, adapted from Datta et al., was used for viscosity dataset splitting to develop a robust model capable of predicting the properties of unknown ILs [14]. This study diverges further by leveraging trained and tested models to predict the viscosity and LPTT of novel ILs, with the ultimate goal of validating these predictions experimentally and feeding the results back into the models in future research.

The current study focuses on the development of ML models for IL properties particularly relevant to pumped fluid loops, i.e., viscosity and LPTT. During model developments, the effects of data-splitting methods and model configurations on performance were analyzed. Once high-performance models were identified, novel ILs were generated from unique cation and anion combinations. These models were then used to predict the viscosity and LPTT of the novel ILs, establishing a workflow of training and evaluating models, followed by their application to identify ILs with advantageous properties. This workflow aims to address the significant research gap in ILs as heat transfer fluids and contribute to long-term advancements in this field.

II. Methodology

A. Dataset gathering and preparation

The pyILT2 Python library was used to access the NIST (National Institute of Standards and Technology) ILThermo Database (SRD#147) v2.0 and extract datasets for 2,620 unique pure ILs in different phases [12, 23]. The datasets consist of chemical and physical properties documented from experimental studies in the literature. In pursuit of designing novel ILs with superior thermophysical properties for use as working fluids in spacecraft single-phase pumped fluid loop thermal control systems, data on melting point, equilibrium temperature, viscosity, thermal conductivity, and heat capacity were identified for further processing. The distributions of thermophysical properties were studied to understand the relationship between temperature and the targeted properties. Since this research focuses on atmospheric conditions, datasets were filtered to a pressure range of 101.325 ± 2 kPa. Table 1 provides detailed information on the number of unique ILs, data points, temperature ranges, and property ranges for each targeted thermophysical property of pure ILs in the liquid phase under atmospheric conditions. The data in the table does not encompass all available data under liquid phase and atmospheric conditions. Some data for unique ILs were excluded because the feature generation method, which is described in the next section, could not be applied to those ILs.

Table 1 Filtered ILThermo database of target thermophysical properties of pure ILs in liquid phase under atmospheric conditions.

Property	# Unique ILs	# Data Points	Temp. Range [K]	Data Range
Melting Point	759	1,135	215.95 – 581.6	215.95 – 581.6 [K]
Equilibrium Temp.	1,000	1,796	120 – 651.1	120 – 651.1 [K]
Viscosity	1,139	19,926	191.2 – 537	$10^{-4} - 2 \times 10^9$ [Pa·s]
Thermal Conductivity	85	753	273.15 – 438.07	0.1032 – 0.97 [$\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$]
Heat Capacity	305	12,263	167.7 – 663.1	23.1 – 7,970 [$\text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$]

Given the application to proposed pumped fluid loops, our goal was to identify ionic liquids (ILs) with the widest temperature range in their liquid phase. The melting point represents the crystal-to-liquid transition temperature for ILs. The equilibrium temperature property also includes crystal-liquid and liquid-glass phase transition information. Although liquid-glass transitions differ significantly, as it requires supercooling to achieve, the glass transition data were used when melting point information was unavailable for a given IL. Consequently, the melting point and equilibrium temperature data were merged, giving 466 unique ILs with melting point, 699 with equilibrium temperature, and 293 present in both. For ILs in both datasets, melting point data were prioritized to avoid the lower glass transition temperatures, ensuring consistency in capturing the highest liquid-solid transition temperature. After merging and filtering, the new dataset was named “Liquid Phase Transition Temperature (LPTT)”.

The datasets before data cleaning are shown in Fig. 1, which illustrates the regions with sparse data, particularly at low and elevated temperatures. After examining the dataset, predictions of LPTT and viscosity were chosen as the primary focus for this study. LPTT is the most critical property for determining the operability of the working fluid in spacecraft thermal control systems, and the viscosity also has a direct impact on loop system performance. During the comprehensive study of viscosity and LPTT data, inconsistencies were detected due to multiple sources for a given IL and temperature. For LPTT, multiple temperature values were reported for the same IL, while viscosity data showed different values for the same IL at the same temperature. To resolve inconsistencies, the median of these values was computed. The datasets were then grouped by unique ILs and, for viscosity, by temperature as well. This data cleaning process improved data consistency but reduced the number of data points for both properties. As a result, the final viscosity dataset included 1,139 unique ILs with 13,605 data points, while the LPTT dataset contained 1,458 unique ILs and equal number of data points.

B. Featurization

Featurization transforms physical and chemical attributes of ILs into vector representations. These vectors enable ML models to uncover complex relationships between IL properties and operating conditions, beyond the scope of traditional analysis. In molecular featurization, the unique chemical information encoded of each IL is translated into a numerical and machine-readable format that can be used to distinguish them. To achieve this, the chemical structures

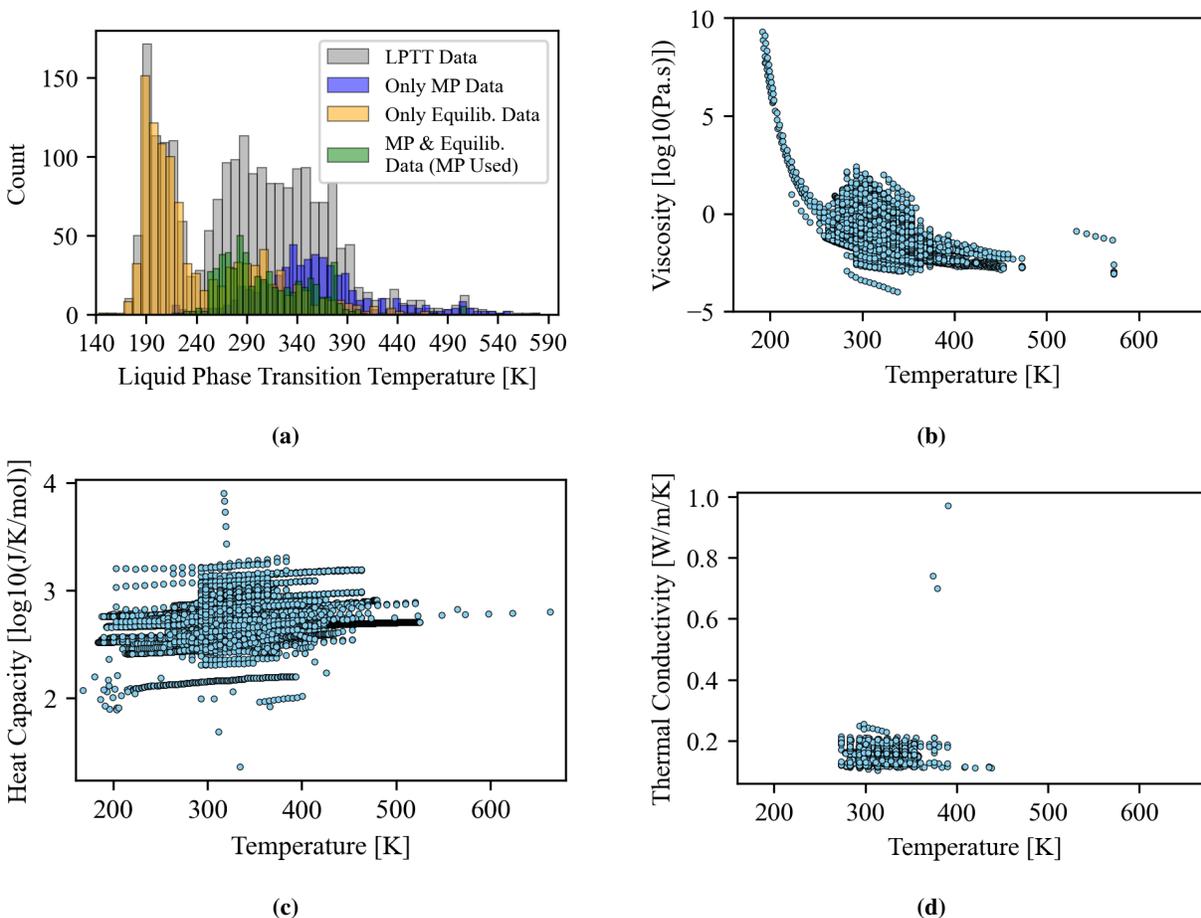


Fig. 1 Data distributions of thermophysical properties of pure ILs in liquid phase at atmospheric pressure. (a) The LPTT distribution is shown (gray), as well as with a distinction between data sources: only melting point data (purple), only equilibrium temperature data (yellow), and combined data where melting point was prioritized (green). (b) and (c) Viscosity and heat capacity as a function of temperature, scaled to \log_{10} due to their wide data range. (d) thermal conductivity as a function of temperature.

