Computing and Informatics, Vol. 43, 2024, 611–632, doi: 10.31577/cai_2024_3_611

CORPORATE FRAUD DETECTION BASED ON IMPROVED BP NEURAL NETWORK

Wei Liu, MingMing Liu

College of Computer Science and Engineering
Shandong University of Science and Technology
Qingdao 266590, China
e-mail: liuwei_doctor@yeah.net, lmm@sdust.edu.cn

Chun Yan*

College of Mathematics and System Science
Shandong University of Science and Technology
Qingdao 266590, China
e-mail: yanchunchun9896@sina.com

Man Qi

School of Engineering, Technology and Design
Canterbury Christ Church University
Canterbury CT1 1QU, UK
e-mail: man.qi@canterbury.ac.uk

LuLu Zhang

College of Computer Science and Engineering
Shandong University of Science and Technology
Qingdao 266590, China
e-mail: 2508020352@qq.com

* Corresponding author

Abstract. Corporate fraud risk detection is a branch of fraud. It may exist in various industries and cause economic problems. Effective identification of corporate fraud can protect the safety of funds for investors in some sense. This paper proposes a classifier model of a fractional-order immune BP neural network based on the self-attention mechanism to improve efficiency. The improved artificial immune algorithm with dynamic region contraction strategy is used to optimize the initialization process of the BP neural network. Furthermore, it combines the self-attention mechanism to design the input layer. Finally, Caputo fractional non-causal calculus is used to optimize the parameter updating process in BP neural network. The experiment results indicate that our model has fast convergence rate and powerful capacity of detection, and performs efficiently in detecting fraud behaviors.

Keywords: Intelligent optimization algorithm, self-attention mechanism, BP neural network, fractional-order, fraud detection

1 INTRODUCTION

People have always been concerned about and sought to avoid corporate fraud problems. Because it can cause serious consequences and malicious influence to individuals or society. Although people have the awareness of fraud prevention, all kinds of fraud methods still emerge in an endless stream. Fraud involves all aspects of people's lives, such as consumer fraud, insurance fraud and financial fraud.

With the development of artificial intelligence, many researchers have tried to use methods of machine learning to achieve accurately fraud detection [1, 2]. Probabilistic graphical model and long short-term memory network are used to detect fraudulent behaviors of credit cards [3]. Boosted decision trees and feedforward neural network were used to detect abnormal behavior in network performance data [4]. Alharbi et al. proposed the new method to convert text data into image format. Then, the pipeline features were input into the convolutional neural network to realize credit card fraud detection [5]. BP neural network has good ability of self-learning, self-adaptation and generalization [6]. Therefore, it is widely used and has become an effective tool for fraud detection [7].

Yang used BP neural network combined with logistic regression in health insurance fraud identification [8]. Yan et al. proposed an improved recognition algorithm which combined an adaptive genetic algorithm with BP neural network algorithm. And it is used for the identification of auto insurance fraud [9]. Hu used the probit and BP neural network models to analyze auto insurance fraud. Moreover, the selected identification factors were used as input data to build the neural network model [10]. Ma et al. used BP neural network to identify the financial fraud of listed companies. And it increased the recognition rate of fraud in financial statements [11]. Mohamed et al. used backpropagation neural network to perform telecommunication interpolation based on local telecommunication network services. And it is adjusted accordingly to control the speed of obtaining solution [12].

There are many people who have conducted research on the fraud risk detection. However, BP neural network still has some deficiencies in fraud detection, such as weak generalization capacity of model, fixed parameters and low indicators [13, 14, 15, 16]. And the existing corporate fraud risk detection accuracy is low. This paper proposes a fractional-order immune BP neural network classifier with a self-attentive mechanism (SFIBP classifier). The improved artificial immune algorithm with dynamic region contraction strategy is used to optimize the initialization process of the BP neural network. Furthermore, it combines the self-attention mechanism to design the input layer. Finally, Caputo fractional non-causal calculus is used to optimize the parameter updating process in BP neural network. It applies to public data and implements the process of predicting corporate fraud risk. Furthermore, it compares with four other different classifiers. The specific contributions are highlighted below:

- The structure, training method and improvement method of BP neural network are fully analyzed. The Improved artificial immune algorithm, Self-attention mechanism and Caputo score non-causal calculus are used to optimize the BP neural network. Through the comparison experiment on UCI public dataset, it is verified that the classifier has good convergence performance.
- The improved BP neural network classifier is used to predict corporate fraud risk on public data. And it is compared with four different classifiers. Experiments prove the effectiveness and superiority of the improved BP neural network classifier in fraud detection.

This paper is organized as follows. Section 2 introduces the BP neural network and details specific improvements of the SFIBP classifier. Section 3 verifies the validity of the classifier. Section 4 introduces the application of the SFIBP classifier in corporate fraud problems and compares it with other classifiers. Section 5 concludes the paper.

2 SELF-ATTENTION FRACTIONAL BP NEURAL NETWORK CLASSIFIER

BP neural network is improved in several aspects, such as the local optimum, poor performance in high-dimensional data and slow convergence speed. Therefore, we introduce the traditional neural networks and relevant optimization measures [17, 18, 19, 20].

2.1 BP Neural Network

The BP neural network algorithm is a multi-layer Feedforward neural network. And the method of gradient descent is used to train the weight of network connection. The network error sum of squares is minimized in BP neural network [21].

The BP neural network is one of the most widely used types of neural networks. It implements a mapping function from input to output. The training process consists of two main parts: the forward transmission of information and the reverse transmission of errors [22]. The topology of the BP neural network is shown in Figure 1.

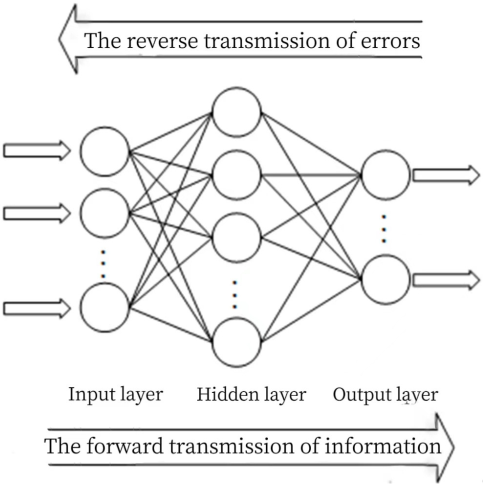


Figure 1. BP neural network topology

2.2 SFIBP Initialization Based on DIAIA Weights

The improved artificial immune algorithm with dynamic region contraction strategy (DIAIA algorithm) has better performance on high dimensional problems. So it is used to optimize the initialization process of the BP neural network to avoid falling into local extremum. At the same time, it makes the classifier achieve better classification effect. The steps of DIAIA algorithm to optimize the weights of BP neural network are as follows.

