Simultaneous Prediction of the Density and Viscosity of the Ternary System Water-Ethanol-Ethylene Glycol Using Support Vector Machine

Ehsan Kianfar

Department of Chemical Engineering, Arak Branch, Islamic Azad University, Arak, Iran
Email: ehsan_kianfar2010@yahoo.com, e-kianfar94@iau-arak.ac.ir

Abstract: Ethylene glycol is an organic solvent used in extractive distillation to separate water-ethanol mixtures. An appropriate process description requires accurate physical property data. In this paper, experimental liquid densities and dynamic viscosities of pure ethylene glycol as well as the ternary system water-ethanol-ethylene glycol are presented over a wide temperature range (298.15 to 328.15 K) at atmospheric pressure. Density and viscosity data of the ternary system of water-ethanol-ethylene glycol solutions were obtained from the literature and a support vector machine model was used to predict the density and viscosity of this system. The determination coefficients for density and viscosity of the ternary system of water-ethanol-ethylene glycol are 0.9854 and 0.9892, respectively. The mean square errors of density and viscosity are obtained 4.6572×10^{-4} and 3.4920×10^{-4}, respectively. The results confirmed that the proposed method can predict the density and viscosity of the ternary system of water-ethanol-ethylene glycol as a function of temperature, using a support vector machine.

Keywords: water, ethanol, ethylene glycol, SVM support vector machines, density, viscosity

1. Introduction

Ethylene glycol (EG) is an organic solvent which is used as an antifreeze, a heat transfer fluid, a precursor of polymers and which is envisaged for CO_2 absorption[1-2]. Besides many other applications, this solvent is used as a trainer for the separation of water-ethanol mixtures using extractive distillation[3-9]. To design or simulate this process, sound knowledge of the physical and transport properties of this ternary mixture is important. Previous work has demonstrated that in the extractive distillation process, the mass transfer efficiency is strongly affected by the viscosity of the liquid phase[10]. For this reason, reliable mass transfer modeling requires accurate viscosity data. Dynamic viscosities and densities of the binary mixture water-EG at several temperatures are well known[11-16]. To our knowledge, experimental ternary dynamic viscosities and densities as function of temperature are only published in a previous study for a limited range of compositions and temperatures[17-20] in spite of the importance of EG in thermal separations. This limitation could not give a good representation of these properties in an extractive distillation column. In principle, these ternary dynamic viscosities at any temperature can be obtained by summing the mole fractions of the pure dynamic viscosities assuming ideal mixture behavior.

In this paper, based on the linear combination, the decisions on the category and regression have been adopted. The supported vector machine used in this work belongs to the Grendel method, which has systems for the efficient execution of linear learning machines in the kernel space so that they maintain the properties of optimality and generalizability. Thermodynamic properties of triple mixtures are considered the dependent variables, and viscosity and density variables as independent variables. So, variables of the backup vector machine model should be optimized.

2. Experimental

2.1. Methodology validation

The required data was found from the paper with temperature-based density and viscosity (Table 1)[21].

Temperature (k)	water	Ethanol (ρ/g_cm^{-3})	Ethylene glycol (ρ/g_cm^{-3})	density	viscosity
298.15	0.4024	0.2725	0.3251	0.98438	4.578
308.15	0.4024	0.2725	0.3251	0.97668	3.383
318.15	0.4024	0.2725	0.3251	0.96882	2.566
328.15	0.4024	0.2725	0.3251	0.96078	2.225
298.15	0.8369	0.1135	0.0496	0.97758	2.375

Table 1. Density and viscosity values for water and ethanol and ethylene glycol in terms of temperature[21]

Support-vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. The Support Vector Machine (SVM) algorithm is a popular machine learning tool that offers solutions for both classification and regression problems. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.

SVMs can be used to solve various real-world problems:

1. SVMs are helpful in text and hypertext categorization, as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings. Some methods for shallow semantic parsing are based on support vector machines.