were converted to the Simplified Molecular Input Line Entry System (SMILES) convention, which expresses the 3-D connectivity of a molecule as a canonical string. Leveraging the py2opsin Python library, chemical names of the ILs were converted to SMILES strings, which were then used to obtain molecular descriptors through the RDKit package [24, 25]. Molecular descriptors were obtained using two methods, both of which included the generation of SMILES string from chemicals. In the first method, SMILES strings were directly obtained from ILs, which resulted in 210 molecular descriptors. Whereas, in the second method, ILs were separated into their cation and anion components, individual SMILES strings were generated for each, and 420 molecular descriptors were obtained. No significant difference in accuracy was found between the two methods. However, the cation-anion separation method reduced the number of unique ILs due to difficulties in obtaining some SMILES strings. Therefore, to maximize the number of unique ILs and minimize overfitting risks, the whole ILs method was chosen to generate the feature set.

In Fig. 2, the steps of the featurization process are illustrated, and four out of 210 molecular descriptors are provided for illustration. This process was applied to all unique ILs in both the viscosity and LPTT datasets. Following this, feature reduction methods were applied to reduce complexity and improve performance by removing the least important features, thereby mitigating issues associated with high dimensionality, such as overfitting. First, zero variance features were eliminated because they provide no informational value. Next, a correlation matrix was generated to identify fully correlated pairs, and one feature from each pair was removed to avoid redundancy. Additionally, features containing missing data were dropped since some ML models cannot process datasets with missing values. Finally,

feature importance was assessed using Least Absolute Shrinkage and Selection Operator (LASSO) regression model to determine the most important features and eliminate the rest [26, 27].

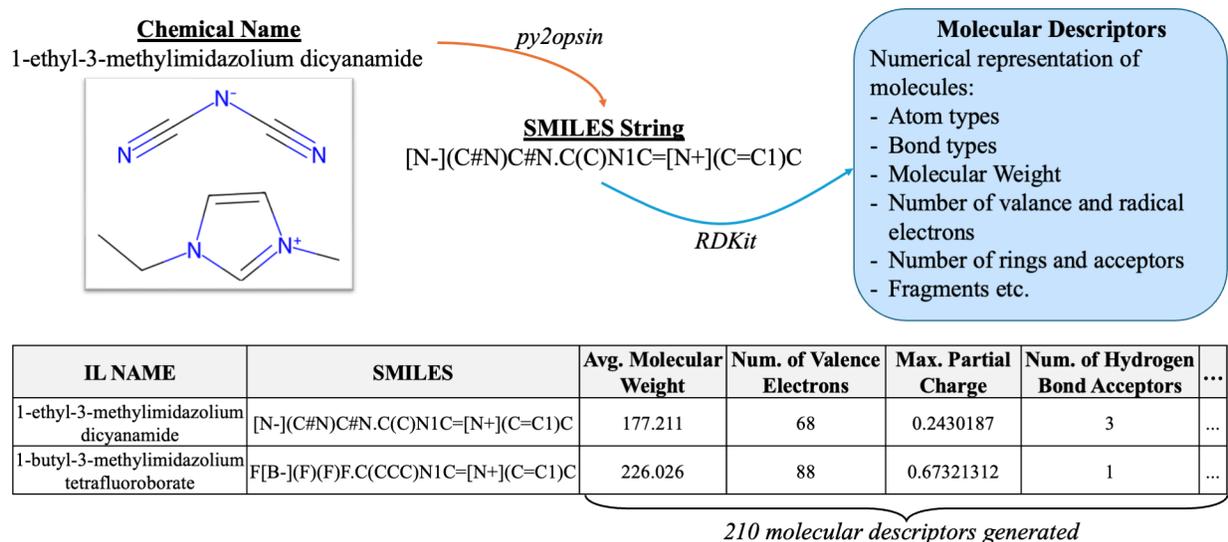


Fig. 2 Schematic representation of the featurization process used for molecular descriptor generation. Ionic liquids (ILs) are converted to SMILES strings, enabling numerical encoding of their molecular structures.

C. Training and test dataset splitting and scaling

To evaluate and test the model's performance, the viscosity dataset was split into the training and test sets using two methods: data-points-split and IL-split [14]. The data-points-split method, an entirely random approach, potentially allows the same ionic liquid to appear in both training and test sets. Hence, this method provides a broad variety of ILs and features for both sets. It is ideal for predicting the properties of already-known ILs at unmeasured temperatures. In contrast, the IL-split method groups data points by their corresponding unique ILs and ensures no overlap between training and test sets, which is crucial for predicting properties of completely new ILs. Also, during the splitting, care was taken to exclude extreme low and high-temperature data from the test set to ensure the ML model is trained with a broader range of temperature. An 80% training, 10% test, and 10% validation split ratio was used for both methods. Note that validation data was randomly split using the data-points-split method after separating the test set, as it is part of the training process to tune the model. On the other hand, the LPTT dataset was randomly split using the data-points-split method with a 90% training and 10% test ratio. Since each unique IL in the LPTT dataset is linked with a single temperature value, random splitting overlap between the training and test sets cannot occur. Additionally, a validation set was not required for the ML model of LPTT.

In Table 2, temperature statistics for the IL-split five training and test sets of the viscosity dataset are listed. These sets vary in the number of unique ILs, features, and temperature ranges. The variance in the number of features arises from the application of feature reduction methods to the training set after splitting. In addition to differences in the number of ILs, as shown in Fig. 3, each set contains a variety of cation and anion cores. The most common cation and anion cores were identified as imidazolium and NTf_2 derivatives, which are also the most frequent pairings in the training set. To compare the differences between the IL-split and data-points-split methods, statistics for a randomly split training and test set are provided in Table 3. Unlike the IL-split method, the data-points-split method includes unique ILs that appear in both the training and test sets. While IL-split test sets contain over 100 unseen ILs, the randomly split test set includes only 21 unique ILs, with the rest also present in the training set. Additionally, the randomly split test set covers a broader temperature range compared to the IL-split test sets. For LPTT, the statistics for the randomly split training and test sets are given in Table 3.

Next, the features were scaled using the min-max normalization method, which transformed each feature into a range of values between 0 and 1. The exact minimum and maximum values were then used to transform the test set. Additionally, viscosity was scaled using the logarithmic base ten (\log_{10}) due to its wide data range.

Table 2 Temperature statistics for IL-split train and test sets of the viscosity dataset.