The first step is to set the objective function. It is the affinity function between antibody and antigen. This paper uses the cross-entropy function as the antibody-antigen affinity function.

$$L_{CrossEntropy} = \frac{1}{N} \sum_{i} -(y_i \log p_i + (1 - y_i) \log(1 - p_i)), \tag{1}$$

where y_i is the class of sample i. If it is positive, y_i is 1. If it is negative, y_i is 0. p_i is the probability that sample i is a positive class. i is a positive class. N is the total number of training samples.

In the second step, the neural network weight is encoded.

In the third step, the algorithm starts to run. And the update process of the antibody is equivalent to the correction process of the parameter.

In the fourth step, the algorithm outputs the optimized network parameter after the algorithm reaching the iteration times or the calculation accuracy.

In the fifth step, the optimized parameter is decoded. And then, it is passed to the network.

2.3 SFIBP's Input Layer Design Based on a Self-Attentive Mechanism

The neural network's generalization capacity is weak after training. So this paper introduces Self-attentive mechanism to adjust the input of neurons [23].

Suppose the data sequence is $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$, where x_i^* represents the i^{th} element in the sequence. Self-attention operation is performed on sequence x^*. The obtained sequence is represented by $x = (x_1, x_2, \ldots, x_n)$. Where x_i represents the i^{th} element in the new sequence.

Firstly, the correlation α between sequences is calculated according to Equation (2).

$$a_{i,j} = q_i \cdot k_j \Rightarrow Vector form : A = K^T \cdot Q, \tag{2}$$

$$a_{ij} = x_i^{*T} x_j^*. \tag{3}$$

Then, according to Equation (4), the weight coefficient β_{ij} is obtained by normalization.

$$\beta_{ij} = \frac{a_{ij}}{\sum_j a_{ij}}. \tag{4}$$

Finally, according to Equation (5), the new sequence expression is obtained by the linear weighted sum of the weighting coefficient and the data sequence itself.

$$x_i = \sum_{j}^{L_x} \beta_{ij} x_j^*, \tag{5}$$

where L_x is the length of the data sequence. Since the network input layer itself has dot multiplication operation of the data sequence, it is only necessary to assign the corresponding β_{ij} to each node of the input layer.

In the practical application of BP neural network, the features of the sample data are not equally important. Therefore, the input neurons are given weights based on the importance of the various sample characteristics. The weighting structure of the BP neural network after adding the attention mechanism is shown in Figure 2.

2.4 SFIBP Weight Update Based on Caputo's Fractional Order Non-Causal Calculus

Gradient descent is the core of weight updating within BP neural network. But the traditional gradient descent method has many drawbacks [24]. Hence, Caputo

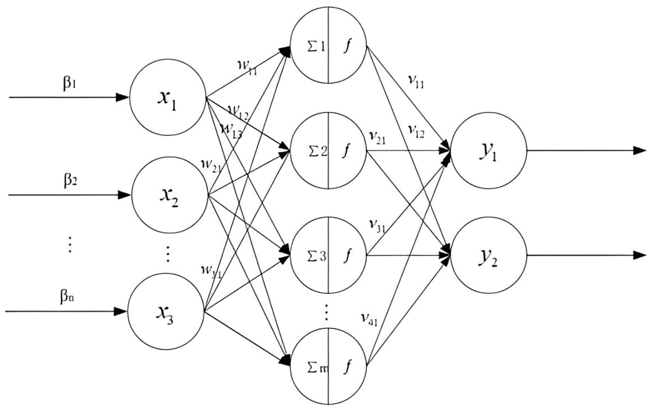


Figure 2. Structure of BP neural network weights after adding weights β

fractional-order non-causal calculus is introduced to update the parameters of the BP neural network [25]. This improves the scientific accuracy of prediction.

The fractional-order non-causal calculus is the linear weighted sum of anti-causal and non-causal calculus. The fractional-order $\alpha \in (0,1)$ is defined by Equation (6).

$$ {}_{d}CaputoD_t^{\alpha} f(t) = \frac{1}{2\sin\frac{\alpha\pi}{2}} ({}_{l}CaputoD_t^{\alpha} f(t) - {}_{r}CaputoD_t^{\alpha} f(t)), \tag{6} $$

where the ${}_{d}CaputoD_t^{\alpha} f(t)$ is the Caputo fractional-order non-causal operator. The ${}_{l}CaputoD_t^{\alpha} f(t)$ is the Caputo fractional-order causal operator and calculated as Equation (7). ${}_{r}CaputoD_t^{\alpha} f(t)$ is the Caputo fractional-order anti-causal operator and calculated as Equation (8).

$$ {}_{l}Caputo_{a}D_t^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \int_{\alpha}^{t} (t-\tau)^{-\alpha} \cdot f'(\tau)\,\mathrm{d}\tau, \tag{7} $$

$$ {}_{r}Caputo_{b}D_t^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \int_{t}^{b} (t-\tau)^{-\alpha} \cdot f'(\tau)\,\mathrm{d}\tau, \tag{8} $$

where $[a,t]$ is the integration intervals of $f(t)$ in fractional causal operators. $[t,b]$ is the integration intervals of $f(t)$ in anti-causal operators.