2. Classification of images can also be performed using SVMs. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback. This is also true for image segmentation systems, including those using a modified version SVM that uses the privileged approach as suggested by Vapnik. Figure 1 shows the test design flowchart.
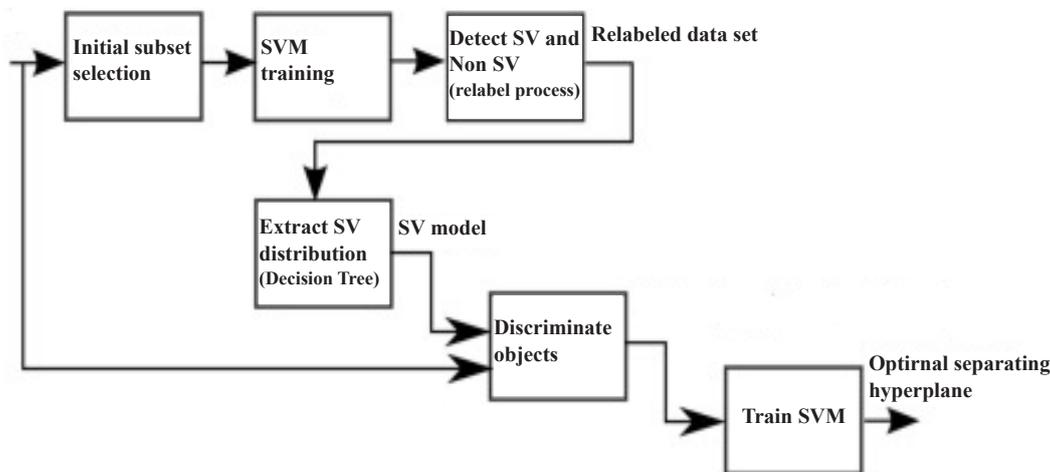
Figure 1. Flowchart for the design of experiments

3. Results and discussions

In this paper, the backup machine model was used to create the model in Assam, which is explained in the terms of determination coefficient (R^2), Mean Square Error (MSE), sigma, and gamma for density and viscosity in the training category and the test category. 78 data as inputs for this paper, 62 data were used as training groups and 16 data were used as test classes. Data input to the backup machine and computation by it were performed 15 times with the results obtained for the density and the viscosity and are shown in Tables 2 and 3. Given the obtained values, R^2, MSE, sigma, gamma, where the lowest MSE and the highest R^2 coefficient were obtained, as the optimum values in the calculation of the ninth stage of software implementation for the density and the fifth stage of software implementation for viscosity. The optimal R^2 values for the training and testing arm in both cases showed that there is a good correlation between the predictive and experimental values. Therefore, the proposed model can be a successful model for predicting the density and viscosity of the triple system of water, ethanol, and ethylene glycol at the same time. Figure 2 shows the correlation between the experimental data and the predicted values of the trivalent density of water, ethanol, and ethylene glycol by the proposed

model. The determination coefficient for the density of this system was obtained in the 1.0000 instruction class, and as shown in Figure 3, the coefficient of determination for the density of this system was obtained in the test category 0.9854.

Figure 4 shows the correlation between the experimental data and the predicted viscosities values of the training category of water, ethanol and ethylene glycol by the proposed model. The determination coefficient for the training category of this system is obtained in the 1.0000 instruction class, and as shown in Figure 5, the coefficient of determination for the viscosity of this system in the test category was obtained 0.9892. As shown in Figure 2-5, and according to the coefficients of determination and mean square error obtained in this paper, the results show the success of the proposed method in simultaneously predicting the density and viscosity of the ternary system of water, ethanol and ethylene glycol.

