Journal of Molecular Liquids 388 (2023) 122747

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Deep eutectic solvent viscosity prediction by hybrid machine learning and group contribution

Ahmadreza Roosta [a], Reza Haghbakhsh [b,c,*], Ana Rita C. Duarte [c], Sona Raeissi [a,*]

[a] School of Chemical and Petroleum Engineering, Shiraz University, Mollasadra Ave., Shiraz 71348-51154, Iran
[b] Department of Chemical Engineering, Faculty of Engineering, University of Isfahan, 81746-73441 Isfahan, Iran
[c] LAQV, REQUIMTE, Departamento de Química da Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Keywords:
DES
Physical property
Machine learning
Artificial neural network
Support vector machine

ABSTRACT

In this study, hybrid machine learning nonlinear models were developed to predict the viscosity of DESs by combining the group contribution (GC) concept with the multilayer perceptron, a well-known feedforward artificial neural network, and the Least Squares Support Vector Machine (LSSVM) algorithm. Deep Eutectic Solvents (DESs) have come to the forefront in recent years as appealing substitutes for conventional solvents. It is imperative to have a thorough grasp of the essential properties of DESs to expand the employment of these compounds in new procedures. Most frequently, one of the crucial physical properties of a DES that must be precisely determined is its viscosity. To develop the models, a dataset of 2533 viscosity data points for 305 DESs at various temperatures (from 277.15 to 373.15 K) was gathered to build the models. By using temperature, molar ratios, and functional groups as inputs, the results indicate that the suggested models can determine the viscosity of DESs with high accuracy. The models yield average absolute relative deviations below 10% and squared correlation coefficients higher than 0.98.

1. Introduction

The chemical industry has been essential to modern living and works as a growth engine for contemporary society, but it has also contributed significantly to environmental pollution. It is imperative, yet very challenging, for the chemical industry to adjust to minimize potential environmental impact [1]. One step that reduces the production of chemical pollutants is to minimize or completely stop the use of toxic solvents in industrial processes. Therefore, the search for green solvents for use in chemical processes is gaining popularity. The ideal green solvent would be sustainable, non-toxic, biodegradable, readily available, economical, and easily synthesized. Deep Eutectic Solvents (DESs) mostly meet these criteria as prospective green solvents. DESs are mixtures of two or three hydrogen bonding components. The component that accepts the hydrogen bond is known as the hydrogen bond acceptor (HBA), whereas the component that donates the hydrogen bond is known as the hydrogen bond donor (HBD) [2,3].

It was the discovery by Abbott et al. [4] of the exceptionally significant melting point depression at eutectic composition, due to hydrogen bond interactions between the HBA and HBD molecules, that spurred the rise of interest in DESs over the past two decades. Apart from desirable features like biodegradability, non-toxicity, low vapor pressure, inflammability, low cost, and ease of preparation, their ability to be tailored for particular physicochemical properties makes them even more interesting for a variety of industrial applications, including gas solubility [5], electrodeposition [6], biocatalysis [7], pharmaceuticals [8], biodiesels [9], polymers [10], batteries and power systems [11].

But the use of DESs necessitates a thorough comprehension of their fundamental properties. One of the most important properties is viscosity, however, a major issue with DESs is their high viscosity, resulting in for example, higher pumping costs and decreased solute mass transfer rate during solvent extraction. The intrinsic complexity in the mechanism of liquid phase momentum transfer, as well as the very large variations in viscosity with temperature, have made it a most difficult property to model [12,13]. Mjalli and Naser [14] first suggested models for estimating the viscosity of choline chloride-based DESs with the aid of the Eyring-Wilson and the Vogel-Fulcher-Tamman equations. They took into account the temperature and the salt mole fraction in their

* Corresponding authors at: School of Chemical and Petroleum Engineering, Shiraz University, Mollasadra Ave., Shiraz 71348-51154, Iran (S. Raeissi) Department of Chemical Engineering, Faculty of Engineering, University of Isfahan, 81746-73441, Isfahan, Iran and LAQV, REQUIMTE, Departamento de Química da Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal (R. Haghbakhsh).
E-mail addresses: r.haghbakhsh@eng.ui.ac.ir, r.haghbakhsh@fct.unl.pt (R. Haghbakhsh), raeissi@shirazu.ac.ir (S. Raeissi).

https://doi.org/10.1016/j.molliq.2023.122747
Received 29 April 2023; Received in revised form 17 July 2023; Accepted 2 August 2023
Available online 3 August 2023

Table 1
Different functional groups of the HBA and HBD used in this study.

Number	Groups	Number	Groups
	Not within a ring		
1	CH_3	25	P═O
2	—CH_2	26	F
3	>CH—	27	Cl
4	>C<	28	Br
5	═C<	29	P
6	>C═O	30	O═S═O
7	═CH_2	**In an unsaturated ring**	
8	═CH—	31	═CH—
9	O—C═O	32	═C<
10	COOH	33	—N═
11	OH	34	>NH
12	—O—/O^-	35	>CH—
13	>N<	36	—CH_2—
14	>N—/[>NH—]$^+$	37	OH
15	>NH/>NH_2^+	38	Cl
16	═NH	**In a saturated ring**	
17	—NH_2/—NH_3^+	39	—CH_2—
18	H_2O	40	>CH—
19	Cr	41	>C<
20	—C≡N	42	>C═O
21	—S—/—S	43	—O—
22	K	44	>NH_2^+/>NH
23	CO_3^{2-}	45	OH
24	Na		

models. In an effort to model twenty-seven different DESs, Haghbakhsh et al. [15] followed the approaches of the perturbed chain-statistical associating fluid theory (PC-SAFT) and the cubic plus association (CPA) equations of state in conjunction with the friction theory. Both models' deviations were 4.4%. Despite the acceptable errors, their DES viscosity databank contained only 253 viscosity data points, which is relatively small. Furthermore, they had fitted the friction theory parameters for each DES, which made it a correlative approach. With a databank of 156 DESs, Bakhtyari et al. [12] suggested a correlative model using critical pressure, critical temperature, and one reference viscosity data point at a reference temperature as the input parameters. Their model's average absolute relative deviation (*AARD%*) was 10.4% for the entire databank. Despite being reliable simply as a correlative tool, this model does not consider the effect of DES's molecular structure on viscosity. Eqs.(1)–(3) present the relations of Bakhtyari et al.'s model [12],

$$\mu = \left[\mu_{ref}^{A} + B\left(\frac{T_{ref} - T}{T \times T_{ref}}\right)\right]^{1/A} \tag{1}$$

$$A = {}^{-0.817}/_{P_c} - 0.123 \tag{2}$$

$$B = -1.595T_c \tag{3}$$

where P_c and T_c stand for critical pressure (bars) and critical temperature (kelvin), respectively, and μ and μ_{ref} represent the viscosities (Pa.s) at T and T_{ref}, respectively.

From the standpoint of molecular structure, the quantitative structure–property relationship (QSPR) model, based on the conductor-like screening model for real solvents (COMSO-RS), was put forth by Lemaoui et al. [16] to predict the viscosities of hydrophobic eutectic solvents (HESs). This model, however, is only applicable to hydrophobic eutectic solvents. The group contribution (GC) methods, commonly referred to as additive methods, divide a compound's molecular structure into its building blocks (such as functional groups), and estimate properties of the compound by summing the contribution of each building block [17]. Due to the possibility of using the range of defined functional groups to describe the molecular structure of numerous compounds, the GC methods are inherently predictive [18]. As a result, the GC approach is among the best and most effective techniques for estimating viscosity. The GC method, together with transition state theory, influenced a model presented by Yu et al. [13]. For a database of 183 DESs, they achieved *AARD%* values of 8.12% and 8.64% for their training and testing data sets, respectively. However, because of the numerous factors in this model and the requirement that users have a good understanding of the HBA and HBD structure, this model can be complicated and time-consuming. The mathematical form of this model is expressed by Eqs. (4)–(8) [13],

$$\begin{aligned}\ln\eta_{DES} &= x_{HBA}\left(A'_{HBA} + \frac{B'_{HBA}}{T} + \frac{C'_{HBA}}{T - E'} + D'_{HBA}\ln(x_{HBA}M_{HBA})\right)\\ &+ x_{HBD}\left(A'_{HBD} + \frac{B'_{HBD}}{T} + \frac{C'_{HBD}}{T - E'} + D'_{HBD}\ln(x_{HBD}M_{HBD})\right) + x_{HBA}x_{HBD}G_{HBA,HBD}\end{aligned} \tag{4}$$

$$A' = a_1\left(\sum_{j=1}^{k} n_j\Delta A'_j\right) \tag{5}$$

$$B' = b_1\left(\sum_{j=1}^{k} n_j\Delta B'_j\right) \tag{6}$$

$$C' = c_1\left(\sum_{j=1}^{k} n_j\Delta C'_j\right) \tag{7}$$

$$D' = d_1\left(\sum_{j=1}^{k} n_j\Delta D'_j\right) \tag{8}$$

where η_{DES} is the DES's viscosity, $\Delta A'_j$, $\Delta B'_j$, $\Delta C'_j$, and $\Delta D'_j$ represent the