The updating stage of parameter is improved by fractional non-causal calculus. The parameter updating equation of the improved BP neural network is as follows.

$$ \begin{cases} w^{k+1} = w^k - \eta \cdot_d CaputoD_{w^k}^{\alpha} E_{sample}, \\ v^{k+1} = v^k - \eta \cdot_d CaputoD_{v^k}^{\alpha} E_{sample}, \end{cases} \tag{9} $$

where $\eta > 0$ is the learning rate. $\alpha \in (0,1)$ is the fractional order. k is the number of current iterations. *Example* is the error function obtained from Equation (10). Both ${}_dCaputoD^{\alpha}_{w^k}E_{sample}$ and ${}_dCaputoD^{\alpha}_{v^k}E_{sample}$ are Caputo fractional-order non-causal operators, which are calculated based on Equation (11) and Equation (12).

$$E_{sample} = \frac{1}{2}\sum_{j=1}^{n}(d_j - f(V \cdot g(W \cdot x_j)))^2. \tag{10}$$

X is the n input samples of the BP neural network. d is the corresponding desired output. $W = w_{n*p}$ is the connection weight between the input layer and the hidden layer. p is the number of nodes in the hidden layer. $V = v_{p*q}$ is the connection weight between the hidden and output layers. q is the number of nodes in the output layer. g is the hidden layer transfer function. f is the output layer transfer function.

$$\begin{cases} {}_dCaputoD^{\alpha}_{w^k}E_{sample} = \frac{1}{2\sin\frac{\alpha\pi}{2}}({}_lCaputoD^{\alpha}_{w^k}E_{sample} - {}_rCaputoD^{\alpha}_{w^k}E_{sample}), \\ {}_dCaputoD^{\alpha}_{v^k}E_{sample} = \frac{1}{2\sin\frac{\alpha\pi}{2}}({}_lCaputoD^{\alpha}_{v^k}E_{sample} - {}_rCaputoD^{\alpha}_{v^k}E_{sample}), \end{cases} \tag{11}$$

$$\begin{cases} {}_lCaputo_{cmin}D^{\alpha}_{w^k}E_{sample} = \frac{1}{(1-\alpha)\Gamma(1-\alpha)} \cdot \sum_{j=1}^{n} f'_j(V \cdot g(WX_j)) \\ v^k g'(w^k \cdot X_j)x_j(cmin - w^k)^{1-\alpha}, \\ {}_rCaputo_{cmin}D^{\alpha}_{w^k}E_{sample} = \frac{1}{(1-\alpha)\Gamma(1-\alpha)} \cdot \sum_{j=1}^{n} f'_j(V \cdot g(WX_j)) \\ v^k g'(w^k \cdot X_j)x_j(cmax - w^k)^{1-\alpha}, \\ {}_lCaputo_{cmin}D^{\alpha}_{v^k}E_{sample} = \frac{1}{(1-\alpha)\Gamma(1-\alpha)} \cdot \sum_{j=1}^{n} f'_j(V \cdot g(WX_j)) \\ g(w^k \cdot X_j)(v^k - cmin)^{1-\alpha}, \\ {}_rCaputo_{cmax}D^{\alpha}_{v^k}E_{sample} = \frac{1}{(1-\alpha)\Gamma(1-\alpha)} \cdot \sum_{j=1}^{n} f'_j(V \cdot g(WX_j)) \\ g(w^k \cdot X_j)(cmax - v^k)^{1-\alpha}. \end{cases} \tag{12}$$

$cmin = \min(w^k_{ij}, v^k_{jh})$, $cmax = \max(w^k_{ij}, v^k_{jh})$, $k \in N$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, p$, $h = 1, 2, \ldots, q$. In this paper, the fractional value α is set to 2/9.

3 SFIBP NEURAL NETWORK CLASSIFIER PERFORMANCE TESTING AND ANALYSIS

3.1 Selection and Processing of Datasets

The Iris dataset and Wine dataset are typical datasets. They are often used to test the performance of classifiers. Therefore, these two datasets are used to evaluate the effectiveness of neural network classifiers. The description of the dataset is given in Table 1.

Dataset	Attribute Quantity	Category	Training Sample	Test Sample	Total
Iris	5	setosa	40	10	150
		versicol	40	10	
		virginic	40	10	
Wine	13	class1	41	18	178
		class2	50	21	
		class3	34	14	

Table 1. Dataset description

3.2 Analysis of Results

To verify the effectiveness of the proposed SFIBP classifier, we compared it with the traditional BP neural network algorithm and other improved BP neural network algorithms on the dataset. These algorithms used for comparison are shown in Table 2.

Algorithm Name	Abbreviation	From Literature
Traditional BP neural network	BP	[26]
BP neural network improved by genetic algorithm	GA-BP	[27]
BP network based on inverse time chaotic Coyote	ICCOABP	[28]
SA-BP neural network based on Fuzzy Theory	SA-BP	[29]

Table 2. Comparison of algorithms

Table 3 describes the classification accuracy and standard deviation of the different classifiers on the test dataset. It can be seen that the classification accuracy of the SFIBP classifier is improved. In terms of robustness performance, the classifier exhibits high robustness on the wine dataset.

	Iris		Wine	
Algorithm	ACC	SDev	ACC	SDev
BP	93.25	±1.67	95.83	±1.73
GA-BP	95.28	±1.45	96.34	±1.25
SA-BP	93.33	±1.38	89.12	±1.67
ICCOABP	95.35	±1.72	97.01	±1.16
SFIBP	96.67	±1.64	98.08	±0.94

Table 3. Comparison of the results of the dataset classification experiments

The Radviz technique is used to visualize the actual classification situation and the predicted situation of the SFIBP classifier. Each color represents a category shown in Figures 3, 4, 5 and 6.