Table 2. The data obtained by the rectifying machine vector in the training and test category for the density of the triple system of water and ethanol and ethylene glycol

Steps to run the software	gamma	Sigma2	Train		test	
			R^2	MSE	R^2	MSE
1	247.6291	0.2101193	1.0000	5.1328×10^{-7}	0.8972	0.0026
2	81.8542	0.1505280	1.0000	2.7551×10^{-6}	0.8120	0.0043
3	200.3167	0.2001038	1.0000	6.9913×10^{-7}	0.8858	0.0029
4	329.5379	0.1716492	1.0000	2.3653×10^{-7}	0.8477	0.0036
5	108.0751	0.1671832	1.0000	1.7645×10^{-6}	0.8399	0.00038
6	206.4259	0.1982543	1.0000	5.5601×10^{-7}	0.8836	0.0029
7	277.3569	0.2117261	1.0000	4.2543×10^{-7}	0.8990	0.0026
8	438.0806	0.3766458	1.0000	5.2785×10^{-7}	0.9819	5.7082×10^{-3}
9	452.842	0.3980327	1.0000	5.5770×10^{-7}	0.9854	4.6572×10^{-4}
10	225.5518	0.2065906	1.0000	5.9090×10^{-7}	0.8933	0.0027
11	177.3426	0.1934183	1.0000	8.3294×10^{-6}	0.8775	0.0030
12	300.861	0.2093218	1.0000	3.6289×10^{-7}	0.8964	0.0026
13	410.7357	0.3413949	1.0000	4.8611×10^{-7}	0.9740	7.9614×10^{-4}
14	361.4749	0.2482923	1.0000	3.4360×10^{-7}	0.9314	0.0019
15	388.7204	0.2525492	1.0000	3.1245×10^{-7}	0.9344	0.0018

Table 3. The data obtained by the rectifying machine vector in the training and test category for the viscosity of the triple system of water and ethanol and ethylene glycol

| Steps to run the software	gamma	Sigma2	Train		test	
			R^2	MSE	R^2	MSE
1	200.3167	0.2001038	1.0000	6.9913×10^{-7}	0.8858	0.0029
2	231.7204	0.2025492	1.0000	2.7551×10^{-6}	0.8887	0.0028
3	302.4856	0.2632014	1.0000	5.0897×10^{-7}	0.9414	0.0016
4	279.5518	0.2265906	1.0000	4.0555×10^{-7}	0.8934	0.0027
5	488.1297	0.4280164	1.0000	5.7609×10^{-7}	0.9892	3.4920×10^{-4}
6	445.4749	0.4133784	1.0000	6.1395×10^{-7}	0.9875	4.0275×10^{-4}
7	227.7399	0.1918428	1.0000	5.3016×10^{-7}	0.8756	0.0031
8	390.1423	0.2462117	1.0000	2.9747×10^{-7}	0.9299	0.0019
9	457.3605	0.3797392	1.0000	5.0260×10^{-7}	0.9824	5.5380×10^{-4}
10	376.3066	0.2994503	1.0000	4.4222×10^{-7}	0.9599	0.0012
11	420.8256	0.3692034	1.0000	5.4027×10^{-6}	0.9805	6.1293×10^{-4}
12	248.242	0.208641	1.0000	5.0629×10^{-7}	0.8956	0.0027
13	296.7531	0.2439018	1.0000	4.6624×10^{-7}	0.9281	0.0019
14	436.7166	0.3976163	1.0000	5.8695×10^{-7}	0.9853	4.6806×10^{-4}
15	355.0319	0.2709318	1.0000	4.0961×10^{-7}	0.9460	0.0015
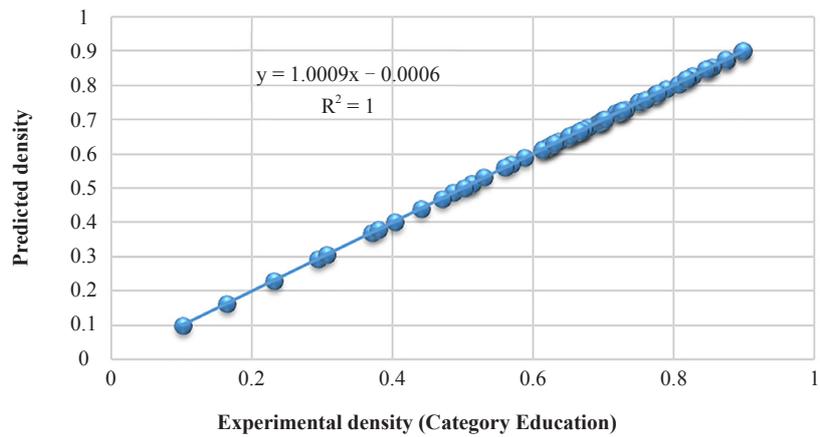
Figure 2. The relationship between experimental density and density predicted by the proposed model in the training category
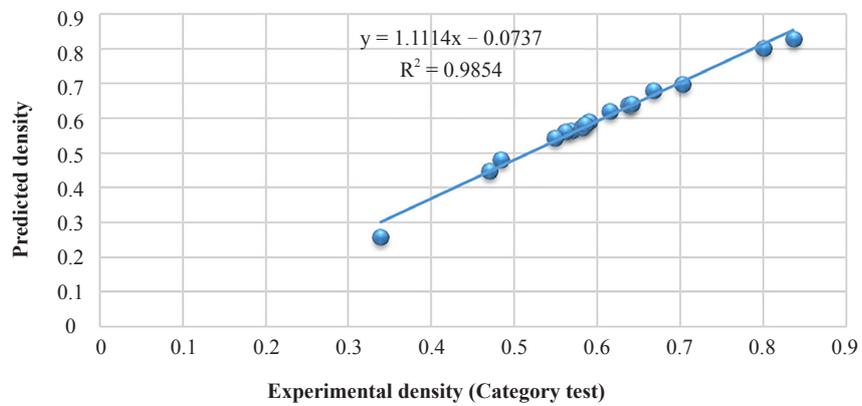