Table 2
Statistical metrics (including *AARD%*, R^2, and *RMSE*) for correlating and/or predicting the viscosities of DESs by the various developed models.

Optimization techniques	AARD%			R^2			RMSE		
	Training set	Testing set	All data	Training set	Testing set	All data	Training set	Testing set	All data
MLPANN (BR)	**8.53**	**13.02**	**9.43**	**0.98**	**0.98**	**0.98**	**58.48**	**31.26**	**54.14**
MLPANN (LM)	9.97	14.16	10.81	0.98	0.98	0.98	58.01	27.70	53.34
MLPANN (Rprop)	12.80	16.96	13.63	0.96	0.97	0.96	79.05	44.60	73.46
LSSVM (CSA-gridsearch)	**5.59**	**12.02**	**6.88**	**0.99**	**0.97**	**0.99**	**44.33**	**42.10**	**43.90**

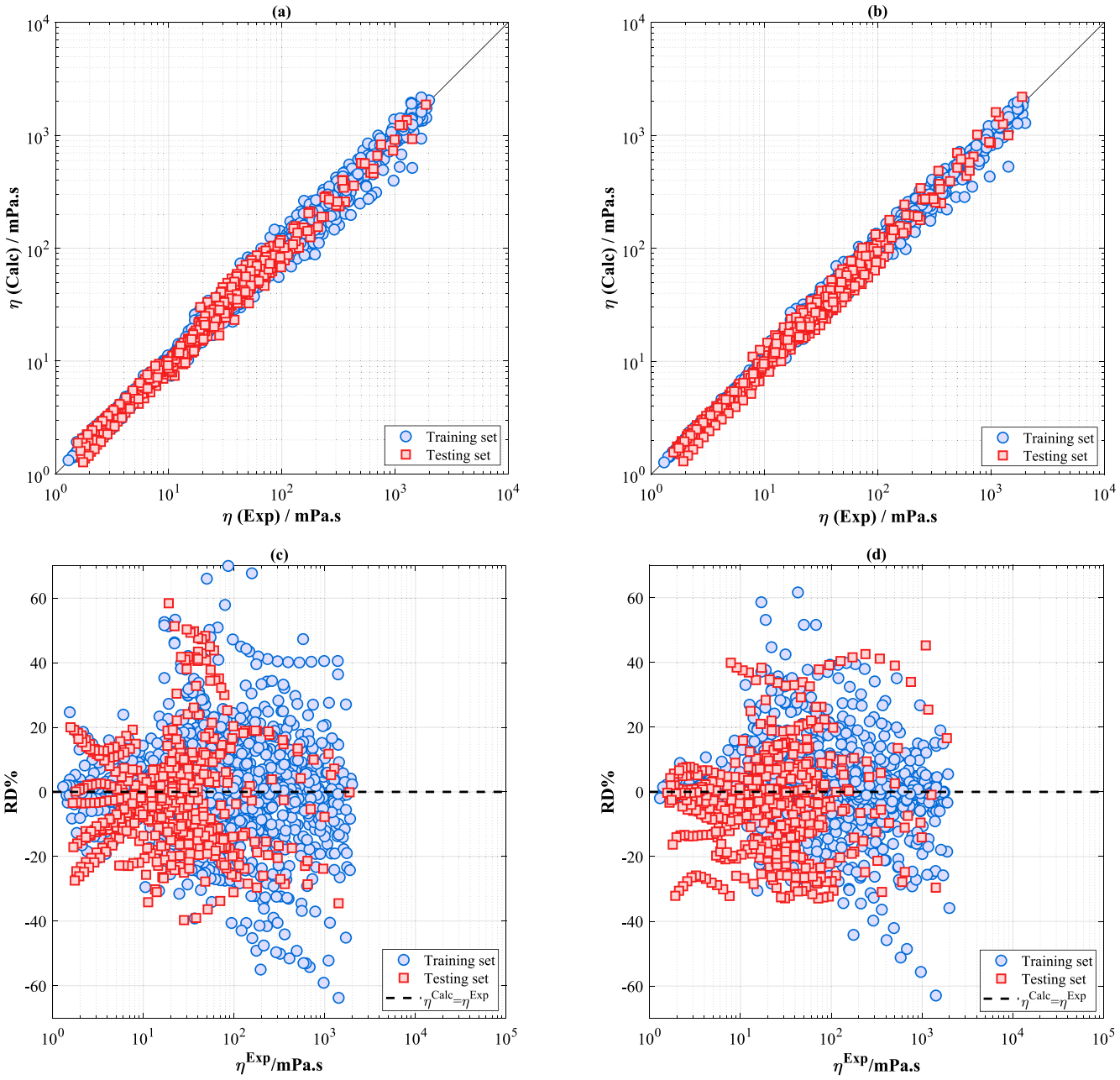

Fig. 1. Calculated versus experimental viscosities and the relative deviations of the estimated DES viscosities for the proposed MLPANN-GC and LSSVM-GC models: (a), (c) MLPANN-GC and (b), (d) LSSVM-GC.

various group contributions, and n_j is the number of occurance of the group of type *j,* while *k* represents the overall number of groups in the molecule. x_{HBA} and x_{HBD} are the constituent components' mole fractions, and M_{HBA} and M_{HBD} are the molecular weights of the HBA and HBD. $A^{'}$,$B^{'}$,$C^{'}$, and $D^{'}$ are adjustable parameters, while $E^{'}$ is a constant. The interactions between the HBA and HBD components is represented by the term $G_{HBA,HBD}$. More detailed information on the parameters of this model can be found in the corresponding literature [13]. In another study [19], the same group estimated the viscosities of 183 DESs using a neural network influenced by the theory of transition states, resulting in an *AARD%* of 6.84%. In this way, both of their models rather accurately estimated the DES viscosities with acceptable *AARD%* values of less than 9%. Although they did incorporate a GC approach, it should be noted that they did not carry out predictions in their work.

There is no denying that we live in a data flooded era, and machine learning (ML) has made it possible to quickly develop models that can evaluate sizeable complicated information to produce accurate results. It can identify patterns and learn complex behavior from a large set of data. ML employs an algorithm among a variety of possible algorithms, including for example, decision tree, support vector machine (SVM), random forest, k-means, and artificial neural network (ANN). Two studies over the past few years covered the viscosities of DESs using the ANN and gradient boosting (GB) [20,21]. However, only minimal effort has been put into using ML techniques to estimate the viscosities of numerous DESs over wide ranges of temperatures, and the topic deserves further investigation.

