A Review of Artificial Neural Networks for Chemical Process Optimization and Compound Property Prediction

Abstract Artificial neural networks are widely used in chemical processes because of their powerful data processing capabilities, fault tolerance, nonlinear relationship processing capabilities, and learning capabilities. This paper will introduce the development history and important models of artificial neural network, and focus on its application in chemical process optimization and prediction of physical properties of compounds.

Keywords: artificial neural network; BP neural network ; CNN; RBF; process optimization; property prediction.

1 Introduction

Artificial neural networks, inspired and developed from the principles of biology (Figure 1), are nonlinear, adaptive information processing systems composed of a large number of interconnected processing units, which mimic the abstraction of the neuronal network of the human brain, and then create some kind of mathematical model to process information by adjusting the interconnections between a large number of nodes within the model [1].

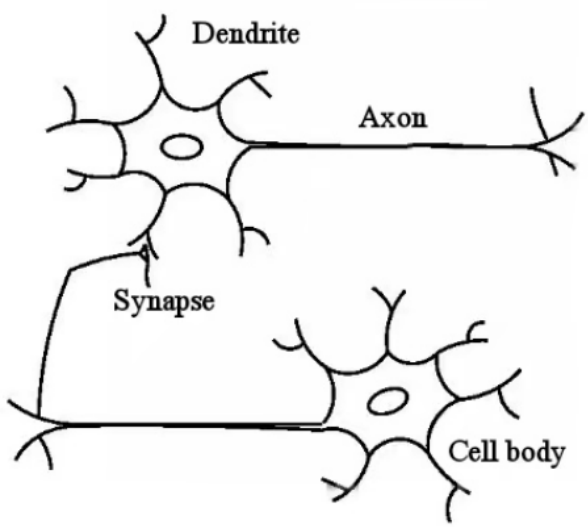

Figure 0 **Inter-neuronal signaling**

Artificial neural network has self-learning function and fault-tolerant characteristics, and can approach any nonlinear mapping relationship. The corresponding relationship between the physical property data of chemical products, such as boiling point, spontaneous combustion point and saturated vapor pressure, and the material structure is nonlinear, as is the case between the objective function of chemical processes and the influencing factors. Neural networks can be used to optimize or predict. In this paper, the development history and important models of artificial neural network will be introduced and its applications in physical property prediction and chemical process optimization will be summarized.

2 Artificial neural network

In 1943, physiologist McCulloch and mathematician Pitts [2] published an article, proposing the first neuron model (M-P model). Their efforts laid the foundation for

the development of network models and later neural networks, and began the research on artificial neural networks. Hebb [3] proposed the Hebb rule, which only changes the weight according to the activation level of the neuronal connections. Therefore, this method is also called correlation learning or parallel learning. In 1954, biologist Eccles [4] proposed a real synaptic shunt model, which was confirmed by electrophysiological experiments on synapses. In 1960, Widrow and Hoff [5] proposed the Adaptive linear component Adaline network model, which is a continuous-valued adaptive linear neuron network model, they are useful for training segmented linear networks, this method is faster and has high accuracy. Kohonen [6] proposed the self-organizing mapping (SOM) theory. Anderson [7] proposed a similar neural network called "interactive memory". In 1980, Kunihiko Fukushima[8] published "Neocognitron", a model of visual pattern recognition mechanism, which was combined with biological vision theory to synthesize a neural network model with certain pattern recognition ability like human. Hopfield [9] proposed the Hopfield model theory, and proved that the network could reach a stable state under certain conditions. In 1986, the idea of Parallel Distributed Processing (PDP) network was proposed by Rumelhart and McCkekkand [10]. After years of development, hundreds of neural network models have been proposed and applied [11,12], among which, BP neural network, RBF neural network and CNN neural network are more commonly used [13].