Set	Features	Train			Test		
		ILs	Mean \pm Std [K]	Min – Max [K]	ILs	Mean \pm Std [K]	Min – Max [K]
#1	97	1033	322.19 \pm 30.57	191.2 – 573.0	106	319.79 \pm 23.75	278.15 – 573.0
#2	117	1032	322.12 \pm 30.64	191.2 – 573.0	107	320.46 \pm 22.98	278.15 – 433.15
#3	96	1035	322.07 \pm 30.70	191.2 – 573.0	104	320.98 \pm 22.26	278.15 – 398.15
#4	96	1033	322.02 \pm 30.69	191.2 – 573.0	106	321.39 \pm 22.34	278.15 – 423.1
#5	99	1032	321.94 \pm 30.47	191.2 – 573.0	107	322.16 \pm 25.07	278.15 – 573.0

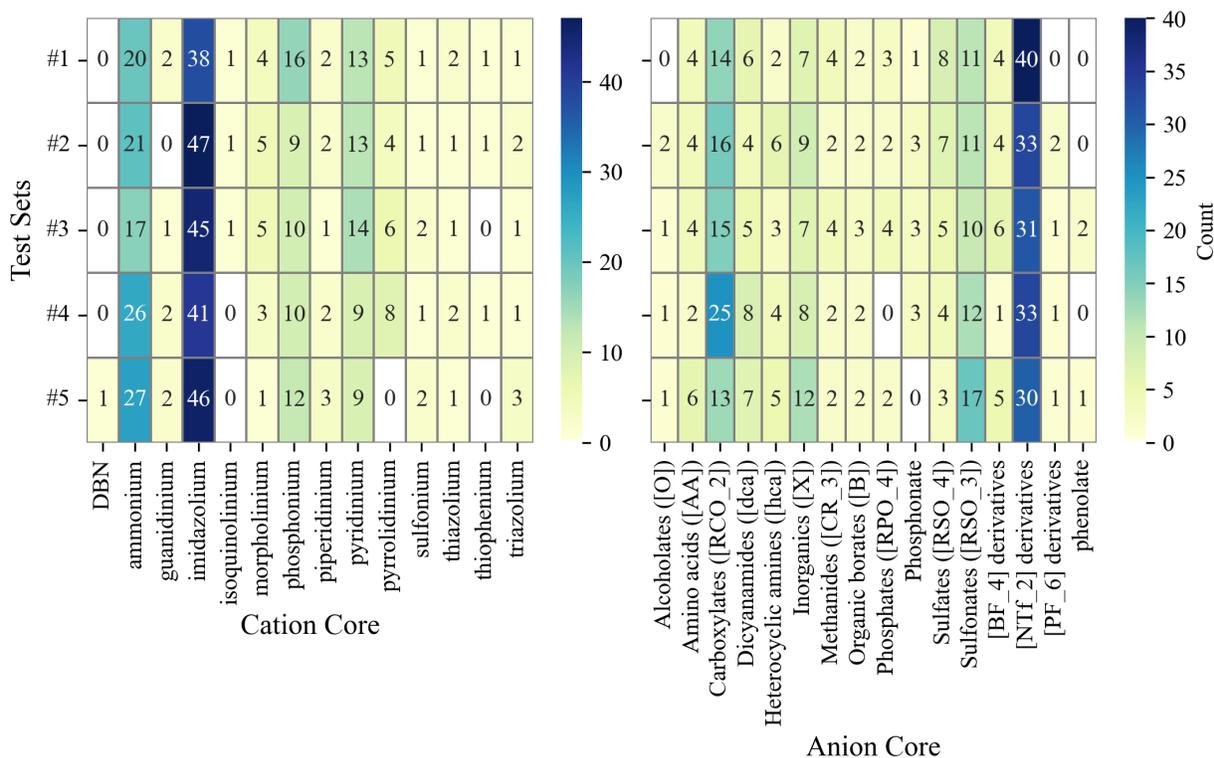


Fig. 3 Cation and anion cores of IL-split test sets for viscosity.

Table 3 Temperature statistics for data-points-split train and test sets of the viscosity and LPTT datasets. For the viscosity dataset, 475 ILs in the training set are not present in the test set, and 21 ILs in the test set are not in the train set.

Dataset	Train			Test		
	ILs	Mean \pm Std [K]	Min – Max [K]	ILs	Mean \pm Std [K]	Min – Max [K]
Viscosity	1118	321.86 \pm 39.15	191.2 – 573.0	664	322.89 \pm 28.51	201.2 – 458.1
LPTT	1326	293.11 \pm 75.79	170 – 581.6	132	305.61 \pm 62.09	193 – 411.1

D. Machine Learning model development

ML model development process is explored separately for viscosity and LPTT due to fundamental differences in their data sizes and nature.

1. Model development for viscosity

Regression models were investigated for viscosity. Given the complex and non-linear nature of the viscosity data, the deep neural network (DNN) was selected as the ML model for this research. The DNN models were developed with TensorFlow 2.16.1 with several key adjustable hyperparameters, including the number of hidden layers, nodes per layer, optimizer, activation function, and loss function [28]. In Fig. 4, the architecture of the DNN model developed for viscosity is shown. The ReLU activation function, Adam optimizer, and Mean Squared Error (MSE) loss function, as indicated in the diagram, are fixed hyperparameters in this study [29]. Other hyperparameters, such as the number of hidden layers and nodes per layer, were tuned to optimize model performance. During this process, 200 epochs and a batch size of 32 were found to be reliable across all models investigated. The hyperparameter tuning process began with model configurations consisting of 3, 4, and 5 hidden layers with 16, 32, 64, 128, 256, 512, 1024 nodes per layer. Once the model with best performance was identified, more detailed tuning was conducted around that configuration. These more detailed configurations are listed in Table 4, with varying hidden layers and nodes per layer.

Table 4 List of DNN model configurations for detailed hyperparameter tuning.

Model	Hidden Layer Configuration	Model	Hidden Layer Configuration
1	[128, 256, 256, 128]	8	[64, 128, 256, 128, 64]
2	[200, 200, 200, 200]	9	[128, 256, 320, 256, 128]
3	[256, 256, 128, 128]	10	[256, 128, 64, 32, 16]
4	[256, 256, 256, 256]	11	[256, 256, 256, 256, 256]
5	[256, 512, 512, 256]	12	[64, 128, 256, 512, 128, 64]
6	[280, 280, 280, 280]	13	[128, 256, 256, 256, 256, 128]
7	[320, 320, 320, 320]	14	[256, 256, 256, 256, 256, 256]

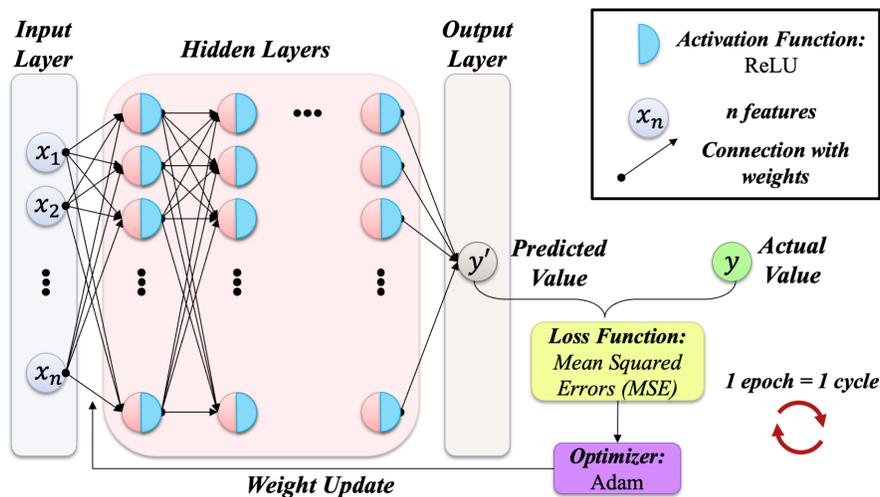


Fig. 4 Architecture of the DNN model for viscosity.

2. Model development for LPTT

Regression models, including Deep Neural Networks (DNN), Random Forest Regressor, and Support Vector Machine, were initially evaluated for predicting Liquid Phase Transition Temperature (LPTT). However, due to the relatively small size of the LPTT dataset, these models yielded unsatisfactory performance, even after extensive hyperparameter optimization. Consequently, a classification approach was adopted, focusing on categorizing ILs based on whether their LPTT was above or below a predetermined threshold.

The Random Forest Classifier was selected for this task, as shown in Fig.5. Hyperparameter tuning via grid search revealed minimal sensitivity to parameter variations. The optimal model configuration included 400 estimators, square root scaling for the maximum number of features per split, and unrestricted tree depth. Table 5 summarizes four model configurations, each defined by a different temperature threshold for classification. LPTT values were labeled as “high” or “low” relative to the threshold, encoded as 1 and 0, respectively, prior to feature selection. The final feature set varied with the threshold, as detailed in Table 5.