In this study, the least square support vector machine (LSSVM) and multilayer perceptron artificial neural network (MLPANN) are employed, as two robust and reliable ML algorithms, to predict the viscosities of different types of DESs. To this end, a comprehensive databank, containing 2533 data points pertaining to the viscosities of various families of DESs, was gathered from the literature in order to

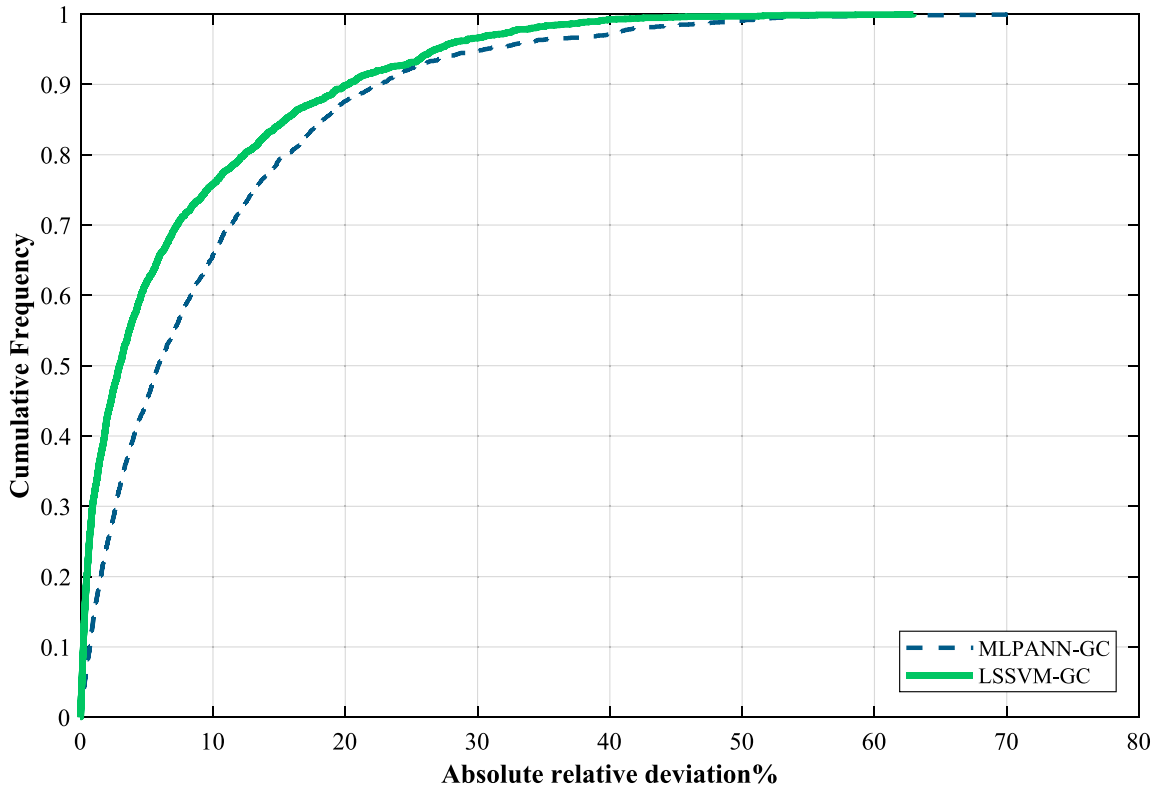

Fig. 2. Cumulative frequency of absolute relative deviation (ARD%) for the proposed MLPANN-GC and LSSVM-GC models.

develop and test the models' reliabilities. Furthermore, in order to enhance the models' predictabilities by considering the black-box nature of the LSSVM and MLPANN methods, the GC method was also incorporated as a hybrid into the LSSVM and MLPANN models.

2. Theory

This section provides a brief explanation of the two machine learning methods considered in this work, namely Artificial Neural Network (ANN) and Least Square Support Vector Machine (LSSVM), as well as the optimization strategies.

2.1. Artificial neural network (ANN)

As an intelligent computing strategy, the artificial neural network (ANN) is a modeling tool that involves mathematical calculations with similar characteristics to the human brain. Due to their simplicity and versatility, ANNs have found extensive use in various fields, including classification, pattern recognition, time series prediction, and function approximation. In addition to extracting complex information and nuanced patterns from the data, ANN can build empirical relationships between input and output variables [22]. The multilayer perceptron (MLP), is the most widely used class of ANN. In MLP, due to the information only moving in one direction from inputs to outputs, it can be considered as a fully connected class of feedforward ANN. The MLP structure consists of inputs, outputs, and one or more hidden layer(s) in between them. The input layer receives the input signals, which are then processed by the hidden layer(s) that determines the complex dependency between the dependent and independent variables. Finally, the output layer produces the model's output. Each layer contains a number of neurons, and each neuron contains its bias and weight parameters. It is the responsibility of the neurons to weigh and sum the incoming signals before sending them through the activation function in the neuron body [23]. Different activation functions, either linear or nonlinear, are applied to the MLP's hidden and output layers. Several of these activation functions are represented by the formulae given in Eqs. (9)–(11). The weights and biases are modified during the training process. The most common training algorithm used with MLPs is the Levenberg-Marquardt algorithm [24].

$$Linear : f(x) = x \tag{9}$$

$$\text{Tan}sig : f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{10}$$

$$Logsig : f(x) = \frac{1}{1 + e^{-x}} \tag{11}$$

Furthermore, Eq. (12) represents the MLP mathematical operation to generate an output, for example, for a three-layer neural network.

$$Y = f_2(W_2 f_1(W_1 x + b_1) + b_2) \tag{12}$$

Two weight matrices (W_1 and W_2) and two bias vectors (b_1 and b_2) are used in Eq. (12). Furthermore, the activation function f_1 and f_2 must be determined.

2.2. Support vector Machine (SVM)

One of the powerful supervised learning techniques that may be applied to classification and regression problems is the support vector machine (SVM). SVM differs significantly from neural networks in that it does not require prior knowledge of the network architecture and has a theoretical background that allows for excellent generalization capabilities [25]. SVMs have higher computing costs since they are calculated by solving quadratic programming problems. This is one of the main disadvantages of this approach [26]. Suykens and Vandewalle [27], who enhanced the SVM techniques, suggested using the Least Squares Support Vector Machine (LSSVM). The LSSVM employs linear equations and reduces the complexity of the optimization process in contrast to the original SVM, which requires the implementation of a nonlinear and quadratic set of equations [28]. Tuning and training are