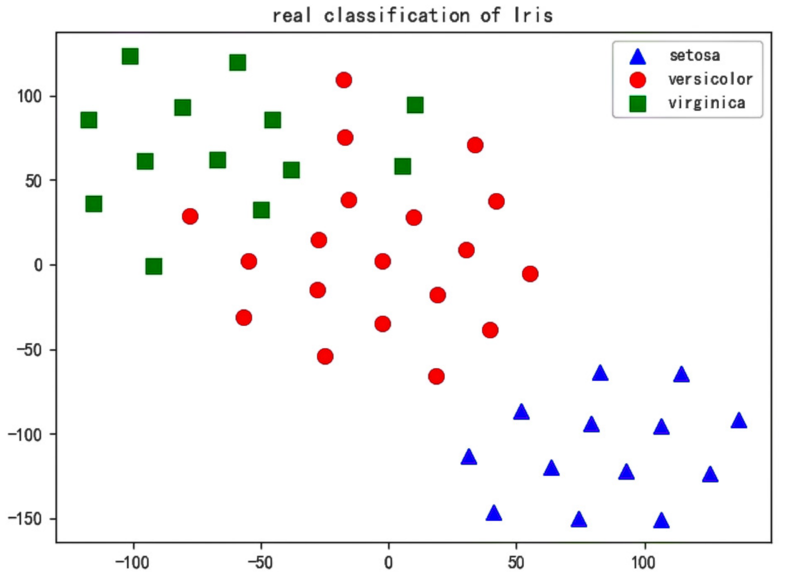


Figure 3. Real classification of Iris test data

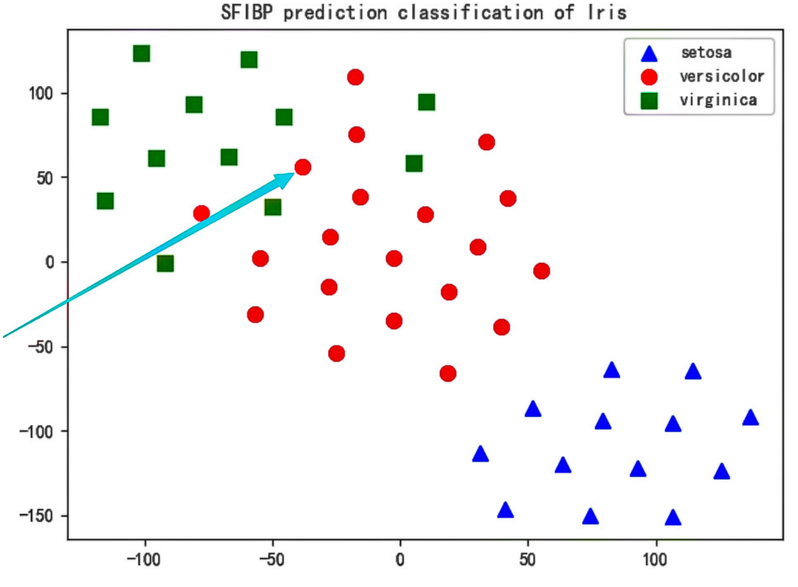


Figure 4. Prediction and classification of Iris test data by SFIBP

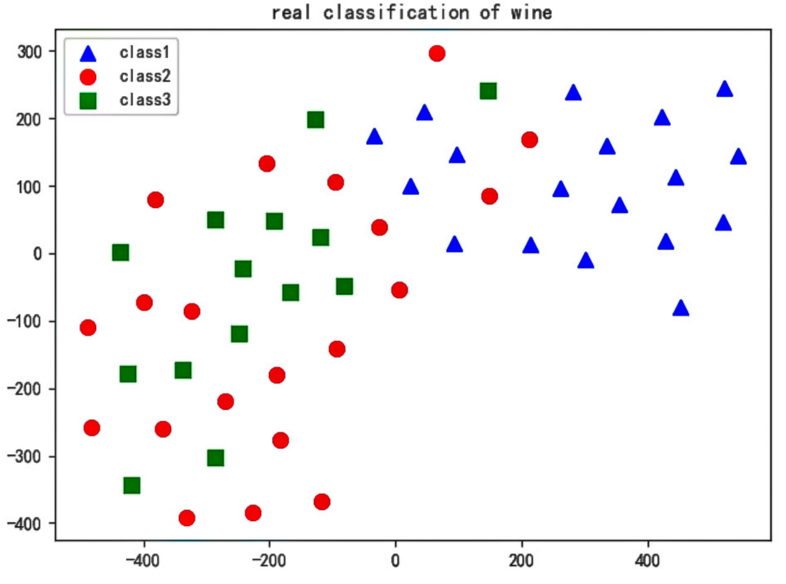


Figure 5. The real breakdown of Wine test data

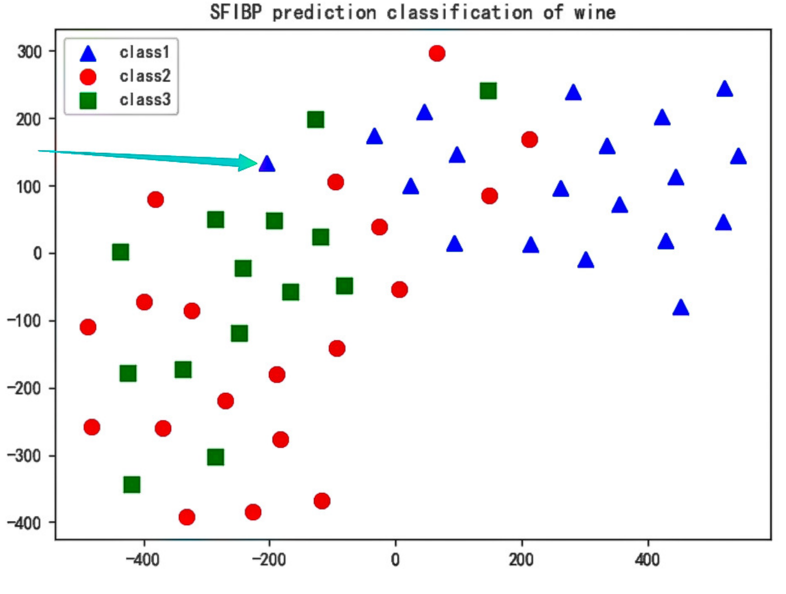


Figure 6. Prediction and classification of Wine test data by SFIBP

4 APPLICATION OF SFIBP NEURAL NETWORK CLASSIFIER TO THE CORPORATE FRAUD PROBLEM

The classification prediction model is designed for the risk detection of corporate fraud in audit work. This section compares the results of the four algorithms on the public dataset of the Audit Office of India with the prediction result of the SFIBP classifier. These four classifiers are AdaBoost algorithm (AB), Random Forest algorithm (RF), Support Vector Machine algorithm (SVM), and Bayesian Network algorithm (BN).

4.1 Experimental Data Selection

The dataset used in this paper is based on publicly available data from the comptroller and auditor general of India. It is the data of company revenue and expenditure in 2018. And it is collected by the audit commission of India. Data sources from the annual data of 776 companies in 46 different cities [28]. It includes 14 sectors, such as public health, livestock, fisheries and industry. The content is shown in Table 4.

ID	Information	Number of Enterprises
1	Irrigation	114
2	Public Health	77
3	Buildings and Roads	82
4	Forest	70
5	Company	47
6	Animal Husbandry	95
7	Signal Communication	1
8	Electrical	4
9	Land	5
10	Science and Technology	3
11	Travel	1
12	Fisheries	41
13	Industry	37
14	Agriculture	200

Table 4. Data information

The dataset contains 26 different attribute features. It includes the previous records of the National Audit Office, audit paragraphs, environmental status report, etc. Table 5 explains some of the attribute features. This dataset has a large amount of data and complex case characteristics. Therefore, the data is preprocessed and dimension reduced by factor analysis. After preliminary processing, the data still has 25 attribute features and 773 samples. The dimension-reduced data has 7 attribute features including 85.125 % of the original data information.