BP neural network is a multilayer feed-forward neural network. The main features of this network are signal forward transmission and error back propagation. In forward transmission, the input signal is processed layer by layer from the input layer through the implicit layer to the output layer. The state of neurons in each layer only affects the state of neurons in the lower layer. If the desired output is not obtained at the output layer, it is transferred to back propagation, and the network weights and thresholds are adjusted according to the prediction error, so that the predicted output of the BP neural network continuously approximates the desired output. The bracket structure of the BP neural network is shown in Figure 2 [14] where X and Y represent the input and predicted value of the BP neural network

respectively, and the ω represents the neural network weights.

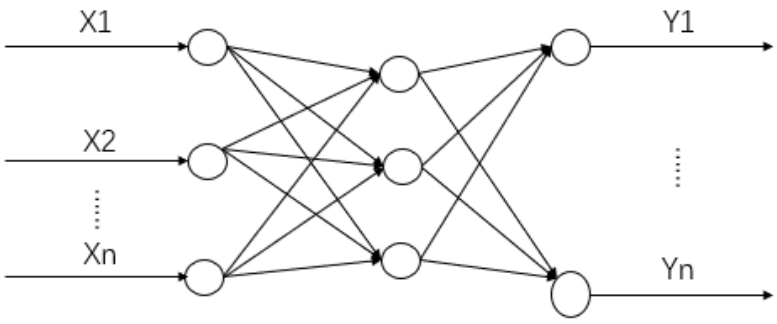

Figure 1 **Bracket structure of BP neural network**

In 1988, Broomhead and Lowe [15] first applied the radial basis function (RBF) to the neural network design to form the RBF neural network. The hidden node of RBF neural network uses the distance between the input mode and the center vector as the independent variable of the function, and uses the radial basis function (such as Gaussian function) as the activation function. It has a "local mapping" property and can approximate any continuous function with any accuracy, which is particularly suitable for solving classification problems.

Convolutional Neural Network (CNN) is a kind of feedforward neural network, which includes convolutional layer and pooling layer, and has excellent performance for large image processing [16]. In 1998, Yann LeCun et al. [17] built a more complete convolutional neural network LeNet-5 and succeeded in the recognition of handwritten digits.

RBF neural networks and BP neural networks are both nonlinear multilayer forward networks. However, BP approximates the minimum error by continuously adjusting the weights of neurons, which is generally a gradient descent method, while RBF is a feed-forward neural network, that is, it does not constantly adjust the weights to approximate the minimum error, and its excitation function is generally a Gaussian function, unlike the S-shaped function of BP, which calculates the weights by the distance between the input and the center of the function. The RBF neural network is not only theoretically the best network among the forward networks

compared to BP networks, but also the learning method avoids the problem of local optimality.CNN is mainly used for image data processing.

So far, artificial neural network has been widely used in various fields, mainly including intelligent driving[18,19], auxiliary medical treatment[20,21], voice information processing[22,23], Chinese medicine processing[24,25], chemical process optimization and process control[26-28].

3 Application of Artificial neural network in chemical industry

Due to the superior self-learning ability of neural networks and the powerful information processing capability, artificial neural networks are also widely used in the chemical industry. In this chapter, some applications of artificial neural networks in chemical industry will be described specifically.

3.1 Prediction of physical properties of compounds

Computer-aided molecular design is a method and technique for designing biologically active compounds prior to compound synthesis based on the laws obtained from the study of the structure-bioactivity relationship of molecules. Traditional research methods attribute the structure-activity relationship to a linear relationship and then perform a linear analysis, and it is clear that this linearization of eight has obvious shortcomings. In contrast, artificial neural network methods are characterized by the ability to model any nonlinear problem. This avoids the shortcomings associated with linearizing the force problem. Thus the relationship between chemical structure and activity can be simulated by learning and understanding a series of chemical structures and their activities through artificial neural networks, forming a structure-activity relationship for the prediction of the activity of a new compound. [29]

The regular relationship between the structure of compounds and their physical and chemical properties is nonlinear, which can be predicted by artificial neural networks. The relationship between molecular structure parameters and properties is modeled by correlating the intrinsic quantitative relationships between molecular structure parameters and experimental data on the properties under study. Once a

reliable quantitative structure-property correlation model is established, only the structural information of the molecule is needed to use it to predict various properties of compounds.