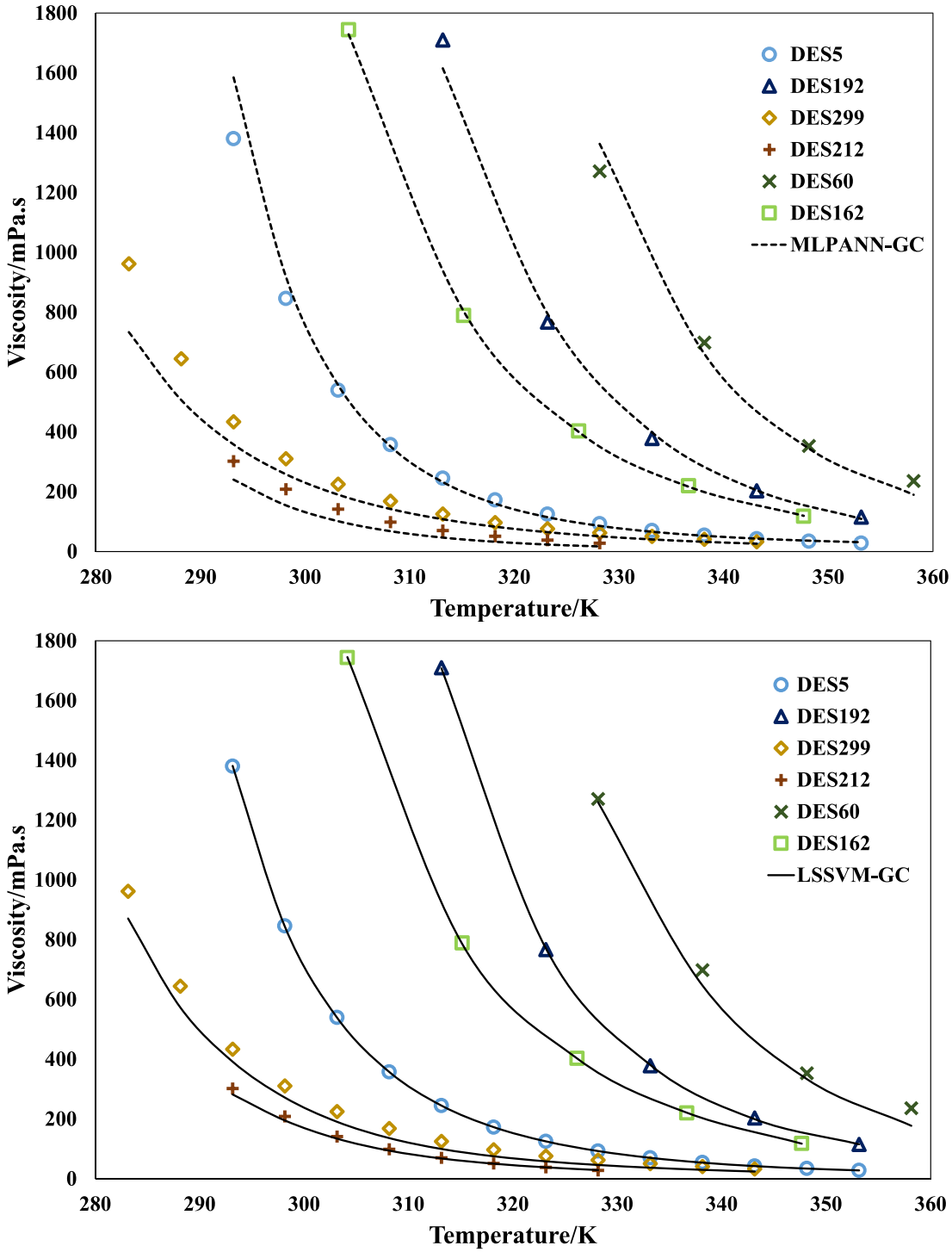

Fig. 3. Comparison of the experimental viscosity trend for six randomly selected DESs at various temperatures with the estimated viscosity from the suggested models. DES5, acetyl choline chloride: d-fructose (1:1); DES192, potassium carbonate: glycerol (1:6); DES299, benzyltributylammonium chloride: diethylene glycol (1:3); DES212, lidocaine: decanoic acid (1:3), DES60, choline chloride: d-fructose (1.5:1); DES162, choline chloride: chromic chloride hexahydrate (1:2.45).

the two stages of the model generation process for LSSVM-based models. In the case of the radial basis function (RBF) kernel, the regularization parameters of LSSVM, namely the regularization constant γ and the kernel function parameter σ^2, are tuned using the simplex algorithm [29], with the initial values for both parameters coming from coupled simulated annealing (CSA) [30,31]. The final model is developed after LSSVM training by solving a linear system of equations. Further details of the LSSVM method are provided in the Supplementary Material.

2.3. Model assessment

Statistical metrics, such as average absolute relative deviation percent (*AARD%*), determination coefficient (R^2), root mean square error (*RMSE*), absolute relative deviation percent (*ARD%*), and mean square error (*MSE*) were utilized to assess the validity of the proposed models. These statistical metrics are given below,

Table 3
The probability percent of the presence of AARD% in the specified range for two proposed models and two literature models.

Range of AARD%	Proposed models		Literature models	
	MLPANN-GC	LSSVM-GC	Bakhtyari et al. [12]	Yu et al.[13]
0–4	27.87	52.46	27.87	14.75
4–8	27.87	17.70	25.25	24.26
8–12	17.05	10.16	16.39	17.38
12–16	11.48	6.89	10.16	10.16
16–20	5.57	2.95	6.89	4.26
More than 20	10.16	9.84	13.44	29.18

$$AARD\% = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{\eta_i^{Calc}-\eta_i^{Exp}}{\eta_i^{Exp}}\right| \tag{13}$$

$$R^2 = 1-\frac{\sum_{i=1}^{N}\left(\eta_i^{Calc}-\eta_i^{Exp}\right)^2}{\sum_{i=1}^{N}\left(\bar{\eta}-\eta_i^{Exp}\right)^2} \tag{14}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\eta_i^{Calc}-\eta_i^{Exp}\right)^2} \tag{15}$$

$$ARD\% = 100\left|\frac{\eta_i^{Calc}-\eta_i^{Exp}}{\eta_i^{Exp}}\right| \tag{16}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\eta_i^{Calc}-\eta_i^{Exp}\right)^2 \tag{17}$$

where N signifies the number of data points and η^{Exp} refers to the experimentally determined viscosity values of DESs. $\bar{\eta}$ is the average of the experimental viscosity data and η^{Calc} stands for the estimated viscosity values by the model.

2.4. Optimization techniques

2.4.1. Levenberg-Marquardt

The Levenberg-Marquardt (LM) approach is one of the effective algorithms for optimizing the weights and biases of feed-forward neural networks because of its fast convergence [32,33]. It is significant to point out that LM may often find only the local minima, instead of ensuring that a global minimum has been determined [34]. The Hessian matrix needs not to be calculated for the LM algorithm [35]. When the performance function takes the form of a sum of squares, the gradient is determined, and the Hessian matrix is approximately calculated as follows [35,36].

$$H = J^T J \tag{18}$$

$$g = J^T e \tag{19}$$

In the preceding equations, J denotes a Jacobian matrix with first-order derivatives of network errors with respect to weights and biases, and e represents a vector of network errors. In the following Newton update, the LM employs the previously indicated Hessian matrix approximation,

$$x_{n+1} = x_n - \left(J^T J - \lambda I\right)^{-1} J^T e \tag{20}$$

where λ stands for a constant and x represents connection weights. After each successful step, λ falls, while it only rises when a tentative step improves the performance function.

2.4.2. Bayesian regularization algorithm

As a training technique, the Bayesian regularization (BR) algorithm adjusts weights and biases by LM optimization [37]. In order to create a network that generalizes well, it first identifies the best optimal combination of weights and squared errors [37]. The following shows how BR incorporates network weights into the cost function [38],

$$F(w) = \nu E_w + \kappa E_D \tag{21}$$

where E_w and E_D represent the sum of the squared network errors and the sum of the network weights, respectively, and $F(w)$ is the objective function. The parameters of the objective function are considered to be ν and κ, and based on Bayes's theorem, these factors are defined. In previous literature, it was discussed in detail how BR could be used to train neural networks [38].

2.4.3. Resilient backpropagation

Tangent and logarithmic sigmoid transfer functions are commonly utilized in the hidden layers of multilayer artificial neural networks. The domain of these functions is between negative and positive infinity, while the outputs of these functions have a limited range. As the functions increase monotonically, and as the values of the inputs increase, their slopes get closer to zero. As a result, biases and weights may slightly change when the network is trained using the steepest descent because the gradient has a low value [35]. Resilient Backpropagation (Rprop) eliminates the negative impact of the partial derivative on the values of weights and biases. The weight update can only be directed in a certain direction based on the sign of the derivative. The weight update is unaffected by the magnitude of the derivative. Other references provide more information on this algorithm [39].

2.4.4. Coupled simulated annealing

Simulated Annealing (SA) is a stochastic optimization method to solve nonconvex problems. SA imitates the annealing procedures, which are employed in metallurgy to harden glass and metals. These materials are heated to a high temperature during this procedure, then cooled over time to get to a state of low energy [40]. Coupled Simulated Annealing (CSA) is a class of optimization techniques based on SA with the intention of increasing SA accuracy without significantly reducing convergence speed and making it simpler to escape local optima. The generation and acceptance of a single current state are required for each optimization step in CSA, and each of these operations is carried out separately. The acceptance probability is the only difference between SA and CSA processes [30].