Attribute Name	Meaning
Para A value	Planned expenditure and differences found in report A
Para B value	Planned expenditure and differences found in report B
Total	Total differences found in other reports
Number	Historical difference score
Money value	Amount involved in misstatement in past audits
Sector score	Historical risk score
Loss	The amount of losses suffered by the company last year
History	Average historical losses suffered in the past ten years
District score	Historical risk score of a region
Sector ID	Unique ID of the target Department
Location ID	Unique ID of the city or province
Audit ID	Unique ID assigned to the audit case
ARS	Risk score

Table 5. Explanation of attribute characteristics in the model

4.2 SFIBP Classifier Calculation

In order to better detect risk, the overall structure of the network needs to be determined. Because the proper network structure can provide better results, this paper adopts the network structure with a single hidden layer.

1) **Determination of input and output layer nodes.** In this paper, seven characteristics are extracted from the annual data of different companies. Accordingly, the number of nodes in the network's input layer is 7. The output layer nodes are set to 2 corresponding to two actual output categories: risky and risk-free companies.
2) **Determination of hidden layer nodes.** Since there is no uniform theoretical guidance for selecting hidden layer nodes, this paper learns them from the empirical formula given by some researchers.

In order to select the appropriate number of hidden layer nodes, different numbers of hidden layer nodes are used for experiments on dataset [30]. The variation in loss value is obtained, as shown in Figure 7.

Figure 7 shows the change in the loss value of using different nodes under the same dataset. As can be seen from the figure, the number of 21 hidden layer nodes is large. So it obviously falls into the local extremum before 4 000 iterations. Although it is out of the local optimal solution later, its loss value is still higher than the loss value of hidden layer nodes of 8 or 15. The variation in loss values for 8 and 15 hidden layer nodes are very close. Consequently, they are compared and shown in Figure 8.

As shown in Figure 8, the weight combination with low loss value is initially found when 8 hidden layer nodes are used. However, the late convergence ability is poor. Therefore, the 15 hidden layer nodes are chosen for risk detection.

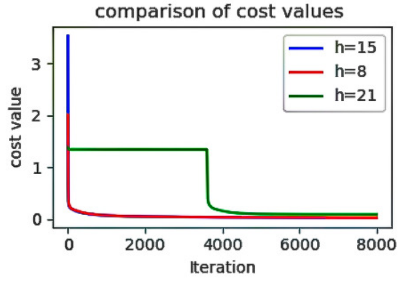


Figure 7. Comparison of the change in loss values for hidden layer nodes of 15, 8 and 21, respectively

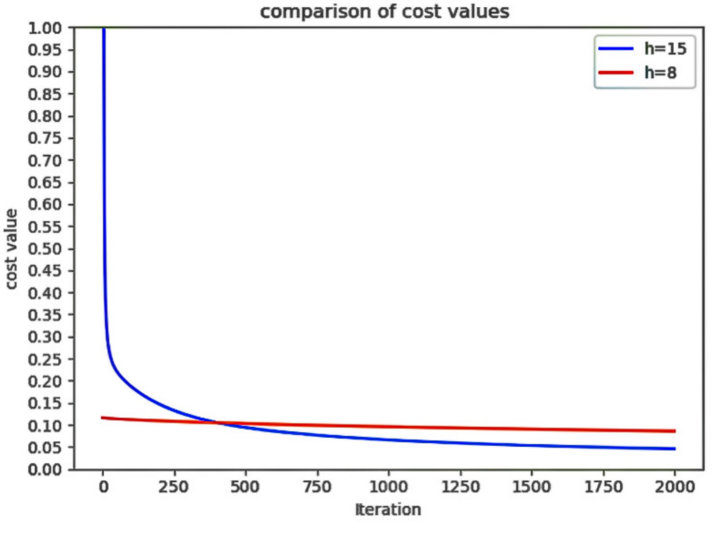


Figure 8. Comparison of the change in loss values for hidden layer nodes of 15 and 8, respectively

4.3 Evaluation Index System of Corporate Fraud Risk Prediction Model

In order to test the performance of the classifier, nine performance metrics are set to validate it. Those are Type I error, Type II error, accuracy, error rate, sensitivity, specificity, F1 score, F2 score and Matthew correlation coefficient. The performance metrics are shown in Table 6.

The classifier is trained to distinguish between fraudulent and non-fraudulent companies. The confusion matrix summarises the results of category fraud. X (True positive) corresponds to the hit value, it gives the number of fraudulent companies correctly predicted to be fraudulent. Z (False negative) gives the number of

Performance Index	Formula
Type I error (Type-I)	Q
Type II error (Type-II)	Z
Sensitivity (Sens)	$X/(X+Z)$
Specificity (Spec)	$Y/(Q+Y)$
Accuracy (Acc)	$(X+Y)/(X+Q+Y+Z)$
Error rate (Error R)	$1-\text{accuracy}$
F1 score	$(2X)/(2X+Q+Z)$
F2 score	$(5X)/(5X+4Z+Q)$
Matthew correlation coefficient (MCC)	$(XY)-\frac{(QZ)}{\sqrt{(X+Q)(X+Z)(Y+Q)(Y+Z)}}$

Table 6. Performance evaluation indicators

True	Forecast	
	Fraud exists	No fraud
Fraud exists	True positive X	False negative Z
No fraud	False positive Q	True negative Y

Table 7. Confusion matrix

fraudulent companies that are wrongly marked as non-fraudulent companies. It is called Type II error. Q (False positive) gives the number of non-fraudulent companies incorrectly labeled as fraudulent. It is called Type I error. Moreover, Y (True negative) indicates the number of correctly classified non-fraudulent companies.

4.4 Experimental Analysis

In this paper, nine performance evaluation indicators consider the results of corporate fraud risk detection. The K-fold cross-validation is a common method to verify model robustness in machine learning. 70 % of the total dataset is taken as the training set. And the other 30 % data is used as the testing set. We compare the results of the AB algorithm, RF algorithm, SVM algorithm and BN algorithm with the prediction of the SFIBP classifier on this dataset. The results are shown in Table 8.