Tetteh et al. [30] applied RBF (radial basis function) artificial neural network technology to predict the flash point and boiling point of 400 kinds of organic substances at the same time.The molecular connectivity index and 25 functional groups were used as molecular structure descriptors to predict 133 substances in the test set, and the average absolute error was 11.9 ℃.

Pan *et al.* [31] applied the principle of group contribution method to extract 32 possible molecular groups in carbon, hydrogen and oxygen compounds as structure descriptors, and combined BP artificial neural network method to successfully predict the flash points of 258 such compounds. Subsequently, Pan et al. [32] proposed to use group bonds as molecular descriptors to characterize molecular structure characteristics, and studied the correlation between the flash points of 44 alkanes and their molecular structures by BP artificial neural network method.

Albahri and George [33] used the artificial neural network method to predict the spontaneous ignition point of 490 organic substances.

Deng Tong [34] predicted the solubility of compounds in water based on the artificial neural network model. On this basis, the experimental data obtained by Baghban et al. [35] were used to establish a data set of CO_2 solubility in ionic liquids, and a prediction model of CO_2 solubility combining CNN and the critical characteristics of ionic liquids was established. The fitting coefficients R2 was 0.999.

In 2022, Xu et al. [36] predicted the bioactivity values of 50 new compounds based on PSO optimized BP neural network with a model fit of 0.834 and a root mean square error of 0.732, which was more relevant than the predicted values of the BP neural network before optimization. Subsequently, to improve the success rate of drug development, an ADMET classification prediction model was constructed based on the existing ADMET property data using PSO-optimized SVM, and the algorithm cross-validated CV accuracy reached 94 %, and the prediction accuracy of all five index models was above 79%. The results showed that the developed model had

better prediction performance than the baseline model, and the adopted prediction strategy was effective, which could provide a reference for the development of anti-breast cancer drugs.

Hou [37] designed various neural networks including single-layer neural networks, multilayer neural networks and convolutional neural networks to investigate the prediction performance of 15 properties of 138,850 organic small molecules in the QM9 database. In this paper, they achieved the efficient prediction of multiple properties of molecules in a large volume database based on machine learning, completed the establishment and large-scale expansion of a small database, and performed high-throughput screening of hydrocarbons and nitrogen-containing compounds to provide theoretical guidance for the development of new high-density hydrocarbon fuels and high-energy nitrogen-containing explosives. The application of the new machine learning method to the molecular design, property prediction and high-throughput screening of energetic compounds is realized in this paper. The proof-of-concept results show that the strategy will accelerate the design and development of novel energetic materials.

Li et al. [38] used the substituted substrate segment value P and the atomic type electrical topological state index Em as the input layer variables and the thermodynamic properties as the output layer variables of the neural network, and the equations were all in a 5:13:1 network structure, and three satisfactory QSPR models were obtained using the BP algorithm, and their total correlation coefficients were 0.9986, 0.9911, and 0.9795, respectively, The standard errors were 2.123, 3.237 and 3.952, respectively. The relative average errors of the predicted values calculated with these three neural network models and the literature values were 0.30%, 1.85% and 1.14%, respectively, which indicated that the models had good stability and predictive ability. The results obtained by this neural network model are better than those obtained by multiple regression methods and can be used for theoretical analysis and prediction of the properties of polychlorocarbazole compounds.

3.2 Optimization of Chemical Process

The synthesis process of a chemical product includes many factors such as

temperature, material ratio and so on. In addition to product yield, energy consumption should also be considered as indicators to characterize the process. Therefore, the relationship between the independent variables and the objective function related to process optimization is extremely complex, which is exactly the area where the artificial neural network can play a role [39-42].