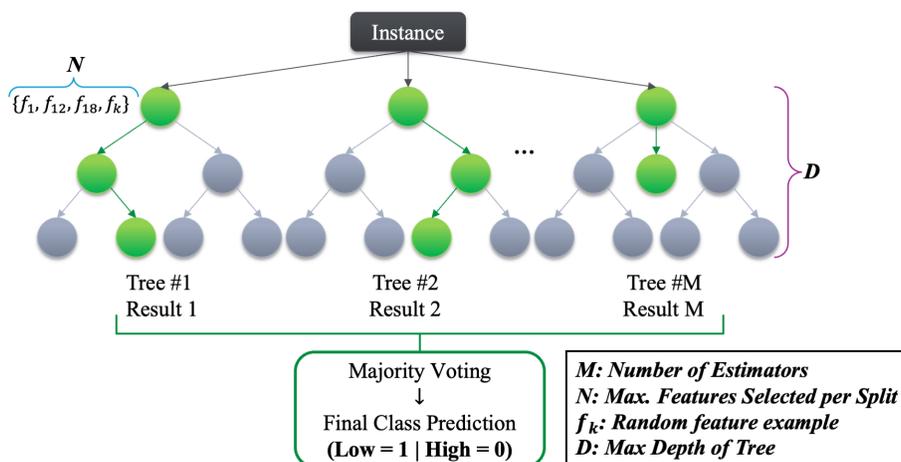


Fig. 5 Architecture of the Random Forest Classifier model for LPTT.

Table 5 LPTT RF Classifier model configurations with various thresholds.

Threshold [K]	Number of Features	Number of Estimators	Max. Features Selected per Split	Max. Depth of Tree
323	92	400	'sqrt'	None
293	75	400	'sqrt'	None
243	75	400	'sqrt'	None
223	70	400	'sqrt'	None

3. ML model performance evaluation criteria

The MSE loss function is used to evaluate progress during training of the viscosity DNN model. The loss values for both training and validation sets are examined through the epochs to identify potential overfitting or underfitting. If the model demonstrates a good fit, it is then tested using additional metrics. Criteria such as Root Mean Squared Error (RMSE), the coefficient of determination R^2 , and residual error are used to evaluate the model's performance. The equations for these criteria are provided in Eqs.(1 - 4), where y_i represents the actual value, \hat{y}_i the predicted value, and n is the number of data points.

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

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

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

$$r_i = y_i - \hat{y}_i \quad (4)$$

For LPTT predictions, classification metrics such as accuracy, precision, recall, and F1-score were used. The equations for these metrics are provided in Eqs.(5 - 8). These metrics are derived from the confusion matrix generated from the results, which is shown in Fig. 6 with label definitions.

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

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

Actual High (0)	True Negative (TN)	False Positive (FP)
	False Negative (FN)	True Positive (TP)
	Predicted High (0)	Predicted Low (1)

Fig. 6 Confusion matrix with label definitions.

E. Generation and preprocess of novel ILs

To achieve the objective of designing novel ILs with specific thermophysical properties that surpass those of current heat transfer fluids, 328,570 ionic liquids (ILs) were generated through cation and anion combinations. Then standard preprocessing steps were applied to generate molecular descriptors and prepare the datasets for ML predictions, with a slight modification to the feature reduction process, keeping only the features used to train the models. For the viscosity dataset, the feature set was expanded by adding four target temperatures (153 K, 193 K, 243 K, and 293 K) for predictions.

III. Results and Discussion

A. Viscosity

1. Impact of data-splitting method

To compare the outcomes of IL-split and data-points-split methods, a DNN model with four hidden layers and 256 neurons per layer was trained and tested on datasets partitioned using these two approaches. IL-split training-test sets #3 from Table 2 and a randomly split sets from Table 3, both utilizing the same 96 features, were employed in this process. The results, shown as parity plots in Fig. 7, illustrate the relationship between actual and predicted viscosity values with temperature variations. The randomly split test set achieved an R^2 value of 0.981, indicating strong predictive performance across a broader temperature and viscosity range. In comparison, the IL-split method accomplished an R^2 value of 0.902, which indicates a relatively weaker performance. The difference arises from the nature of the splits: the randomly split test set contains only 21 unique ILs, while the IL-split test set includes 104. The smaller number of unique ILs in the random split allows for better model performance due to greater feature overlap with the training set. Therefore, the results from the randomly split test set are closer to the perfect fit line, with minor deviations, while the IL-split method are more scattered. Some of these deviations in the IL-split results can be attributed to unique IL-specific trends, which implies the model’s difficulty in generalizing patterns that connect molecule descriptor features to viscosity. Although the data-points-split method performs better when the test set has fewer unseen ILs, the IL-split method was chosen to evaluate performance on unseen ILs for novel design. Despite its lower R^2 value of 0.902, which corresponds to $\sim 90\%$ accuracy and acceptable performance given dataset limitations.

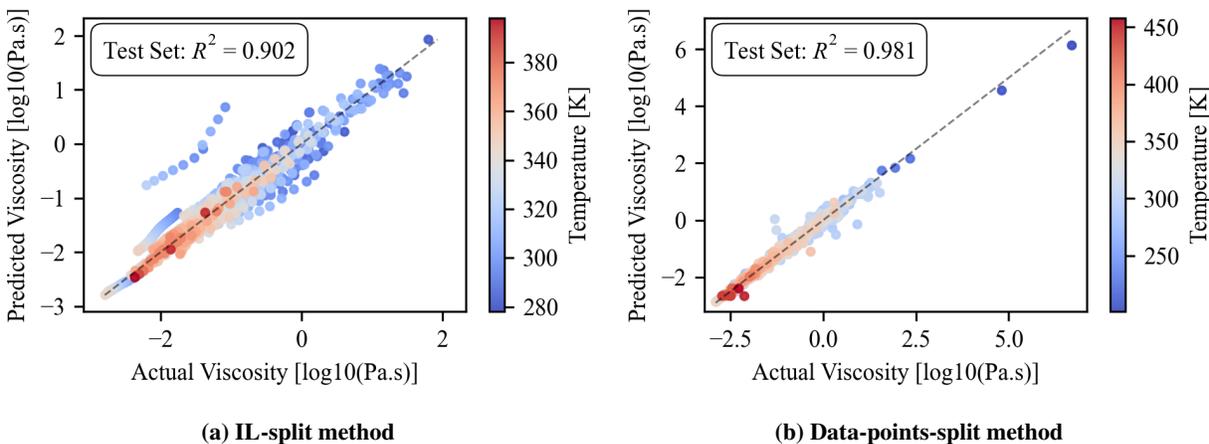


Fig. 7 Parity plots in logarithmic base ten scale comparing DNN models with 4 hidden layers, 256 neurons per layer, using (a) IL-split and (b) data-points-split methods with the same 96 features.

2. Impact of DNN configurations

Several configurations were tested, including 3, 4, and 5 hidden layers with 16, 32, 64, 128, 256, 512, and 1024 nodes per layer. The models demonstrated consistent performance, with an average R^2 score of 0.87. Notably, configurations with 4 hidden layers and 256 nodes per layer, as well as 3 hidden layers and 1024 nodes per layer, achieved $R^2 > 0.90$. The model with 4 hidden layers and 256 nodes per layer was selected as the base model for further tuning, as it offered comparable performance while avoiding unnecessary complexity and excessive computational costs.