Figure 3. The relationship between experimental density and predicted density by the proposed model in the test category
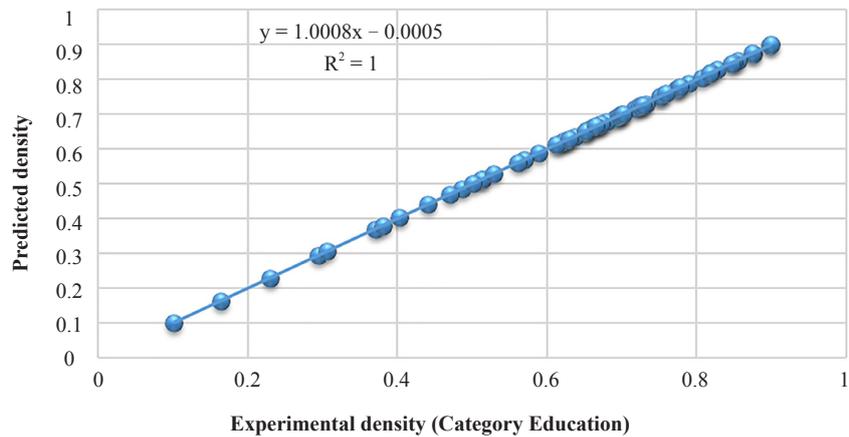