Classifier	Type-I	Type-II	Sens	Spec	Acc	Error R	F1	F2	MCC
AB	0.05	0.11	0.87	0.95	0.91	0.09	0.92	0.89	0.81
RF	**0.01**	0.12	0.88	**0.99**	0.92	0.08	0.93	0.90	0.85
SVM	0.10	0.14	0.85	0.90	0.87	0.13	0.89	0.86	0.73
BN	**0.01**	0.07	0.91	**0.99**	0.94	0.06	0.95	0.92	0.88
SFIBP	**0.01**	**0.002**	**0.99**	0.98	**0.99**	**0.01**	**0.98**	**0.99**	**0.97**

Table 8. Comparison of the performance of different classifiers

In the case of audit issues, false positive only lead to further scrutiny of trustful company. However, false negative may allow the fraudulent company to lurk. By contrast, the latter is more dangerous. Therefore, Type II error should be given greater attention when analyzing Type I and Type II error.

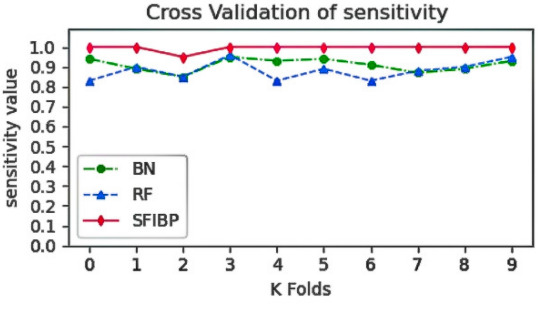
As shown in Table 8, the SFIBP classifier has a relatively low error rate in Type II error. Though the specificity is not as high as the RF and BN algorithm. But specificity is not as important as sensitivity. And the SFIBP classifier has a significant improvement in sensitivity. Overall, the SFIBP classifier has higher classification accuracy and better robustness.

To sum up, the BN and RF algorithm perform better than the AB and SVM algorithm. At the same time, the BN algorithm, RF algorithm and SFIBP classifier have the same Type I error. But the specificity of these two algorithms is relatively higher. Therefore, the indicators of the three algorithms are further graphically analyzed.

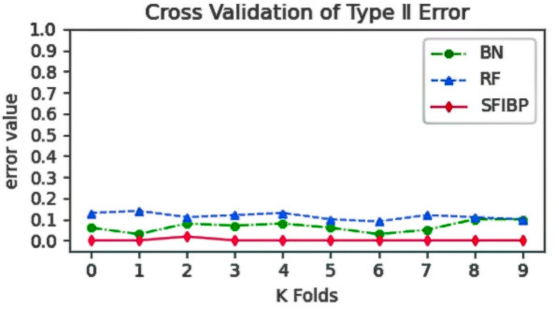
Figure 9 compares the sensitivity, Type II error, and accuracy for ten cross-validations of three algorithms during the training process. Figure 9 a) compares the sensitivity of the three algorithms. It can be seen that the BN algorithm has a relatively high sensitivity. However, the sensitivity of SFIBP neural network classifier is less than 1 only once. The classifier misidentified fraudulent firms as non-fraudulent firms only once out of 10 validations. This indicates that the SFIBP classifier has a high sensitivity. Figure 9 b) compares the Type II error for the three algorithms. It can be seen that Type II errors of BN algorithm appear most among the three algorithms. And the RF algorithm is second. The SFIBP classifier has minor Type II error. There is only one Type II error, which is a meager percentage. Figure 9 c) compares the accuracy of the three algorithms. It can be seen that the accuracy of SFIBP classifier is the highest among the three algorithms. In the case of only one Type II error, there are four instances where the Type II error is less than 100 % accurate. This indicates that Type I error occurs several times. But Type I error has a little effects on risk for detection. Overall the SFIBP classifier has the highest accuracy of the three algorithms. The average accuracy of the SFIBP classifier is 98.7 % in the training process.

After network training, the remaining 30 % of 232 new samples are used for testing our model. This can check whether the classifier appears to be overfitting. At the same time, the validity of the classifier in corporate fraud detection can be tested. At the completion of the test, four prediction errors are found in the 232 samples. And the error rate is 1.7 %. In order to facilitate observation, Figure 10 and Figure 11 are drawn as follows. Figure 10 shows the proper classification of the new sample. And Figure 11 shows the classifier's predicted classification of the new sample.

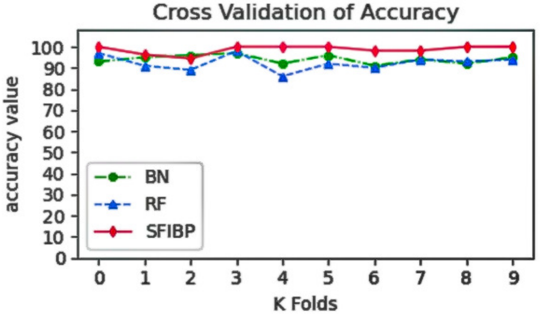
It is clear from the comparison between Figure 10 and Figure 11. The accuracy of the SFIBP classifier on new samples is 98.28 %. This demonstrates its high performance on the issue of corporate fraud risk detection. And it has dramatically improved the efficiency of the detection.



a) Sensitivity



b) Type-II Error



c) Accuracy

Figure 9. Ten-fold cross-validation of BN, RF and SFIBP classifiers on sensitivity, Type II error and accuracy