After determining a range of nodes per layer and hidden layers for the model, more detailed tuning was conducted and 14 configurations, listed in Table 4, were investigated. These models were trained on 5 different training-test sets, given in Table 2. It is important to note that during the training process, the training and validation loss functions were monitored throughout the epochs, and no signs of underfitting or overfitting were detected. In Fig. 8, model performance for both training and test sets is given for 14 model configurations and 5 training-test splits. The training performances of all model configurations showed consistency, with high R^2 values (most exceeding 0.99) and low variability, which indicates that the models fit the training data well. Compared to the training results, test performance showed overall lower R^2 values and higher variability across configurations and training-test splits. Nonetheless, given the diversity

of unique ILs in the test sets, this difference was expected. Furthermore, most models (e.g., 1, 2, 3, 5, 8, 9, 10, 11, 12, 13, 14) demonstrated strong performance for test sets #1 and #3 but performed poorly on the other test sets, which resulted in significant accuracy variance. Models 6 and 7 showed better consistency across test sets but did not achieve the highest performance. On the other hand, model 4 had the best performance for test set #3 while maintaining more consistent results across the other sets. Consequently, among the evaluated configurations, model 4 (with 4 hidden layers and 256 nodes per layer) was chosen as the best model for viscosity predictions in this study. It displayed the highest generalization ability and the best performance by achieving $R^2=0.992$ and RMSE = 0.079 [log10(Pa·s)] for training, and $R^2=0.902$ and RMSE = 0.223 [log10(Pa·s)] for test set #3.

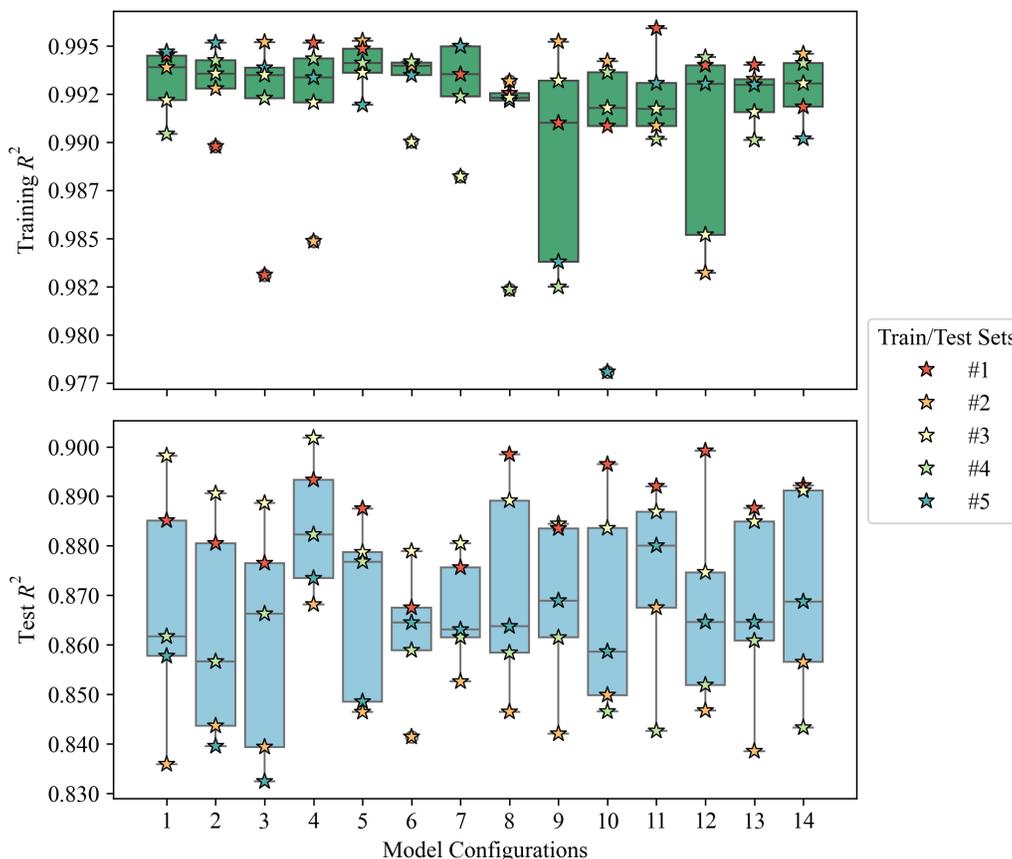


Fig. 8 DNN training and test R^2 performance across different configurations for data sets listed in Table 2.

Figure 9 shows the residual error distribution of the test set (#3) predictions. The residuals exhibit a near-normal distribution, with high symmetry around zero and both the mean and median close to zero. The lack of skewness indicates that the model does not systematically underpredict or overpredict. The distribution's short tails suggest that extreme errors are rare, with most being slightly negative, reflecting instances of overprediction. These outliers are evident in Fig. 7a, where a few predictions deviate notably from the trend. Overall, the model demonstrates robust performance, with minimal random errors and a slight negative bias.

As model's general performance was analyzed, its performance was also evaluated across different temperature ranges. For this purpose, RMSE and R^2 results for temperature ranges between 250 to 400 K, divided into 50 K intervals, are given in Fig. 10. The model performs better at higher temperatures in the test set, which is also apparent in Fig. 7a, where the more diverged predictions occurred at lower temperatures. This was expected, as the mean temperature of the training set is around 322 K. While model performance at low temperatures is particularly important for this project due to the focus on lower temperatures, an R^2 of 0.85 is still reasonably acceptable and not indicative of a poor performance.

Shapley Additive Explanations (SHAP) values, which quantify each feature's contribution to predictions, were used to interpret the model's performance [30]. Figure 11 shows the SHAP value distributions for 20 out of 96 features. Out of all features, temperature stands out as the most dominant, having the greatest influence on predictions of

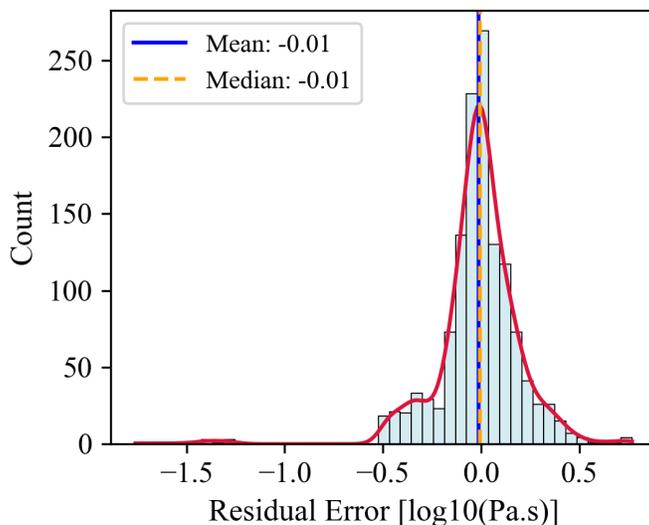


Fig. 9 Residual error distribution of the predictions (DNN model: 4 hidden layers, 256 nodes per layer).

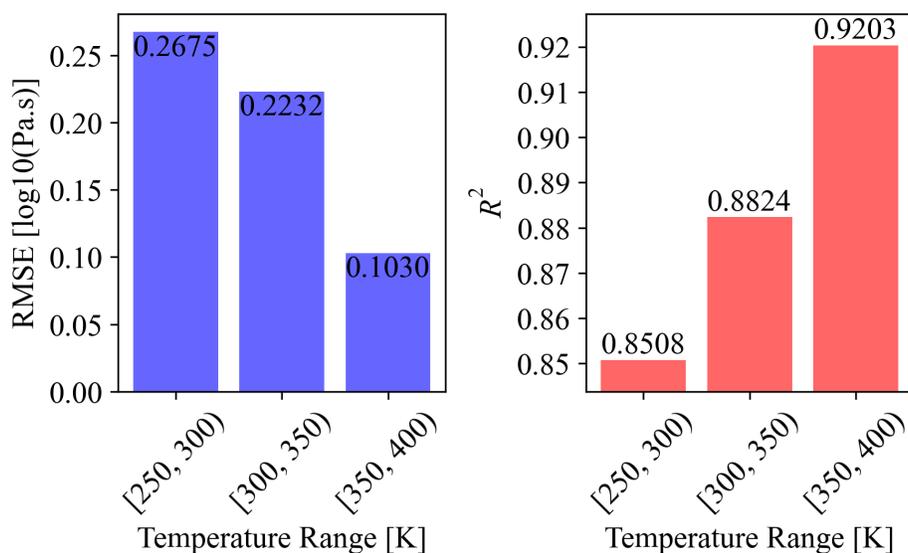


Fig. 10 RMSE and R^2 distribution across temperature ranges (DNN model: 4 hidden layers, 256 nodes per layer).

viscosity. The SHAP values for temperature follow an anticipated trend: when temperature increases (red), SHAP values shift to negative, which indicates a decrease in viscosity. This supports the well-established inverse relationship between viscosity and temperature. In opposition, lower temperatures (blue) with positive SHAP values signifies higher viscosity. Among the molecular descriptors, maximum absolute partial charge (“MaxAbsPartialCharge”) shows the next highest influence on predictions. The significantly higher influence of temperature compared to molecular descriptors highlights a limitation of the model, as its performance on predicting viscosity across unique ILs is relatively lower. This emphasizes the need to identify more influential features that better characterize ILs and their relationship with viscosity.