Figure 4. The relationship between experimental viscosity and viscosity predicted by the proposed model in the training category
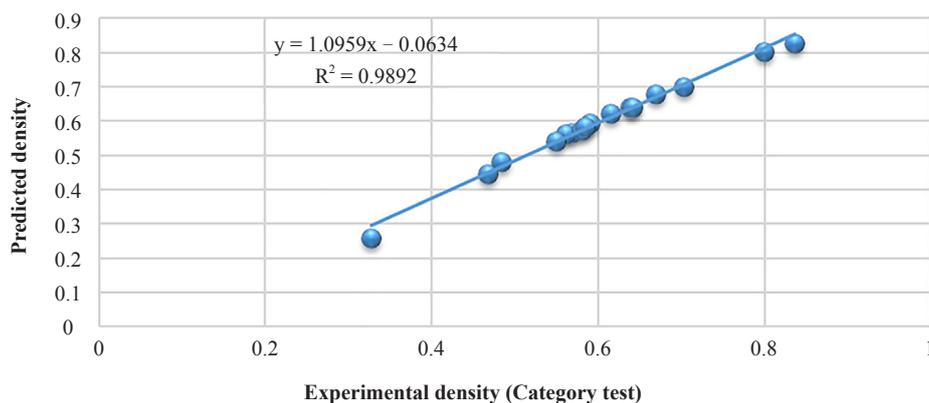
$$y = 1.0959x - 0.0634$$
$$R^2 = 0.9892$$

Figure 5. The relationship between experimental viscosity and predicted viscosity by the proposed model in the test category

4. Conclusions

In this paper, it was possible to simultaneously predict the density and viscosity of the triple system of water and ethanol and ethylene glycol using Support vector machines. Retest requires a lot of time and money, and the number of parameters is high, and the temperature is considered for each of the parameters separately. This method, which predicts all conditions simultaneously, is of importance. In other words, this method allows researchers in the future to achieve optimal results without re-testing and a waste of time by referring to modeling. The supported vector machine method confirmed the simultaneity prediction of density and viscosity of water, ethanol, and ethylene glycol as a function of temperature. The lower errors and higher accuracy compared to statistical method along with ease of using of supported vector machine makes it a practical method for predicting the viscosity and density of the mixtures.

5. Acknowledgment

Conflict of interest

The author declares no competing financial interest.

References

[1] Abbaspour. A., Baramakeh. L. Simultaneous determination of antimony and bismuth by beta-correction spectrophotometry and an artificial neural network algorithm. *Talanta*. 2005; 65: 692-699. Available from: https://www.ncbi.nlm.nih.gov/pubmed/18969854.

[2] Agrawal. V. K., Louis. B., Khadikar. P. V. Prediction of intrinsic solubility of generic drugs using MLR, ANN and SVM analyses. *European Journal of Medicinal Chemistry*. 2010; 45: 4018-4025. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20584562.

[3] Akbari. E., Buntat. Z., Enzevaee. A., et al. Analytical modeling and simulation of I-V characteristics in carbon nanotube based gas sensors using ANN and SVR methods. *Chemometrics and Intelligent Laboratory*. 2014; 137: 173-180. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0169743914001488.

[4] Balabin. R. M., Lomakina. E. I., Safieva. R. Z. Neural network (ANN) approach to biodiesel analysis: Analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy. *Fuel*. 2011; 90: 2007-2015. Available from: https://www.sciencedirect.com/science/article/pii/S0016236110006538.

[5] Beebe, K. R., Kowalski, B. R. Binding of tissue plasminogen activator to human umbilical vein endothelial cells. *Anal. Chem*. 1987; 705: 241-254. Available from: https://www.ncbi.nlm.nih.gov/pubmed/3111002.

[6] Bordbar. M. M., Khajehsharifi. H., Solhjoo. A. PC-ANN assisted to the determination of Vanadium (IV) ion using an optical sensor based on immobilization of Eriochorome Cyanine R on a triacetylcellulose film. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2015; 151: 225-231. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26142655.

[7] Keskin, S., Kayrak, D., Akman, U., et al. "A review of ionic liquids towards supercritical fluid applications". *The*

Journal of Supercritical Fluids. 2007; 43: 150-180. Available from: https://www.sciencedirect.com/science/article/pii/S0896844607002033.

[8] Khajehsharifi. H., Bordbar. M. M. A highly selective chemosensor for detection and determination ofcyanide by using an indicator displacement assay and PC-ANN and itslogic gate behavior. *Sensors and Actuators B*. 2015; 209: 1015-1022. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0925400514012647.