3. Investigated systems

Most models must have the smallest residual distance between experimental and calculated points instead of traversing all of the given data points [41]. Sometimes, overfitting is one of the issues that arise for models that are developed using ML techniques. Additionally, overfitting happens when a model learns unnecessary information, memorizes noise, and fits a set of data too closely. As a result, the model cannot generalize well to new data. One way to prevent the problem of overfitting is to use a sufficiently large data bank. Consequently, in the current study, 305 DESs (2533 data) over the wide temperature range of 277.15 to 373.15 K were collected from the literature and used to develop the models. In this databank, the DESs with exceptionally high viscosities (higher than 2000 mPas) were excluded. The databank was then divided into two subsets: the training set, which makes around 80%

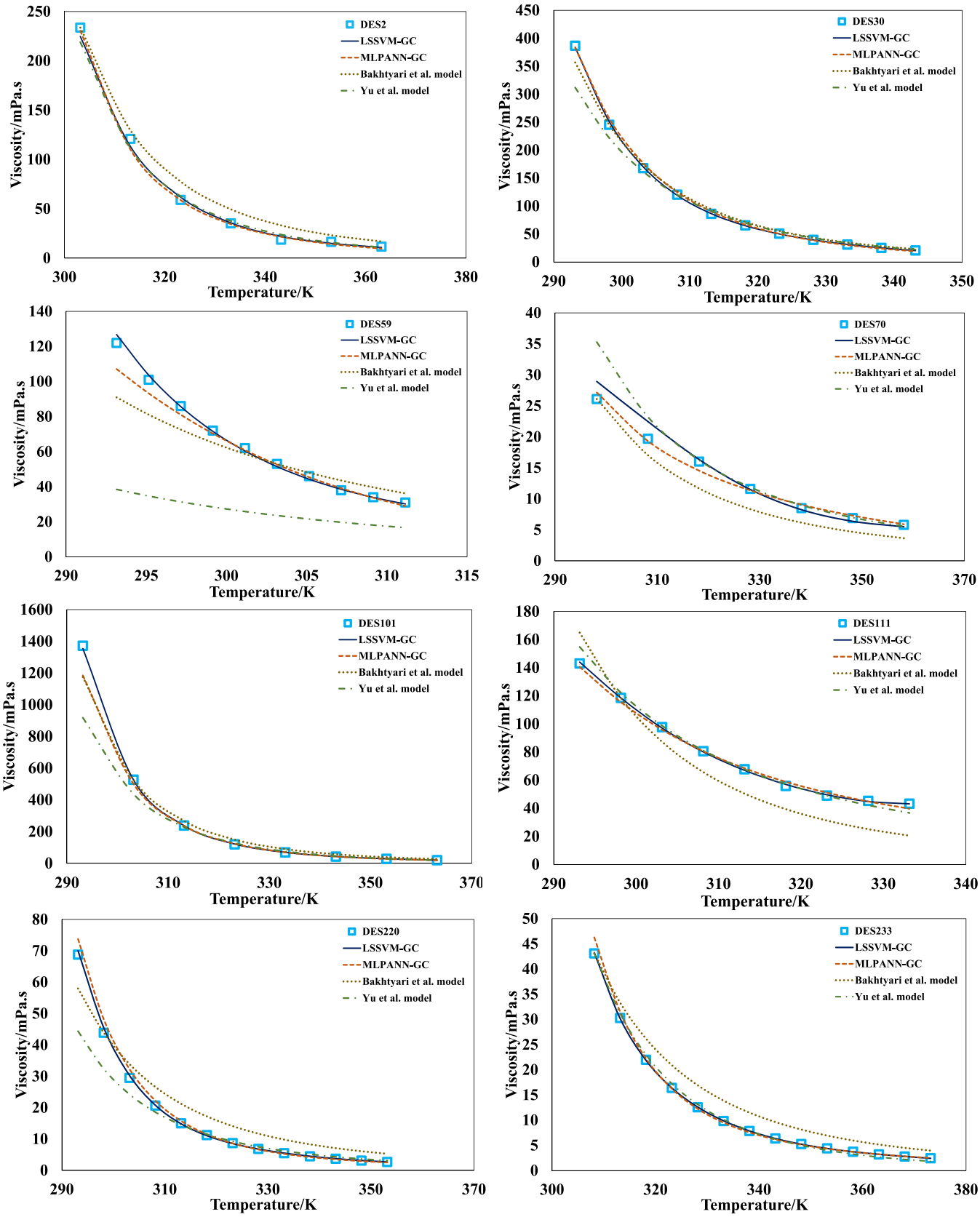

Fig. 4. Comparison of the viscosity-temperature trends for the proposed MLPANN-GC and LSSVM-GC models and also two literature models for eight random DESs consisting of: DES2, Acetyl choline chloride: Imidazole (1:1.5); DES30, Betaine: DL-lactic acid (1: 5); DES59, Choline chloride:2,3-Butanediol (1:9); DES70, Choline chloride: Ethylene Glycol (1:3); DES101, Choline chloride: Urea (1:2); DES111, Methyltriphenyl phosphonium bromide: Acetic acid (1:4); DES220, menthol: Thymol (2:1); DES233, Thymol: L-borneol (1:1).