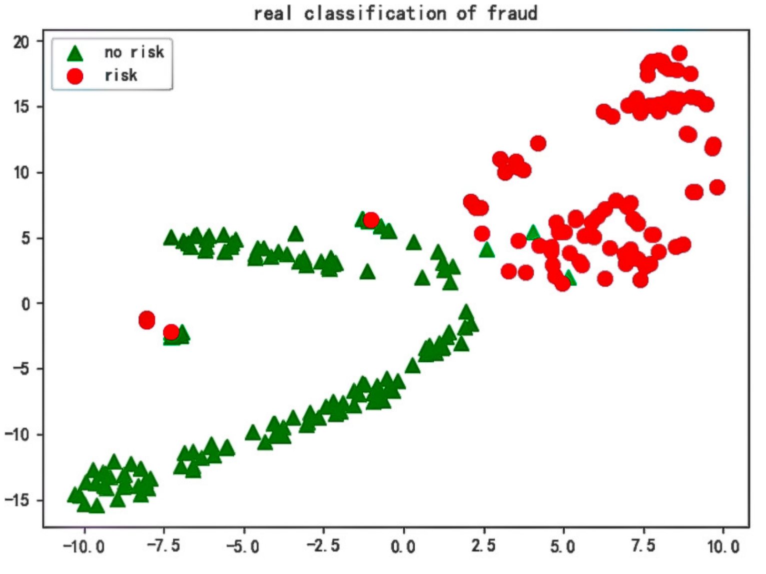


Figure 10. True breakdown of fraud test data

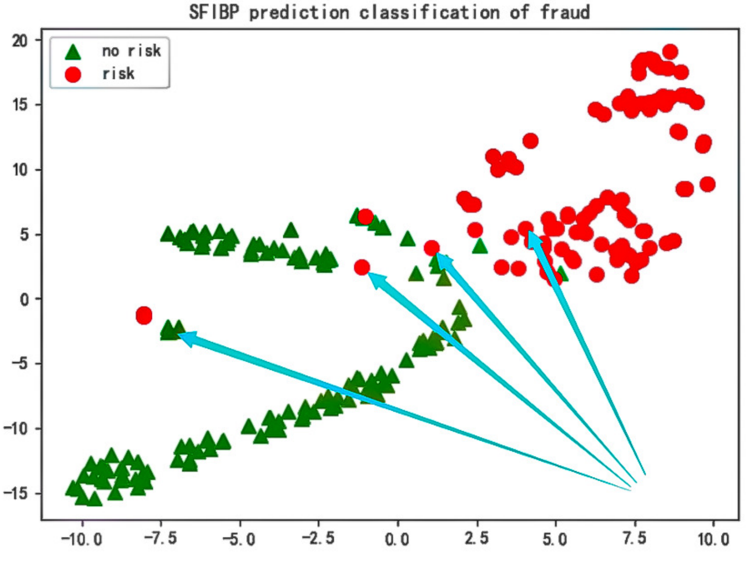


Figure 11. SFIBP's predictive breakdown of fraud test data

5 CONCLUSION

A classification model is designed to address the problem of corporate fraud risk detection. The Improved artificial immune algorithm, Self-attention mechanism and Caputo score non-causal calculus are used to optimize the BP neural network. It can help auditors to predict the fraud risk of company. And it is compared with four different classifiers. Experiments demonstrate that the SFIBP classifier has a higher classification accuracy in the fraud detection problem. And all performance indicators have improved. This also verifies its effectiveness and superiority in detecting corporate fraud risk. On the basis of previous studies, the performance optimization and application of BP neural network are improved in this paper. But there are still some shortcomings which need to be further improved.

Acknowledgement

This research was supported by the Natural Science Foundation of Shandong province under Grants No. ZR2020MF033 and No. ZR2022MG059. Besides, this research was supported by the National Key Project of Statistical Scientific Research of China (No. 2019LZi0). This research was also supported by the General Project of Science and Technology Plan of Beijing Municipal Commission of Education, Grant No. KM202010017001.

REFERENCES

[1] Wang, A.—Ma, B.—Hu, H.: The Progress of Research on Corporate Fraud. Economic Perspectives, Vol. 48, 2019, No. 2, pp. 115–132 (in Chinese).

[2] Liu, B.—He, Y.: Application of Artificial Intelligence Algorithm in the Field of Anti-Fraud under Epidemic Situation. China Financial Computer, Vol. 27, 2020, No. 4, pp. 55–57 (in Chinese).

[3] Forough, J.—Momtazi, S.: Sequential Credit Card Fraud Detection: A Joint Deep Neural Network and Probabilistic Graphical Model Approach. Expert Systems, Vol. 39, 2022, No. 1, Art. No. e12795, doi: 10.1111/exsy.12795.

[4] Zhang, J.—Vukotic, I.—Gardner, R.: Anomaly Detection in Wide Area Network Mesh Using Two Machine Learning Anomaly Detection Algorithms. CoRR, 2018, doi: 10.48550/arXiv.1801.10094.

[5] Alharbi, A.—Alshammari, M.—Okon, O. D.—Alabrah, A.—Rauf, H. T.—Alyami, H.—Meraj, T.: A Novel text2IMG Mechanism of Credit Card Fraud Detection: A Deep Learning Approach. Electronics, Vol. 11, 2022, No. 5, Art. No. 756, doi: 10.3390/electronics11050756.

[6] Yan, C.—Zhang, X.: Life Insurance Customer Churn Prediction Algorithm Based on Improved K-Means and BP-Adaboost. Journal of Shandong University of Scien-

Ce and Technology: Natural Science Edition, Vol. 41, 2022, No. 1, pp. 54–65, doi: 10.16452/j.cnki.sdkjzk.2022.01.006 (in Chinese).

[7] Cheng, P.—Chen, D.—Wang, J.: Clustering of the Body Shape of the Adult Male by Using Principal Component Analysis and Genetic Algorithm–BP Neural Network. Soft Computing, Vol. 24, 2020, No. 17, pp. 13219–13237, doi: 10.1007/s00500-020-04735-9.

[8] Yang, C.: Research on Health Insurance Fraud Identification Based on BP Neural Network. Ph.D. Thesis. Qingdao University, 2014.

[9] Yan, C.—Li, M.—Zhou, X.: Improved Genetic Algorithm for Vehicle Insurance Fraud Identification Model Based on BP Neural Network. Journal of Shandong University of Science and Technology: Natural Science Edition, Vol. 38, 2019, No. 5, pp. 72–80, doi: 10.16452/j.cnki.sdkjzk.2019.05.009 (in Chinese).

[10] Hu, Z.: Research on Auto Insurance Fraud Identification Based on Neural Network. Master Thesis. Guizhou University of Finance and Economics, 2021, doi: 10.27731/d.cnki.ggzcj.2021.000425 (in Chinese).

[11] Ma, X.—Li, X.—Song, Y.—Zheng, X.—Zhang, Z.—He, R.: A BP Neural Network for Identifying Corporate Financial Fraud. 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, pp. 191–193, doi: 10.1109/ISI.2019.8823408.

[12] Mohamed, A.—Bandi, A. F. M.—Tamrin, A. R.—Jaafar, M. D.—Hasan, S.—Jusof, F.: Telecommunication Fraud Prediction Using Backpropagation Neural Network. 2009 International Conference of Soft Computing and Pattern Recognition, 2009, pp. 259–265, doi: 10.1109/SoCPaR.2009.60.

[13] Xu, X.—Arshad, M. A.—Ali, U.—Mahmood, A.: The GM-BP Neural Network Prediction Model for International Competitiveness of Computer Information Service Industry. Algorithms, Vol. 14, 2021, No. 11, Art. No. 308, doi: 10.3390/a14110308.