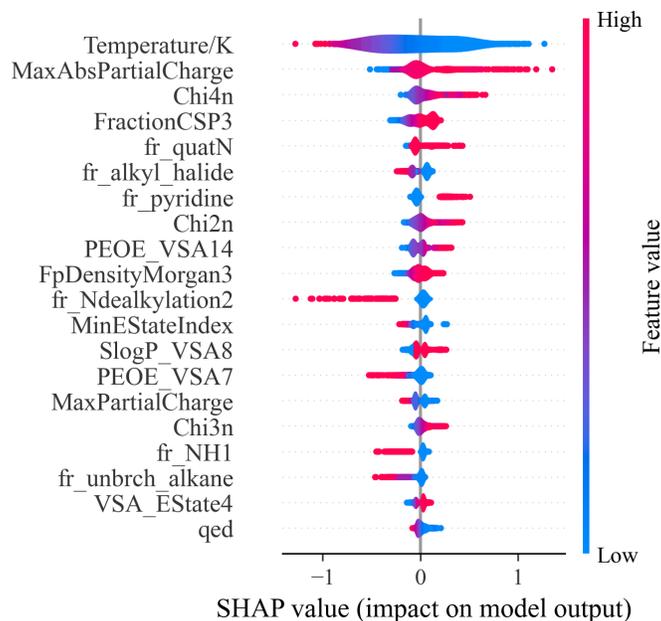


Fig. 11 Impact of features using SHAP values on the viscosity predictions (DNN model: 4 hidden layers, 256 nodes per layer).

B. LPTT

Regression models performed poorly for LPTT data, by obtaining an $R^2 \approx 0.50$ while showing signs of overfitting, likely due to the small dataset size. For classification, model training was conducted for four different temperature thresholds, and performance was evaluated using accuracy, precision, recall, and F1 score criteria for both training and test sets, as listed in Table 6. While the model fits the training set perfectly, the disparity in performance between the training and test sets suggests potential overfitting. The best performance is observed at the highest threshold 323 K, with a decline as the threshold decreased. This trend is likely because the 323 K threshold is above the mean temperature of the training set, resulting in a more balanced label distribution (0: high and 1: low) and providing the model with better patterns to learn. As the thresholds decrease, fewer data points are labeled as low, reducing the model's ability to identify meaningful patterns, particularly given the small dataset size. Despite this, test performance remains acceptable even at the lowest threshold. Figure 12 presents confusion matrices comparing predicted and actual results, showing a reduction in low-temperature predictions as the threshold decreases, which corresponds to the observed decline in model performance.

Table 6 Training and testing metrics for different thresholds.

Features	Threshold [K]	Accuracy		Precision		Recall		F1 Score	
		Train	Test	Train	Test	Train	Test	Train	Test
92	323	1.0	0.86	1.0	0.83	1.0	0.94	1.0	0.88
75	293	1.0	0.82	1.0	0.79	1.0	0.76	1.0	0.77
75	243	1.0	0.89	1.0	0.77	1.0	0.69	1.0	0.73
70	223	1.0	0.88	1.0	0.68	1.0	0.57	1.0	0.62

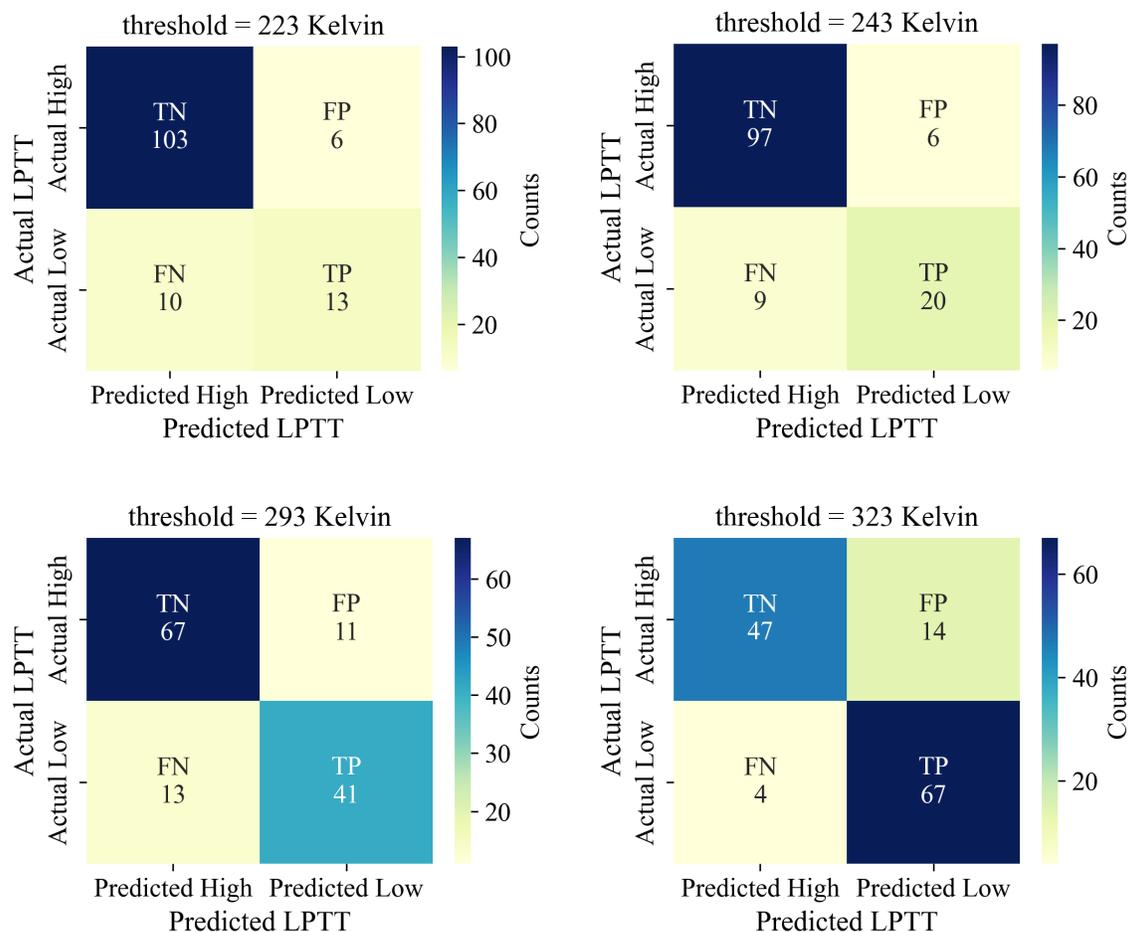


Fig. 12 Confusion matrix of LPTT results for four thresholds listed in Table 6.