[9] Khanchi. A. R., Khayatzadeh Mahani. M., Hajihosseini. M., et al. Simultaneous spectrophotometric determination of caffeine and theobromine in Iranian tea by artificial neural networks and its comparison with PLS. *Food Chemistry*. 2007; 103: 1062-1068. Available from: https://www.sciencedirect.com/science/article/pii/S0308814606006625.

[10] Liu, Q. S., Li, P. P., Welz-Biermann, U., et al. "Density, dynamic viscosity, and electrical conductivity of pyridinium-based hydrophobic ionic liquids". *J. Chem. Thermodynamics*. 2013; 66: 88-94. Available from: https://www.sciencedirect.com/science/article/pii/S0021961413002279.

[11] Ni. Y., Xia. Zh., Kokot. S. A kinetic spectrophotometric method for simultaneous determination of phenol and its three derivatives with the aid of artificial neural network. *Journal of Hazardous Materials*. 2011; 192: 722-729. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21719195.

[12] Raimundo Jr. I., Narayanaswamy. R. Simultaneous determination of relative humidity and ammonia in air employing an optical fibre sensor and artificial neural network. *Sensors and Actuators B*. 2001; 74: 60-68. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0925400500007127.

[13] Rezaei. B., Ensafi. A. A., Shandizi. F. Simultaneous determination of cobalt and nickel by spectrophotometric method and artificial neural network. *Microchemical Journal*. 2001; 70: 35-40. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0026265X01000947.

[14] Safavi. A., Abdollahi. H., Hormozi Nezhad. M. R. Artificial neural networks for simultaneous spectrophotometric differential kinetic determination of Co(II) and V(IV). *Talanta*. 2003; 45: 515-523. Available from: https://www.ncbi.nlm.nih.gov/pubmed/18968936.

[15] Smiglak, M., Hough, W. L., Rodriquez, M. D., et al. "The third evolution of ionic liquids: active pharmaceutical ingredients". *New Journal of Chemistry*. 2007; 31: 1429-1436. Available from: https://pubs.rsc.org/en/content/articlelanding/2007/nj/b706677p#!divAbstract.

[16] Tang. L., Zeng. G. M., Shen. G. L., et al. Simultaneous amperometric determination of lignin peroxidase and manganese peroxidase activities in compost bioremediation using artificial neural networks. *Anal. Chem. Acta*. 2006; 579: 109-116. Available from: https://www.ncbi.nlm.nih.gov/pubmed/17723735.

[17] Tenorio-Borroto. E., Peñuelas Rivas. C. G., Vásquez Chagoyán. J. C., et al. ANN multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the anti-microbial drug G1 in spleen. *Bioorganic & Medicinal Chemistry*. 2012; 20: 6181-6194. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22981917.

[18] Zamaniyan. A., Joda. F., Behroozsarand. A., et al. Application of artificial neural networks (ANN) for modeling of industrial hydrogen plant. *International Journal of Hydrogen Energy*. 2013; 38: 6289-6297. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0360319913005703.

[19] Zhao, H., Xia, S., Ma, P. "Use of ionic liquids as 'green' solvents for extractions". *Journal of Chemical Technology and Biotechnology*. 2005; 80: 1089-1096. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jctb.1333.

[20] Kianfar E, Shirshahi M, Kianfar F. "Simultaneous prediction of the density, viscosity and electrical conductivity of pyridinium-based hydrophobic ionic liquids using artificial neural network." *Silicon*. 2018; 10: 2617-2625. Available from: https://link.springer.com/article/10.1007/s12633-018-9798-z.

[21] Quijada-Maldonado. E., Meindersma. G. W., Haan. A. B. " Viscosity and density data for the ternary system water(1)-ethanol(2)-ethylene glycol(3) between 298.15 K and 328.15 K." *J. Chem. Thermodynamics*. 2013; 57: 500-505. Available from: https://www.sciencedirect.com/science/article/pii/S0021961412003382.