[14] Adamo, J. M.—Anguita, D.: Object Oriented Design of a Simulator for Large BP Neural Networks. In: Mira, J., Sandoval, F. (Eds.): From Natural to Artificial Neural Computation (IWANN 1995). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 930, 1995, pp. 642–649, doi: 10.1007/3-540-59497-3_233.

[15] Liu, T.—Jiang, T.—Li, A. et al.: Remote Sensing Estimation of Forest Stock Volume Based on Neural Network and Different Site Quality. Journal of Shandong University of Science and Technology (Natural Science), Vol. 38, 2019, No. 2, pp. 25–35.

[16] Zhang, M.—Liu, D.—Wang, Q.—Zhao, B.—Bai, O.—Sun, J.: Detection of Alertness-Related EEG Signals Based on Decision Fused BP Neural Network. Biomedical Signal Processing and Control, Vol. 74, 2022, Art. No. 103479, doi: 10.1016/j.bspc.2022.103479.

[17] Zhang, D.—Yang, G.—Zhao, X.—Yang, X.—Kaohsiung, F.: Optimization Design of Vane Diffuser and Volute in Vertical Centrifugal Pump Based on Back Propagation Neural Network. Journal of Agricultural Machinery, Vol. 53, 2022, No. 4, pp. 130–139 (in Chinese).

[18] Dong, C.—Dong, L.—Yang, M.: The Application of the BP Neural Network in the Nonlinear Optimization. In: Cao, B., Wang, G., Guo, S., Chen, S. (Eds.): Fuzzy Information and Engineering 2010. Springer, Berlin, Heidelberg, Advances in

Intelligent and Soft Computing, Vol. 78, 2010, pp. 727–732, doi: 10.1007/978-3-642-14880-4_80.

[19] Jia, W.—Zhao, D.—Shen, T.—Ding, S.—Zhao, Y.—Hu, C.: An Optimized Classification Algorithm by BP Neural Network Based on PLS and HCA. Applied Intelligence, Vol. 43, 2015, No. 1, pp. 176–191, doi: 10.1007/s10489-014-0618-x.

[20] Maru, A.—Dutta, A.—Kumar, K. V.—Mohapatra, D. P.: Software Fault Localization Using BP Neural Network Based on Function and Branch Coverage. Evolutionary Intelligence, Vol. 14, 2021, No. 1, pp. 87–104, doi: 10.1007/s12065-019-00318-2.

[21] Zhao, Y.: Research and Application on BP Neural Network Algorithm. 2015 International Industrial Informatics and Computer Engineering Conference, 2015, pp. 1444–1447, doi: 10.2991/iiicec-15.2015.321.

[22] Zhang, X.—Tan, J.—Han, J.: Fault Diagnosis Method Based on BP Neural Network. Ph.D. Thesis. 2002 (in Chinese).

[23] Bahdanau, D.—Cho, K.—Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate. CoRR, 2014, doi: 10.48550/arXiv.1409.0473.

[24] Tian, Y.—Liang, Y. Q.—Peng, Y. J.: Cuckoo Search Algorithm Based on Stochastic Gradient Descent. In: Krömer, P., Zhang, H., Liang, Y., Pan, J. S. (Eds.): Proceedings of the Fifth Euro-China Conference on Intelligent Data Analysis and Applications (ECC 2018). Springer, Cham, Advances in Intelligent Systems and Computing, Vol. 891, 2019, pp. 90–99, doi: 10.1007/978-3-030-03766-6_10.

[25] Huang, J.—Wang, J.: Fractional-Order Non-Causal BP Neural Networks Model. Journal of Computer Engineering and Applications, Vol. 57, 2021, No. 23, pp. 91–97 (in Chinese).

[26] Li, Y.: Research and Application of Improved Algorithm of Classifier Based on BP Neural Network. Master Thesis. China University of Geosciences, Beijing, 2019, doi: 10.27493/d.cnki.gzdzy.2019.001390 (in Chinese).

[27] Ye, X.—Yao, P.—Long, F.—Zhuang, Z.: Iris Image Real-Time Pre-Estimation Using Compound BP Neural Network. In: Zhang, D., Jain, A. K. (Eds.): Advances in Biometrics (ICB 2006). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 3832, 2005, pp. 450–456, doi: 10.1007/11608288_60.

[28] Hooda, N.—Bawa, S.—Rana, P. S.: Fraudulent Firm Classification: A Case Study of an External Audit. Applied Artificial Intelligence, Vol. 32, 2018, No. 1, pp. 48–64, doi: 10.1080/08839514.2018.1451032.

[29] Xiong, H.: Research and Implementation of Sensitive Text Classification Algorithm Based on Artificial Immune System. Master Thesis. Xi'an University of Electronic Science and Technology, 2021, doi: 10.27389/d.cnki.gxadu.2021.003047 (in Chinese).

[30] Wang, R.—Xu, H.—Li, B.—Feng, Y.: Research on Method of Determining Hidden Layer Nodes in BP Neural Network. Computer Technology and Development, Vol. 28, 2018, No. 4, pp. 31–35 (in Chinese).

Wei Liu is engaged in research work in workflow, service computing, Petri net theory and application, software formal analysis and verification, and Big Data intelligent analysis and processing.

MingMing Liu is now pursuing her Master's degree at the Shandong University of Science and Technology, Qingdao, China. Her research interests include intelligence algorithm optimisation and machine learning.

Chun Yan received her B.Sc., M.Sc. and Ph.D. degrees from the Shandong University of Science and Technology, Qingdao, China, in 2000, 2003 and 2011, respectively. She is presently Associate Professor at the Shandong University of Science and Technology, Qingdao, China. Her research interests include applied mathematics, statistics and economic management. She has about 20 technical papers published in journals and conference proceedings in her research areas.

Man Qi is a Senior Lecturer in computing at Canterbury Christ Church University, UK. Her main research interests are in the areas of intelligent systems and applications, data analytics, cyber security and HCI. Her research has been founded by EPSRC, EU and QR, etc. She is a Fellow of the British Computer Society (FBCS) and a Fellow of the Higher Education Academy (FHEA). She has published over 70 research papers and is on the editorial board of five international journals. She has been an external Ph.D. examiner for a number of universities in the UK and Australia. She has served as a Chair and Program Committee member for over 50 international conferences and has been a long-term reviewer for many international journals.

LuLu Zhang received her Master's degree from the Shandong University of Science and Technology, China, in 2022. Her research interests include intelligence algorithm optimisation.