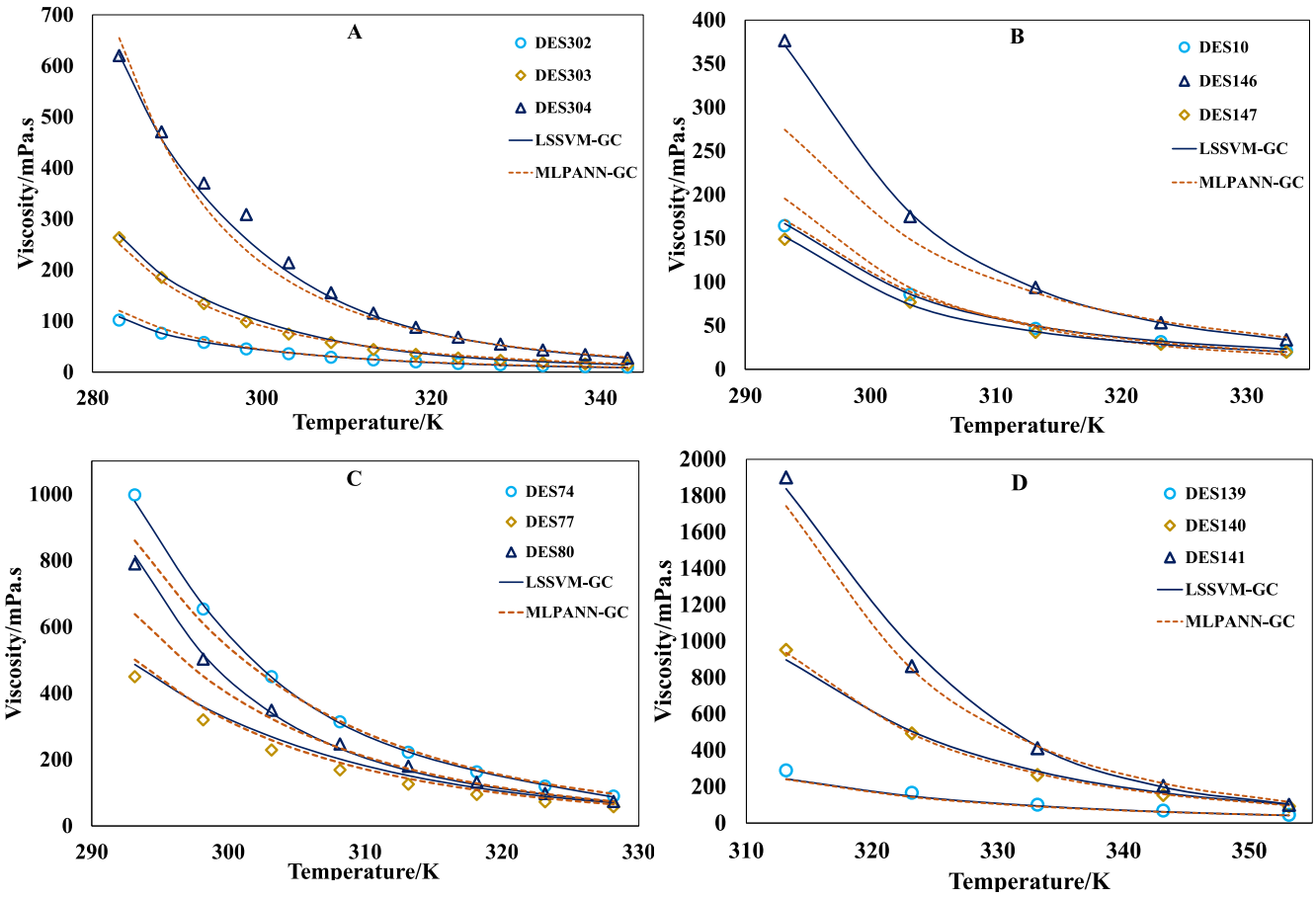

Fig. 5. Comparison of viscosity behavior of: A) various HBDs: in DES302, benzyltrimethylammonium chloride: ethylene glycol (1:3); DES303, benzyltrimethylammonium chloride: diethylene glycol (1:3); DES304, benzyltrimethylammonium chloride: triethylene glycol (1:3), B) various HBAs in: DES10, acetyl choline chloride: levulinic acid (1:3), DES146, tetrabutyl ammonium bromide: levulinic acid (1:3); DES147, tetraethylammonium bromide: levulinic acid (1:3), C) HBD molar ratios, DES74, choline chloride: glycerol (1:19), DES77, choline chloride: glycerol (1:3), DES80, choline chloride: glycerol (1:9), D) HBA molar ratios, DES139, tetrabutyl ammonium chloride: triethylene glycol (1:1), DES140, tetrabutyl ammonium chloride: triethylene glycol (2:1), DES141, tetrabutyl ammonium chloride: triethylene glycol (4:1).

of the entire dataset and contains 245 DESs with 2026 data points, and the test set, which makes around 20% of the whole data set and comprises the 60 remaining DESs with 507 data points. The first subset is used to develop the models and fine-tune the models' parameters, while the second set is used to evaluate the models' performances and predictabilities. The utilized DESs (HBAs and HBDs) are presented in Table S1 of the Supplementary Material, where each DES is also labeled as being used in the training or test dataset. Additionally, Table S5 of the Supplementary Material contains the details for each data point including the names of the HBAs, HBDs, molar ratios, temperature, experimental and calculated viscosities, and the uncertainty of the data.

4. Results and discussion

Based on the training subset (2026 data points), the model structure and tuning parameter adjustments were carried out. Also, according to the testing subset (507 data points), the created models' validities and overfitting issues were tracked.

Since this study considers the two methods of MLPANN and LSSVM, combined with group contribution to predict the viscosities of DESs, we have chosen to name these approaches as MLPANN-GC and LSSVM-GC, respectively. The functional groups of the various HBAs and HBDs are shown in Table 1.

The structures of the HBA/HBD, their molar ratios, and temperature serve as the two models' input parameters. The following definitions apply to the input vector X_1 and X_2 and the output variable Y,

$$X_1 = T(K) \tag{22}$$

$$X_2 = m_{HBA}(n_i) + m_{HBD}(n_j) \quad (i,j = 1,2,\ldots,nG) \tag{23}$$

$$Y = \ln(\eta/mPa.s) \tag{24}$$

where the molar ratios of the substances HBA and HBD, respectively, are denoted by m_{HBA} and m_{HBD}. The terms n_i and n_j indicate how frequently the functional groups of i and j occur in the HBA and HBD, respectively. For the HBA and HBD compounds, nG represents the overall number of defined groups.

The input variable distributions are normalized to fall between 0 and + 1 by applying Eq. (25), which uses the maximum and minimum values for each input data,

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{25}$$

x refers to the individual data point, x_{min} and x_{max} represent the minimum and maximum of each input feature, and x_{norm} stands for the corresponding normalized data point.

4.1. Model development

The proposed MLPANN-GC and LSSVM-GC models were used to predict the viscosities of DESs. This study used BR, LM, and Rprop to optimize the MLP parameters. For the MLP models, different structures

with varying numbers of hidden layers, numbers of neurons in the hidden layers, and a combination of applied transfer functions were examined to find an optimal topology. In most of the MLP models developed in this study, one hidden layer with a Logsig transfer function and a linear transfer function in the output layer performed better than other structures when statistical parameters were taken into account. The 10-fold cross-validation technique was used to determine the optimal number of hidden neurons for the MLP models. This technique trained and cross-validated the MLPs with different numbers of hidden neurons. The architecture with the lowest cross-validation error was then chosen.The mean squared error performance function was minimized during the neural network training. To develop the best model from the training data, the performance function was iteratively minimized by adjusting the weight matrix. It was found that the optimal architectures for the MLPANN-BR, MLPANN-LM, and MLPANN- Rprop were 46–10-1, 46–8-1, and 46–13-1, respectively. The first number represents the number of inputs, the second one is the number of neurons in the hidden layer, and the last indicates the number of outputs. The statistical parameters for all of the developed models, including *AARD%*, *RMSE*, and R^2, are summarized in Table 2. This table shows that the model developed based on the BR yields more accurate results than LM and Rprop. The application of Bayes theorem and LM by the BR in its computing process may be the cause of the better performance of the MLPANN that is optimized with BR. Due to the initial weights and biases being chosen at random, the training of the MLP networks was carried out more than 90 times, to ultimately select the best models. The weights and biases of the hidden and output layers of the MLPANN-GC model, which were optimized with BR, are shown in Table S3 of the Supplementary Material for the trained MLPANN network. For the LSSVM model, the radial basis function (RBF), compared to other kernels, gave more precise results in this nonlinear system, which is why it is used in this study. As a result, the tuning of the parameters of LSSVM (i.e., γ and σ^2) were carried out in two steps. First, CSA selected the appropriate initial values. Then, using the benchmark of MSE with 10-fold cross-validation, these parameters were passed through a grid-search optimization technique to execute a fine-tuning step. The optimal values for γ and σ^2 were 2434020.93 and 5.75, respectively. Furthermore, the bias term, *b,* and the Lagrange multipliers, α, were calculated for the models. The bias term has a value of -8.8 and the α parameter is a vector with a size of $S \times 1$ (S is the number of training examples). The values of the parameters are reported in Table S4 of the Supplementary Material. Table 2 displays the result of this model.

In what is to follow, several graphical perspectives are given to assess the results of the developed MLPANN-GC and LSSVM-GC models for DES viscosity estimations. Fig. 1a and 1b show the predicted values versus experimental data. Distinguish is made between the training and test data sets. The rather close vicinity of the majority of data points to the 45° line for both the training and testing data sets at lower viscosities indicates not only the accuracy, but also the predictability of both models when the viscosity is low. However, as viscosity increases, the accuracies diminish. The lack of accuracy of the models at high viscosities may result from the uncertainty of the experimental data, inherent in the various viscosity measurement methods. Additionally, machine learning algorithms may possibly classify some of these high values as outliers, which would affect the overall data distribution and impede the algorithms' effectiveness. In Fig. 1, the relative errors between the output of the developed models and the corresponding experimental data are plotted. These figures illustrate that the region bound by relative deviations of ± 30% contains the majority of the relative deviations of both models. However, this figure also shows that the LSSVM-GC model has lower relative errors than the MLPANN-GC model. Fig. 2 presents the cumulative frequency of absolute relative error (*ARD%*) for the proposed models. This figure shows that more than 85% of the predicted viscosities by the MLPANN-GC model, and approximately 90% by the LSSVM-GC model have an *ARD%* of less than 20%. As evidenced by the lower estimated errors of the LSSVM-GC model, this confirms its ability to solve nonlinear regression problems.

To ensure whether the developed models follow the expected trends with temperature variations, the results of the models are plotted versus temperature, alongside the experimental viscosities, for six random training and test DESs in Fig. 3. Acceptable agreement between the experimental viscosity trends and the model estimations can be seen in this figure, which demonstrates how well the developed models could follow the behavior of the experimental data in response to temperature variations.