In 2008, Zhao et al. [43] established a product quality forecasting and production process optimization model for iron oxide red using a combination of genetic algorithm and artificial neural network technology (GA-ANN). The network model structures of the iron oxide red product quality forecasting system and the production process optimization system are the same as shown in Figure 3. Then the process of using the combination of ANN and GA for the optimization of iron oxide red production to find the extreme value is shown in Figure 4. The online application shows that the model has high learning accuracy and fast convergence speed, and the hit rate of forecasting the mass fraction of ferric oxide of iron oxide red with an allowable error of less than 3% is over 90%. The use of GA solves the problem that traditional optimization methods are prone to fall into local minima for complex optimization problems. The combination of GA and ANN, which utilizes both the inference and prediction functions of ANN's nonlinear mapping network and the global optimization characteristics of GA, complements each other.

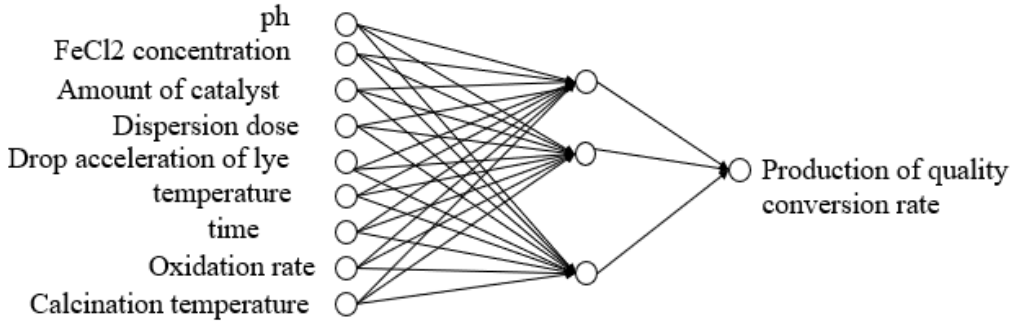

Figure 3 **Iron oxide red production end product quality or conversion rate neural network**