C. Screening generated ILs for target properties

After training and evaluating the viscosity and LPTT ML models, as illustrated in Fig. 13, the feature sets of 328,740 generated unique ILs were fed into the trained models. Viscosity predictions were obtained from the DNN model (model 4 – trained and tested with set #3), while LPTT predictions were classified whether values were lower or higher than the 243 K threshold. The diagram also highlights the model performances and the resulting predictions for the generated ILs. The results from both viscosity and LPTT predictions were merged into a single dataset for further evaluation. ILs with LPTT lower than 243 K were selected, and their viscosity predictions were sorted based on the values at 153 K. From this, the 10 generated ILs with the lowest viscosities were selected and are presented in Fig. 14. The ILs with the “3-azido-1-(2-azidoethyl)-1H-1,2,4-triazol-1-ium” cation were identified as having the lowest viscosities at 153 K and an LPTT lower than 243 K. While no experimental data exist for these ILs to confirm the predictions, related data suggest the results may be promising. Specifically, the IL “3-azido-1-(2-azidoethyl)-1H-1,2,4-triazol-1-ium nitrate” has an LPTT value of 219.1 K, which makes the other results promising, as they share the same cation. However, the influence of different anions on the properties remains uncertain. For comparison, the IL “1-(2-azidoethyl)-1,2,4-triazolium perchlorate” has a melting point of 217 K [31]. The generated IL “3-azido-1-(2-azidoethyl)-1H-1,2,4-triazol-1-ium perchlorate” shares the azidoethyl group and triazolium core, where a similar trend in LPTT could be expected. Furthermore, the generated IL has an asymmetrical structure, further supporting the plausibility of the predictions. However, the accuracy of the viscosity predictions remains uncertain, particularly at 153 K, a temperature at which the ILs are likely to exist in a solid state.

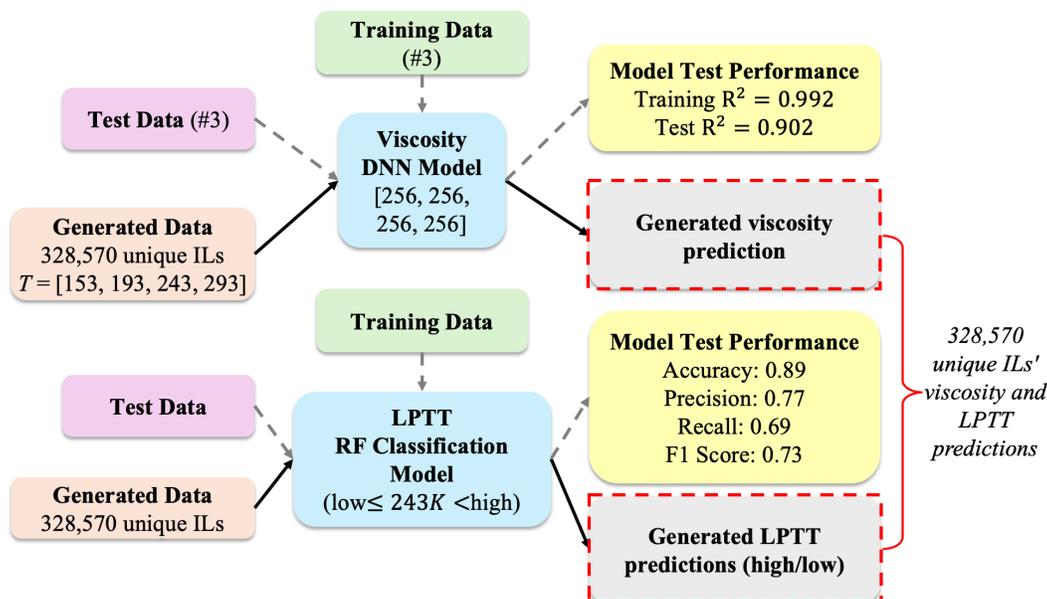


Fig. 13 Workflow of viscosity and PTT prediction models for generated ILs.

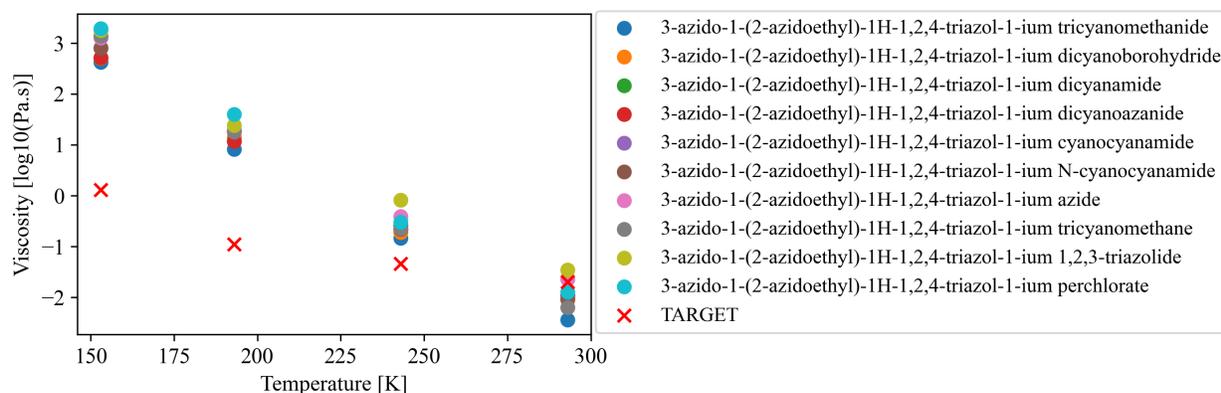


Fig. 14 Log₁₀ scaled viscosity vs. temperature for 10 generated ILs with low viscosity and LPTT below 243 K.

IV. Conclusion

This study presents a comprehensive machine learning (ML) framework to predict the thermophysical properties of ionic liquids (ILs) and identify novel ILs with superior properties for single-phase pumped fluid loops in spacecraft thermal control systems. Deep neural networks (DNN) and random forest classifiers were developed and evaluated for predicting viscosity and liquid phase transition temperature (LPTT), respectively. The models were trained using curated datasets of 1,139 ILs for viscosity and 1,458 ILs for LPTT, incorporating molecular descriptors and temperature as features.

Key findings include the successful application of the IL-split method for evaluating model performance on unseen ILs, achieving acceptable predictive performance (R^2 of 0.902) despite the limitations of small dataset size and feature dominance by temperature. For LPTT, classification at a 243 K threshold yielded robust results, emphasizing the importance of balanced datasets for threshold-based models. Feature importance analysis using SHAP values highlighted temperature as the most influential predictor for viscosity, reinforcing its significance in IL property prediction.

The trained models were applied to a dataset of 328,740 generated ILs to predict viscosity and LPTT, screening for ILs with low viscosity and LPTT below 243 K. The selected ILs, particularly those containing the “3-azido-1-(2-azidoethyl)-1H-1,2,4-triazol-1-ium” cation, demonstrated promising characteristics for spacecraft thermal management. However, the lack of experimental validation underscores the need for further experimental studies to confirm the

predictions, especially for low-temperature viscosity values.

This work establishes a scalable and efficient methodology for identifying high-performance ILs, addressing a critical research gap in heat transfer fluids for extreme environments. Future research will focus on integrating experimental validation with iterative ML model refinement to enhance prediction accuracy and support the development of advanced ILs for thermal management applications.

Acknowledgments

This work was supported by an Early Career Faculty grant from NASA's Space Technology Research Grants Program (Grant Number: 80NSSC23K1501). The authors would like to thank Tom Leimkuehler from NASA JSC as well as Eric Fox and Christopher Henry from NASA MSFC for their valuable discussions and insightful feedback, which contributed to the development of this work. Additionally, the authors thank Yuseok Kim from Cornell University for his assistance during the early stages of this work.