The models of Bakhtyari et al. [12] and Yu et al. [13], presented in Eqs. (1)–(8), were also taken into consideration to compare to the proposed models of this study. Table S2 provides the results of the calculated *AARD%* values for all of these models. Table 3 also includes a brief report on the probability percentage of the presence of AARD% in the given range. Even though the model of Bakhtyari et al. [12] is rather easy to use, it does require on experimental reference viscosity data point at a reference temperature as its input data. Providing an experimental viscosity data, of course, can have a significant positive impact on the results. Despite this, the average errors of the Bakhtyari et al. model are not better than the global models of this work that require absolutely no experimental data input. This also limits the applications of their model to only those DESs for which experimental viscosity data have already been measured. For the Yu et al. [13] model, unacceptable results were obtained for DESs that were not included in their databank. Because of the large number of parameters of Yu et al.'s model and the requirement that users have a thorough understanding of the HBA and HBD structures (including the configurations of chiral carbons, ortho, meta, and para groups and the types of DESs), it can be a challenging model for users, since it is not straightforward. Moreover, this model is highly susceptible to the input variables. Fig. 4 illustrates the exponential decrease of viscosity versus temperature of the proposed and literature models for DES2, DES59, DES101, DES220, DES30, DES70, DES111, and DES233. The results of our models, particularly the LSSVM-GC model, rather closely matched the experimental data and outperformed the other two models in the literature.

Based on the nature of the HBA's and HBD's, DES viscosities vary significantly. Fig. 5 shows the influence of the components and their molar ratios on DESs viscosities. Fig. 5A shows the viscosity behavior as a function of temperature, with the fixed HBA of benzyl trimethylammonium chloride, but with various HBDs, including ethylene glycol, diethylene glycol, and triethylene glycol, where the HBA to HBD molar ratio is 1:3. The viscosity values for benzyl trimethylammonium chloride-based DESs with various HBDs decrease in the order triethylene glycol > diethylene glycol > ethylene glycol. It can be seen that the viscosity values of these DESs rise systematically by adding –OH functional groups. Fig. 5B depicts how various HBAs (namely, acetylcholine chloride, tetrabutyl ammonium bromide, and tetraethylammonium bromide) affect the viscosity of DESs consisting of the fixed HBD of levulinic acid. The molar ratio of HBA to HBD is 1:3. It is seen that the trend of viscosity shows greater changes in the tetrabutyl ammonium chloride-based DES, followed by the acetylcholine chloride and the tetraethyl ammonium bromide-based DESs. Perhaps a large HBA molecule increases the surface area where the molecules come into contact with one another, which helps to decrease the DES molecular mobility and, as a result, increases its viscosity. The influences of HBD and HBA molar ratios on viscosities are shown in Fig. 5C and 5D. In Fig. 5C, the DESs composed of choline chloride as the HBA and glycerol as the HBD are considered at three different molar ratios (1:3, 1:9, and 1:19). This figure indicates how increased viscosities result for DESs with increased ratios of glycerol. Higher viscosities at higher glycerol concentrations may be due to the stronger molecular interactions between the HBA and HBD. Additionally, Fig. 5D shows the DESs made of tetrabutyl

ammonium chloride as the HBA and triethylene glycol as the HBD at 1:1, 2:1, and 4:1 M ratios. It is also evident that increasing the amount of tetra butyl ammonium chloride in this group of DESs increases viscosity. Perhaps this is because when more HBA is added, the fractional free volume in the DES decreases.

5. Conclusions

In the current study, the group contribution method was combined with two machine learning techniques (MLPANN and LSSVM) to predict the viscosities of a wide range DESs of different natures. The proposed viscosity models are functions of temperature, HBA/HBD molar ratio, and weights of the functional groups. These input parameters were selected with the aim to enhance efficiency and predictability. The GC method has the theoretical foundation that each functional group within a chemical structure imparts a specific weight to a given property. As a result, this method can significantly enhance the predictibality of a model. In order to develop the proposed models based on the group contribution theory, a large data bank of 2533 data points from 305 DESs was collected from literature sources. Various graphical and statistical analyses indicate that the proposed models can rather reliably estimate experimental data and even predict viscosities when the viscosities are not too high. The MLPANN-GC model had R^2 and *AARD%* values of 0.98 and 9.43%, while the LSSVM-GC model resulted in corresponding values of 0.99 and 6.88%, respectively. The performance of the different MLP training algorithms is ranked as follows: BR > LM > Rprop. When compared to previous published models in the literature, the key benefit of the methodology utilized in this study is that in addition to relative accuracy and the high calculation speed of the machine learning methods, the incorporation of the GC method into the model significantly boosted the models' predictabilities, even allowing for the screeting of DESs which have not yet been made in the labs. To find the two crucial tuning parameters, γ and σ^2 of the LSSVM method, the coupled simulated annealing and gridsearch optimizer were employed. For chemists and chemical engineers, these developed models, particularly the LSSVM-GC model, can be of great assistance in providing straightforward predictive tools for estimating DES viscosities in the face of challenging thermodynamics techniques. Additionally, when the proposed models were compared to correlations in the literature, it was shown that the developed models have acceptable precision.

CRediT authorship contribution statement

Ahmadreza Roosta: Data curation, Investigation, Methodology, Formal analysis, Software, Writing – original draft. **Reza Haghbakhsh:** Supervision, Conceptualization, Data curation, Investigation, Methodology, Formal analysis, Software, Validation, Writing – review & editing, Supervision. **Ana Rita C. Duarte:** Funding acquisition, Project administration. **Sona Raeissi:** Supervision, Writing – review & editing, Validation, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The authors wish to thank Shiraz University, University of Isfahan and Universidade Nova de Lisboa for the facilities provided.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2023.122747.