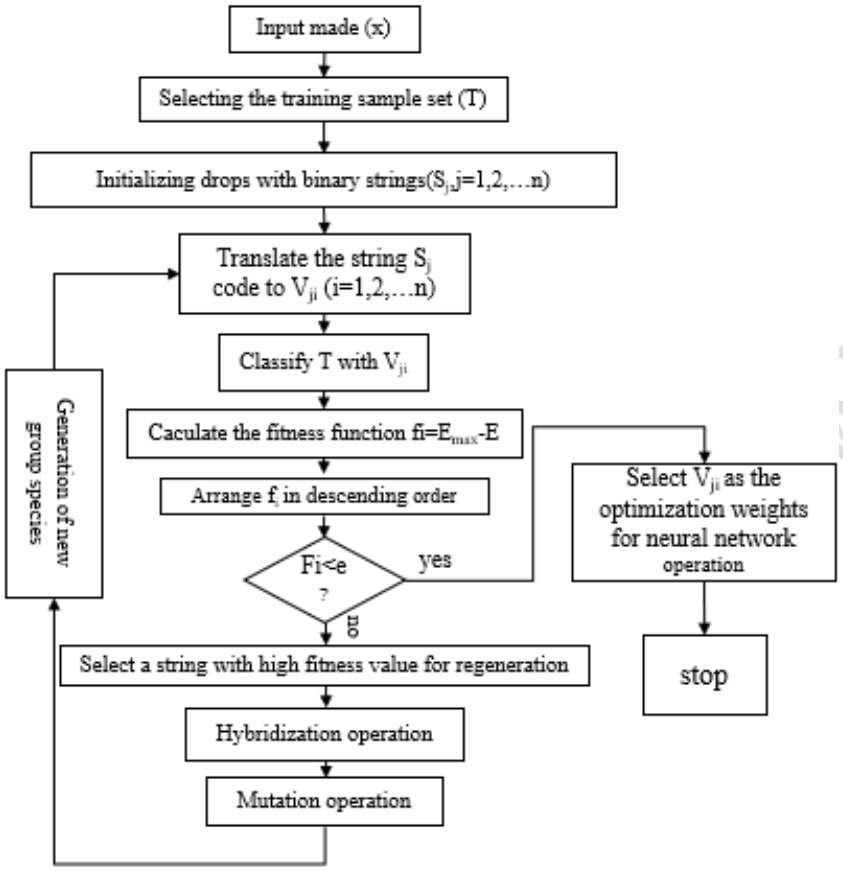

Figure 4 **Computational flowchart of ANN training with GA**

Lu *et al.* [44] explored the optimization problem of combining neural network and genetic algorithm to solve the parameters of gold production process collaboratively. With 64 groups of experimental data as training samples, the BP neural network model established includes 8 input nodes (respectively, sodium sulfate, sodium sulfite, sulfur agent, copper sulfate, oxygen, ammonia concentration, pH value, grinding time), 15 hidden layer nodes and 1 output node (cost). Take the learning parameter c=0.01, input the training samples into the network, and after 200000 steps of training, its mean square error is less than 0.0001 After training, the neural network can correctly identify these training samples, and the fitting rate is 100%.

Hu et al. [45] used artificial neural networks to optimize the production process of new solid fuels. The main factors affecting the mechanical strength and storage performance of the new fuel as input variables, built a quality prediction model by ANN, and inputted a large number of randomly generated process conditions into the

trained ANN, and obtained an optimized stable operating region by counting the input variables corresponding to the output results meeting the quality requirements, and this optimization method effectively improved the stability of product quality. The three-layer feedforward network with a hidden layer is adopted, and the network structure is 6-6-1. The activation function is sigmoid function, and the network is trained with the improved BP algorithm [46]. Through experiments, when the step size is 0.8 and the momentum factor is 0.6, local minima can be avoided. When the training times are 30000, the network converges to $E=4.41 \times 10^{-3}$.

Zheng et al. [47] used response surface methodology and artificial neural network to optimize the processing technology of honey baked Chuanxiong. The content of ferulic acid and ligustrolactone of honey-roasted Chuanxiong was set as evaluation indexes, and Box-Behnken response surface design experiments were carried out to investigate the effects of honey addition, concoction temperature, concoction time and smothering and wetting time on the honey-roasting process. Taking the test data of Box Behnken method as the training sample, using Matlab software programming and Levenberg Marquardt algorithm, the maximum iteration number is 10000, a BP neural network with hidden layer 1 is established, and the network is trained. After training, the network calculation accuracy reaches 0.0001

Wu[48] also used response surface methodology and artificial neural network model to analyze and optimize the reaction conditions for the preparation of epoxy fatty acid isobutyl ester by formic acid autocatalysis using fatty acid isobutyl ester as feedstock.

Wang *et al.* [49] predicted the effect of alloying process parameters on each property index of alloy layers by building a neural network. A 4-layer network structure is adopted to train the network with 20 groups of experimental data as samples. Through self-learning the obtained test data and performance indicators, a model of Fe Si alloying process parameters on relevant performance indicators is established. The predicted saturation magnetization and corrosion current density are close to the actual test results, the network learning accuracy is high, and the relative error between the predicted value and the test value is less than 1.2%. The prediction

result given by the neural network is that the mole fraction of Si is 0.2-0.35 and the temperature is 1000 □.

4 Conclusion

In summary, neural networks have powerful learning ability and nonlinear processing capability in physical property prediction and chemical process optimization. Compared with other methods, the use of neural networks has the advantages of processing a large number of data sets at one time, rapid response, and accurate prediction. However, neural networks still need to be improved in the acquisition and processing of training sets and the evaluation of the reliability of prediction results, and more targeted neural network models for chemical processes are needed. In terms of developing new neural network theory, in addition to expecting breakthroughs in neurophysiology and other research, combining neural networks with other theories and developing new neural network models is also one of the development directions of neural network research. For example, the chaotic neural network theory. Another example is the combination of quantum mechanics and neural networks to form quantum neural networks.

References

[1] Song H , Sun X , Xiang S . Advances in the application of artificial neural networks in chemical processes [J]. Advances in Chemical Engineering, 2016, 35(12): 3755-3762

[2] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. 1943.[J]. Bulletin of Mathematical Biology. 1943;52(1–2):99.

[3] Hebb DO. The organization of behavior: A neuropsychological theory.[M]. New York: John Wiley and Sons, Inc.; 1949.