References

- [1] Birur, G. C., Bhandari, P., Bame, D., Karlmann, P., Mastropietro, A. J., Liu, Y., Miller, J., Pauken, M., and Lyra, J., "From Concept to Flight: An Active Fluid Loop Based Thermal Control System for Mars Science Laboratory Rover," *42nd International Conference on Environmental Systems*, 2012. <https://doi.org/10.2514/6.2012-3514>.
- [2] Ungar, E. K., and Erickson, L. R., "Assessment of the Use of Nanofluids in Spacecraft Active Thermal Control Systems," *AIAA SPACE 2011 Conference & Exposition*, AIAA 2011-7328, 2011. <https://doi.org/10.2514/6.2011-7328>.
- [3] van Gerner, H., Van Benthem, R., van Es, J., Schwaller, D., and Lapensée, S., "Fluid Selection for Space Thermal Control Systems," *44th International Conference on Environmental Systems*, Tucson, 2014.
- [4] Westheimer, D. T., and Tuan, G. C., "Active Thermal Control System Considerations for the Next Generation of Human Rated Space Vehicles," *43rd AIAA Aerospace Sciences Meeting and Exhibit - Meeting Papers*, 2005, pp. 1803–1807. <https://doi.org/10.2514/6.2005-342>.
- [5] Gilmore, D., *Spacecraft Thermal Control Handbook, Volume I: Fundamental Technologies*, American Institute of Aeronautics and Astronautics, Inc., Washington, DC, 2002. <https://doi.org/10.2514/4.989117>.
- [6] Cutbirth, J., "Designer Fluids for Use in a Single Loop Variable Heat Rejection Thermal Control System," Tech. rep., Mainstream Engineering Corporation, 2015. URL <https://techport.nasa.gov/api/file/294638>.
- [7] Stephan, R. A., "Overview of the Altair Lunar Lander Thermal Control System Design," *40th International Conference on Environmental Systems*, AIAA 2010-6080, 2010. <https://doi.org/10.2514/6.2010-6080>.
- [8] Ungar, E. K., "Spacecraft Radiator Freeze Protection Using a Regenerative Heat Exchanger with Bypass Setpoint Temperature Control," 2008. <https://doi.org/10.4271/2008-01-2170>.
- [9] Perry, J. L., "Case Studies in Crewed Spacecraft Environmental Control and Life Support System Process Compatibility and Cabin Environmental Impact," Tech. rep., NASA/TP-2017-219846, Huntsville, 2017. URL <http://www.sti.nasa.gov>.
- [10] Paris, A. D., Bhandari, P., and Birur, G. C., "High Temperature Mechanically Pumped Fluid Loop for Space Applications –Working Fluid Selection," 2004. <https://doi.org/10.4271/2004-01-2415>.
- [11] Minea, A. A., "Overview of Ionic Liquids as Candidates for New Heat Transfer Fluids," *International Journal of Thermophysics*, Vol. 41, No. 11, 2020. <https://doi.org/10.1007/s10765-020-02727-3>.
- [12] Kazakov, A., Magee, J., Chirico, R., Diky, V., Kroenlein, K., Muzny, C., and Frenkel, M., "Ionic Liquids Database - ILThermo (v2.0)," 11 2013. URL <http://trcsv1.boulder.nist.gov/ilthermo/ilthermo.html> (Accessed August 18, 2024).
- [13] Koutsoukos, S., Philippi, F., Malaret, F., and Welton, T., "A Review on Machine Learning Algorithms for the Ionic Liquid Chemical Space," *Chemical Science*, Vol. 12, No. 20, 2021, pp. 6820–6843. <https://doi.org/10.1039/d1sc01000j>.
- [14] Datta, R., Ramprasad, R., and Venkatram, S., "Conductivity Prediction Model for Ionic Liquids Using Machine Learning," *Journal of Chemical Physics*, Vol. 156, No. 21, 2022. <https://doi.org/10.1063/5.0089568>.
- [15] Valderrama, J. O., Muñoz, J. M., and Rojas, R. E., "Viscosity of Ionic Liquids Using the Concept of Mass Connectivity and Artificial Neural Networks," *Korean Journal of Chemical Engineering*, Vol. 28, No. 6, 2011, pp. 1451–1457. <https://doi.org/10.1007/s11814-010-0512-0>.

- [16] Padaszyński, K., and Domańska, U., “Viscosity of Ionic Liquids: An Extensive Database and a New Group Contribution Model Based on a Feed-Forward Artificial Neural Network,” *Journal of Chemical Information and Modeling*, Vol. 54, No. 5, 2014, pp. 1311–1324. <https://doi.org/10.1021/ci500206u>.
- [17] Padaszyński, K., “Extensive Databases and Group Contribution QSPRs of Ionic Liquids Properties. 2. Viscosity,” *Industrial and Engineering Chemistry Research*, Vol. 58, No. 36, 2019, pp. 17049–17066. <https://doi.org/10.1021/acs.iecr.9b03150>.
- [18] Acar, Z., Nguyen, P., Cui, X., and Lau, K. C., “Room Temperature Ionic Liquids Viscosity Prediction from Deep-Learning Models,” *Energy Materials*, 2023. <https://doi.org/10.20517/energymater.2023.38>.
- [19] Venkatraman, V., Evjen, S., Knuutila, H. K., Fiksdahl, A., and Alsberg, B. K., “Predicting Ionic Liquid Melting Points Using Machine Learning,” *Journal of Molecular Liquids*, Vol. 264, 2018, pp. 318–326. <https://doi.org/10.1016/j.molliq.2018.03.090>.
- [20] Venkatraman, V., and Alsberg, B. K., “Predicting CO₂ Capture of Ionic Liquids Using Machine Learning,” *Journal of CO₂ Utilization*, Vol. 21, 2017, pp. 162–168. <https://doi.org/10.1016/j.jcou.2017.06.012>.
- [21] Low, K., Kobayashi, R., and Izgorodina, E. I., “The Effect of Descriptor Choice in Machine Learning Models for Ionic Liquid Melting Point Prediction,” *The Journal of Chemical Physics*, Vol. 153, No. 10, 2020. <https://doi.org/10.1063/5.0016289>.
- [22] Hughes, L. D., Palmer, D. S., Nigsch, F., and Mitchell, J. B., “Why Are Some Properties More Difficult to Predict Than Others? A Study of QSPR Models of Solubility, Melting Point, and log P,” *Journal of Chemical Information and Modeling*, Vol. 48, No. 1, 2008, pp. 220–232. <https://doi.org/10.1021/ci700307p>.
- [23] Dong, Q., Muzny, C., Kazakov, A., Diky, V., Magee, J., Widegren, J., Chirico, R., Marsh, K., and Frenkel, M., “ILThermo: A Free-Access Web Database for Thermodynamic Properties of Ionic Liquids,” *Journal of Chemical & Engineering Data*, Vol. 52, No. 4, 2007, pp. 1151–1159. <https://doi.org/10.1021/jc700171f>.
- [24] Lowe, D. M., Corbett, P. T., Murray-Rust, P., and Glen, R. C., “Chemical Name to Structure: OPSIN, an Open Source Solution,” *Journal of Chemical Information and Modeling*, Vol. 51, No. 3, 2011, pp. 739–753. <https://doi.org/10.1021/ci100384d>.
- [25] Landrum, G., Tosco, P., Kelley, B., Ric, Cosgrove, D., sriniker, Vianello, R., gedeck, NadineSchneider, Jones, G., Kawashima, E., N, D., Dalke, A., Cole, B., Swain, M., Turk, S., Savelev, A., Vaucher, A., Wójcikowski, M., Take, I., Scalfani, V. F., Probst, D., Ujihara, K., godin, g., Pahl, A., Walker, R., Lehtivarjo, J., Berenger, F., strets123, and jasondbiggs, “rdkit/rdkit: Release_2023.09.5,” Zenodo, 2024. <https://doi.org/10.5281/zenodo.10633624>.
- [26] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2825–2830.
- [27] Tibshirani, R., “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 58, No. 1, 1996, pp. 267–288. URL <http://www.jstor.org.proxy.library.cornell.edu/stable/2346178>.
- [28] TensorFlow Developers, “TensorFlow (v2.16.1),” Zenodo, 2024. <https://doi.org/10.5281/zenodo.10794955>.
- [29] Kingma, D. P., and Ba, J., “Adam: A Method for Stochastic Optimization,” 2014. URL <http://arxiv.org/abs/1412.6980>.
- [30] Lundberg, S., and Lee, S.-I., “A Unified Approach to Interpreting Model Predictions,” 2017. URL <http://arxiv.org/abs/1705.07874>.
- [31] Xue, H., and Shreeve, J., “Energetic Ionic Liquids from Azido Derivatives of 1,2,4-Triazole,” *Advanced Materials*, Vol. 17, No. 17, 2005, pp. 2142–2146. <https://doi.org/10.1002/adma.200500789>.