References

[1] W. Zhao, Make the chemical industry clean with green chemistry: an interview with Buxing Han, Nat. Sci. Rev. 5 (2018) 953–956, https://doi.org/10.1093/nsr/nwy045.
[2] M.H. Zainal-Abidin, M. Hayyan, W.F. Wong, Hydrophobic deep eutectic solvents: current progress and future directions, J. Ind. Eng. Chem. 97 (2021) 142–162, https://doi.org/10.1016/j.jiec.2021.03.011.
[3] M.B. Singh, V.S. Kumar, M. Chaudhary, P. Singh, A mini review on synthesis, properties and applications of deep eutectic solvents, J. Indian Chem. Soc. 98 (2021), 100210, https://doi.org/10.1016/j.jics.2021.100210.
[4] A.P. Abbott, G. Capper, D.L. Davies, R.K. Rasheed, V. Tambyrajah, Novel solvent properties of choline chloride/urea mixtures, Chem. Commun. (2003) 70–71, https://doi.org/10.1039/B210714G.
[5] F.P. Pelaquim, A.M. Barbosa Neto, I.A.L. Dalmolin, M.C. da Costa, Gas solubility using deep eutectic solvents: review and analysis, Ind. Eng. Chem. Res. 60 (2021) 8607–8620, https://doi.org/10.1021/acs.iecr.1c00947.
[6] R. Bernasconi, G. Panzeri, G. Firtin, B. Kahyaoglu, L. Nobili, L. Magagnin, Electrodeposition of ZnNi alloys from choline chloride/ethylene glycol deep eutectic solvent and pure ethylene glycol for corrosion protection, J. Phys. Chem. B 124 (2020) 10739–10751, https://doi.org/10.1021/acs.jpcb.0c04784.
[7] Y.P. Mbous, M. Hayyan, A. Hayyan, W.F. Wong, M.A. Hashim, C.Y. Looi, Applications of deep eutectic solvents in biotechnology and bioengineering—Promises and challenges, Biotechnol. Adv. 35 (2017) 105–134, https://doi.org/10.1016/j.biotechadv.2016.11.006.
[8] H.G. Morrison, C.C. Sun, S. Neervannan, Characterization of thermal behavior of deep eutectic solvents and their potential as drug solubilization vehicles, Int. J. Pharm. 378 (2009) 136–139, https://doi.org/10.1016/j.ijpharm.2009.05.039.
[9] A.P. Abbott, P.M. Cullis, M.J. Gibson, R.C. Harris, E. Raven, Extraction of glycerol from biodiesel into a eutectic based ionic liquid, Green Chem. 9 (2007) 868–887, https://doi.org/10.1039/B702833D.
[10] Y. Nahar, S.C. Thickett, Greener, faster, stronger: The benefits of deep eutectic solvents in polymer and materials science, Polymers 13 (2021) 1–24, https://doi.org/10.3390/polym13030447.
[11] A. Boisset, S. Menne, J. Jacquemin, A. Balducci, M. Anouti, Deep eutectic solvents based on N-methylacetamide and a lithium salt as suitable electrolytes for lithium-ion batteries, PCCP 15 (2013) 20054–20063, https://doi.org/10.1039/C3CP53406E.
[12] A. Bakhtyari, R. Haghbakhsh, A.R.C. Duarte, S. Raeissi, A simple model for the viscosities of deep eutectic solvents, Fluid Phase Equilib. 521 (2020), 112662, https://doi.org/10.1016/j.fluid.2020.112662.
[13] L.Y. Yu, X.J. Hou, G.P. Ren, K.J. Wu, C.H. He, Viscosity model of deep eutectic solvents from group contribution method, AIChE J (2022), https://doi.org/10.1002/aic.17744 e17744.
[14] F.S. Mjalli, J. Naser, Viscosity model for choline chloride-based deep eutectic solvents, Asia Pac. J. Chem. Eng. 10 (2015) 273–281, https://doi.org/10.1002/apj.1873.
[15] R. Haghbakhsh, S. Raeissi, K. Parvaneh, A. Shariati, The friction theory for modeling the viscosities of deep eutectic solvents using the CPA and PC-SAFT equations of state, J. Mol. Liq. 249 (2018) 554–561, https://doi.org/10.1016/j.molliq.2017.11.054.
[16] T. Lemaoui, A.S. Darwish, A. Attoui, F. Abu Hatab, N.E.H. Hammoudi, Y. Benguerba, L.F. Vega, I.M. Alnashef, Predicting the density and viscosity of hydrophobic eutectic solvents: Towards the development of sustainable solvents, Green Chem. 22 (2020) 8511–8530, https://doi.org/10.1039/D0GC03077E.
[17] V. Van Speybroeck, R. Gani, R.J. Meier, The calculation of thermodynamic properties of molecules, Chem. Soc. Rev. 39 (2010) 1764–1779, https://doi.org/10.1039/B809850F.
[18] R. Gani, Group contribution-based property estimation methods: advances and perspectives, Curr. Opin. Chem. Eng. 23 (2019) 184–196, https://doi.org/10.1016/j.coche.2019.04.007.
[19] L.Y. Yu, G.P. Ren, X.J. Hou, K.J. Wu, Y. He, Transition state theory-inspired neural network for estimating the viscosity of deep eutectic solvents, ACS Cent. Sci. 8 (2022) 983–995, https://doi.org/10.1021/acscentsci.2c00157.
[20] I. Adeyemi, M.R.M. Abu-Zahra, I.M. AlNashef, Physicochemical properties of alkanolamine-choline chloride deep eutectic solvents: measurements, group contribution and artificial intelligence prediction techniques, J. Mol. Liq. 256 (2018) 581–590, https://doi.org/10.1016/j.molliq.2018.02.085.
[21] X. Xu, J. Range, G. Gygli, J. Pleiss, Analysis of thermophysical properties of deep eutectic solvents by data integration, J. Chem. Eng. Data 65 (2020) 1172–1179, https://doi.org/10.1021/acs.jced.9b00555.
[22] R. Sadiq, M.J. Rodriguez, H.R. Mian, Empirical models to predict disinfection by-products (DBPs) in drinking water: An updated review, 2nd ed., Elsevier Inc., 2019. 10.1016/B978-0-12-409548-9.11193-5.
[23] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Artificial Neural Networks, Data Handling in Science and Technology. 20 (1998) 649–699, https://doi.org/10.1016/S0922-3487(98)80054-3.

[24] A.R. Behesht Abad, S. Mousavi, N. Mohamadian, D.A. Wood, H. Ghorbani, S. Davoodi, M.A. Alvar, K. Shahbazi, Hybrid machine learning algorithms to predict condensate viscosity in the near wellbore regions of gas condensate reservoirs, J. Nat. Gas Sci. Eng. 95 (2021), 104210, https://doi.org/10.1016/j.jngse.2021.104210.
[25] H. Liu, X. Yao, R. Zhang, M. Liu, Z. Hu, B. Fan, Accurate quantitative structure−property relationship model to predict the solubility of C60 in various solvents based on a novel approach using a least-squares support vector machine, J. Phys. Chem. B 109 (2005) 20565–20571, https://doi.org/10.1021/jp052223n.
[26] H. Wang, D. Hu, Comparison of SVM and LS-SVM for regression, Proceedings of 2005 International Conference on Neural Networks and Brain Proceedings, ICNNB'05. 1 (2005) 279–283. 10.1109/ICNNB.2005.1614615.
[27] J.A.K. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Processing Letters 1999 9:3. 9 (1999) 293–300. 10.1023/A:1018628609742.
[28] A. Hemmati-Sarapardeh, S. Hatami, H. Taghvaei, A. Naseri, S.S. Band, K. wing Chau, Designing a committee of machines for modeling viscosity of water-based nanofluids, Engineering Applications of Computational Fluid Mechanics. 15 (2021) 1967–1987. 10.1080/19942060.2021.1979099.
[29] J.A. Nelder, R. Mead, A simplex method for function minimization, Comput. J. 7 (1965) 308–313, https://doi.org/10.1093/comjnl/7.4.308.
[30] S. Xavier-De-Souza, J.A.K. Suykens, J. Vandewalle, D. Bolle, Coupled simulated annealing, IEEE. 40 (2010) 320–335, https://doi.org/10.1109/TSMCB.2009.2020435.
[31] K. Paduszyński, Extensive Databases and Group Contribution QSPRs of Ionic Liquids Properties. 1. Density, Ind. Eng. Chem. Res. 58 (2019) 5322–5338, https://doi.org/10.1021/acs.iecr.9b00130.
[32] Z. Yan, S. Zhong, L. Lin, Z. Cui, R. Alonso, A. Bustos Caballero, J. Meneses Alonso, E. Soriano-Heras, Adaptive Levenberg–Marquardt Algorithm: A New Optimization Strategy for Levenberg–Marquardt Neural Networks, Mathematics. 9 (2021) 2176, https://doi.org/10.3390/math9172176.
[33] A.P. Piotrowski, J.J. Napiorkowski, Optimizing neural networks for river flow forecasting – Evolutionary Computation methods versus the Levenberg–Marquardt approach, J. Hydrol. 407 (2011) 12–27, https://doi.org/10.1016/j.jhydrol.2011.06.019.
[34] A. Hemmati-Sarapardeh, A. Varamesh, M.M. Husein, K. Karan, On the evaluation of the viscosity of nanofluid systems: modeling and data assessment, Renew. Sustain. Energy Rev. 81 (2018) 313–329, https://doi.org/10.1016/j.rser.2017.07.049.
[35] O. Kisi, E. Uncuoğlu, Comparison of three back-propagation training algorithms for two case studies, Indian Journal of Engineering and Materials Sciences. 12 (2005) 434–442.
[36] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Netw. 5 (1994) 989–993, https://doi.org/10.1109/72.329697.
[37] X. Pan, B. Lee, C. Zhang, A comparison of neural network backpropagation algorithms for electricity load forecasting, Proceedings - 2013 IEEE International Workshop on Intelligent Energy Systems, IWIES 2013 (2013) 22–27, https://doi.org/10.1109/IWIES.2013.6698556.
[38] Z. Yue, Z. Songzheng, L. Tianshi, Bayesian regularization BP neural network model for predicting oil-gas drilling cost, BMEI 2011 - Proceedings 2011 International Conference on Business Management and Electronic Information. 2 (2011) 483–487. 10.1109/ICBMEI.2011.5917952.
[39] M. Riedmiller, H. Braun, A direct adaptive method for faster backpropagation learning: the RPROP algorithm, in: 1993 IEEE International Conference on Neural Networks, 1993, pp. 586–591.
[40] S. Russell, P. Norvig, Artificial intelligence : a modern approach, 4th ed., Pearson, n.d.
[41] S. Koutsoukos, F. Philippi, F. Malaret, T. Welton, A review on machine learning algorithms for the ionic liquid chemical space, Chem. Sci. 12 (2021) 6820–6843, https://doi.org/10.1039/D1SC01000J.