[4] Eccles J C, Fatt P, Koketsu K. Cholinergic and inhibitory synapses in a pathway from motoraxon collaterals to motoneurones [J]. The Journal of Physiology, 1954.

[5] Widrow B. Adaptive ` Adaline' Neuron Using Chemical ` Memistors, Stanford Electronics Laboratories Technical Report, No. 1553-2 , 1960.

[6] Kohonen T. Fast evolutionary learning with batch-type self-organizing maps [J].Neural Processing Letters，1999, 9(2) : 153-162.

[7] Anderson JR, Fincham JM, Qin Y, Stocco A. A central circuit of the mind. Trends Cogn. Sci. 2008, 12:136–143.

[8] Fukushima K. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4): 193-202.

[9] Hopfield JJ. Neurons with graded response have collective computational properties like those of two-state neurons. Proc Natl Acad Sci, 1984, 81(10): 3088–3092.

[10] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagation errors, Naturem, 1986, 323: 53–536.

[11] Sun L, Liang F, Cui W. Artificial Neural Network and Its Application Research Progress in Chemical Process, Asian Journal of Research in Computer Science, 2021, 12(4): 177-185

[12] Wang HN, Chen YJ. Application of Artificial Neural Networks in Chemical Process Control. Asian Journal of Research in Computer Science, 2022, 14(1): 22-37.

[13] Wang HH, Mo RY. Review of Neural Network Algorithm and Its Application in Reactive Distillation. Asian Journal of Chemical Sciences, 2021, 9(3): 20-29.

[14] Wang X , Shi F , Yu L , Li Y . 43 case studies of MATLAB neural networks [M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 2013.

[15] Broomhead DS, Lowe D. Multivariate Functional Interpolation and Adaptive Network. Complex Systems 2, 1988, 321-355.

[16] Shi F, Gao J, Huang X. An affine invariant approach for dense wide baseline image matching. International Journal of Distributed Sensor Networks (IJDSN), 2016, 12(12): 1-12.

[17] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[18] Gao J, Tembine H. Distributed Mean-Field-Type Filter for Vehicle Tracking, in

American Control Conference (ACC), Seattle, USA, May 2017.

[19] Kalal Z, Mikolajczyk K, Matas J. “Tracking-learning-detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):1409–1422.

[20] Kamnitsas K，Ledig C，Newcombe V，et al． Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation［J］． Med Image Anal，2017，36: 61-78．

[21] Jamaludin A，Kadir T，Zisserman A． SpineNet: Automated classification and evidence visualization in spinal MRIs． Med Image Anal，2017，41: 63-73．

[22] Pelenis D, Barauskas D, Vanagas G, et al. CMUT-based biosensor with convolutional neural network signal processing[J]. Ultrasonics, 2019, 99:105956.

[23] Gao J, Chakraborty D, Tembine H and Olaleye O. Nonparallel Emotional Speech Conversion. INTERSPEECH 2019, Graz, Austria, September 2019.

[24] Luo N, Cheng L, Fu C, et al. Optimization of Processing Technology for Curcuma phaeocaulis by Orthogonal Test and Artificial Neural Network Model, Chinese Journal of Experimental Traditional Medical Formulae, 2014, 20(24): 10-13.

[25] Sun L, Ren X, Zhang H, et al. Research on processing of Radix Polygoni Multiflori based on UPLC fingerprints and chemometrics, China Journal of Traditional Chinese Medicine and Pharmacy, 2017,32(05): 2194-2197.

[26] Zheng ZL, Qi Y. Study on the Simulation Control of Neural Network Algorithm in Thermally Coupled Distillation. Asian Journal of Research in Computer Science, 2021, 10(3): 53-64.

[27] Feng XD, Sun L. Application Progress of Artificial Neural Network in Chemical Industry, Journal of Engineering Research and Reports, 2022, 23(7): 26-36.

[28] Zhao ZH, Lu X. Research Progress of Chemical Process Control and Optimization Based on Neural Network, Journal of Engineering Research and Reports, 2021, 21(12):10-17.

[29] Dian Jin, Xu JM. Artificial neural network and its Applications in Chemical Industry, Asian Journal of Chemical Sciences, 2022, 12(3): 31-39.

[30] Tetteh J, Suzuki T, Metcalfe E, Howells S. Quantitative structure-property relationships for the estimation of boiling point and flash point using a radial basis function neural network, Journal of Chemical Information and Computer Sciences,1999, 39(3): 491-507.

[31] Pan Y，Jiang J C. Quantitative structure–property relationship study for estimating flash points of organic compounds using back-propagation artificial neural networks[C]//Progress in Safety Science and Technology：Vol. VI. Beijing：Science Press，2006： 1462-1465.

[32] Pan Y, Jiang J C, Wang Z R. Prediction of flash point of alkanes by group bond contribution method based on artificial neural network [J]. Chemical Engineering，2007, 35（4）：38-41.

[33] Albahri T A，George R S. Artificial neural network investigation of the structural group contribution method for predicting pure components auto ignition temperature [J]. Ind. Eng. Chem. Res.，2003，42 (22)：5708-5714.

[34] Deng T. Deep learning-based study on prediction of water solubility and CO_2 solubility of compounds [D]. Sichuan Normal University, 2020.

[35] Baghban A, Ahmadi M A, Shahraki B H. Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches[J]. The Journal of supercritical fluids, 2015, 98: 50-64.

[36] Xu M, Zheng Y, Li Y, Wu W. Anti-breast cancer drug property prediction based on PSO-BP neural network and PSO-SVM [J/OL]. Journal of Nanjing University of Information Engineering (Natural Science Edition): 2022 (10):1-20.

[37] Hou F. Molecular design and property prediction of high-energy compounds based on machine learning [D]. Tianjin University, 2020.

[38] Li J, Plug XH. Neural network prediction of thermodynamic properties of polychlorocarbazole compounds[J]. Journal of Wuhan University (Science Edition), 2014, 60(04): 324-330.

[39] Mu Y, Sun L. Catalyst Optimization Design Based on Artificial Neural Network. Asian Journal of Research in Computer Science, 2022, 13(2): 1-12.

[40] Sun J, Chen Q. Optimization of Extraction Process Based on Neural Network,

Asian Journal of Chemical Sciences, 2022, 11(2): 19-27.

[41] Liang F, Zhang T. Research on Chemical Process Optimization Based on Artificial Neural Network Algorithm, Asian Journal of Research in Computer Science, 2021, 12(4): 12-24.

[42] Li C, Wang C. Application of artificial neural network in distillation system: A critical review of recent progress. Asian Journal of Research in Computer Science, 2021, 11(1): 8-16.

[43] Zhao QM. Optimization model of iron oxide red production process based on neural network and genetic algorithm [J]. Inorganic Salt Industry, 2008(08):56-58.

[44] Lu C, Yun Q, Chen Y , Lan X . Optimization of gold production process parameters [J]. China Mining Industry, 2003(10): 52-54.

[45] Hu L, Sha G, Hu G, Wang W, Xia J. A neural network optimization method for new solid fuel production process [J]. Journal of East China Jiaotong University, 1999(04): 6-9.

[46] Huang D S. Theory of neural network pattern recognition system [M]. Beijing: Electronic Industry Press, 1996: 46-48.

[47] Zheng KX, Zhao YF, Li WB, Li Y, Zhang TT, Chen D, Hu CJ. Box-Behnken response surface method combined with artificial neural network for optimal honey-baked Chuanxiong Ligusticum concoction process [J]. Chinese medicinal materials, 2017, 40(09):2055-2059.

[48] Wu, F. Study on the preparation process of epoxy fatty acid isobutyl ester and its plasticizing properties [D]. Zhejiang University of Technology, 2021.

[49] Wang Y, Han J, Sun Zonghui, et al. Research on prediction of Fe-Si alloy layer properties and process optimization based on artificial neural network [J]. Shandong Metallurgy, 2022, 44(2): 17-20.