Tax Risk Early Warning System for SMEs Using Auto Encoder-Backpropagation Neural Network with Genetic Algorithm Optimization

Zhen Chen
Department of Finance and Economics, Software Engineering Institute of Guangzhou, Guangzhou 510990, China
Email: 15653056654@163.com

Keywords: auto encoder, BP neural network, small and medium-sized enterprises, taxation, risk warning

Received: September 23, 2024

In the current economic circumstances, the tax risk management needs of small and medium-sized enterprises are becoming increasingly urgent. Effectively warning and controlling tax risks has become a critical issue of common concern for enterprise managers and tax departments. In view of this, the study first used the analytic hierarchy process to determine the indicator system and the corresponding weights for each indicator. Secondly, a fusion model for detection and warning was constructed by combining auto encoders with backpropagation neural networks. In the experiment, the study used a self-built data set containing approximately 280,000 samples to compare the performance of models such as support vector machines and gradient boosting decision trees. The performance test results showed that the optimal model parameters after genetic algorithm optimization were 8 hidden layer neurons, 11 hidden layers, dropout rate of 00.48, and learning rate of 0.024. When the amount of iterations was 1000, the loss function value of the model was 0.05, the F1 score was 0.96, the average absolute error was 0.05, and the accuracy was 0.94. In the simulation test, the model had the highest success rate of warning for the manufacturing industry, at 90.12%, and the average success rate of warning for different industries was 86.18%. The experiment findings indicated that the model exhibited high accuracy and reliability in tax risk warning. Therefore, the research not only provides a new technological means for tax risk management of small and medium-sized enterprises, but also provides a certain reference for further research and application in related fields.

Povzetek: Razvit je nov sistem zgodnjega opozarjanja na davčna tveganja za mala in srednja podjetja z združitvijo avtoenkoderjev in nevronskih mrež, optimiziranih z genetskim algoritmom, kar izboljšuje kvaliteto napovedi.

1 Introduction

As the advancement of global economic integration and informatization, tax management has become an important component of government fiscal revenue in various countries [1]. The tax department not only needs to ensure the effective implementation of tax policies, but also continuously raise the level of tax management and services to adapt to the increasingly complex economic environment. As an important component of the national economy, tax management is particularly critical for small and medium-sized enterprises (SEMs) [2]. SEMs not only have a large number, but also take a pivotal part in promoting employment, innovation, and economic growth [3]. In the last few years, with the government's emphasis on the development of SEMs, various supportive policies have been introduced, including tax reductions, financing support, and innovation incentives [4]. Meanwhile, the advancement of information technology has brought new opportunities for tax management, and the intelligence and informatization of tax management have become the trend [5]. However, with the advancement of information technology, the complexity and data volume of tax management have significantly increased, which has raised higher requirements for early warning and control of tax risks. Ding Q et al. raised a risk warning (RW) management model for financial enterprises with fuzzy theory. By combining analytic hierarchy process (AHP) and fuzzy evaluation method, a modeling experiment was built on the financial risks of a listed company, and RW and evaluation were carried out on private enterprise projects. The findings of the research indicated that the utilization of fuzzy theory and contemporary network technology could facilitate a more precise identification and evaluation of the inherent and apparent risks associated with financial enterprises [6]. Zeng H proposed a cash flow-based enterprise RW model and established a financial RW indicator system. Backpropagation neural network (BPNN) was applied to mine financial data, and mobile edge computing service was introduced to improve data processing performance and timeliness of RW. The experiment indicated that the prediction accuracy of the designed warning model reached 91.6% [7]. Li X et al. raised an optimized BPNN as a financial early warning model against the financial uncertainty and risks faced by Chinese enterprises in the context of the "Internet plus" era. By analyzing the financial data of listed companies from 2017 to 2020, the accuracy of the optimized model's predictions exceeded 80%. The findings of the research study indicated that the optimized BPNN was an effective method for predicting financial risk [8]. Scholar Yi X developed an optimized and integrated RW model for financial risk prediction by combining the expected maximization algorithm, Borderline STATE EasyEnsemble sampling method, and support vector machine (SVM). Research results denoted that the predictive ability of the model for a small amount of samples significantly improved, providing effective measures for enterprises to prevent financial risks [9].

BPNN, as a classic supervised learning algorithm, has strong nonlinear modeling capabilities and can achieve accurate prediction of complex data. Du C et al. raised a hybrid genetic algorithm (GA) and particle swarm optimization BPNN algorithm to meet the sensitivity calibration requirements of three-axis accelerometers under different temperature conditions. In the experiment, the accelerometer was calibrated at different temperatures, and the outcomes illustrated that the algorithm had the smallest prediction error, demonstrating effectiveness and reliability in sensitivity calibration [10]. Liu X et al. proposed an ergonomic reliability model based on improved BPNN and human cognitive reliability for the human-machine interface design of medical equipment visual display terminals in high load and high infection rate working environments. The effectiveness of the model was verified through eye tracking experiments, demonstrating the superiority of the improved BPNN in simulating the reliability evaluation of medical equipment operation in high load environments [11]. Ma C et al. raised a new method with the BPNN model of multiple historical test cases. Study collected the variable values during program execution at different breakpoints, and used this data to train an improved BPNN model. The experiment showed that the average prediction accuracy of this method was 95.8%, and the recall rate was 78.9%. The research outcomes indicated that the BPNN model could effectively generate test predictions [12]. Scholar Shen H proposed an active financial RW system based on BPNN to address the importance of financial risk assessment in the financial ecosystem. An efficient RW model was constructed by optimizing parameters using GA. The experiment outcomes indicated that the accuracy of the model reached 97.94%, an improvement of 21.32% compared to other methods, and a reduction of 45.69% in error, demonstrating significant superiority in financial risk analysis [13]. Finally, the research summarized the research areas, indicator testing results, and limitations of the literature review above, as shown in Table 1.

Table 1: Literature summary table

Authors	Year	Algorithms/Methods	Key results	Limitations
Ding [6]	2021	Fuzzy Theory + AHP	Improved accuracy in financial enterprise risk early warning	Increased model complexity and limited application range
Zeng [7]	2022	Edge Computing + BPNN	Prediction accuracy: 91.6% Response rate :2.1%~5%	High resource demands for edge computing
Li, Wang, Yang [8]	2023	Optimized BPNN	Prediction accuracy: ≥80%	Insufficient model stability and interpretability
Yi [9]	2023	Data Mining+Genetic Algorithm+SVM	The small sample prediction ability was improved by 14.83%	Limited applicability to large-scale datasets
Du, Kong [10]	2022	GA-PSO-BPNN Hybrid Algorithm	Prediction error: about ±0.1	Further validation needed for broader environmental applications
Liu et al. [11]	2023	Improved BPNN+HCR	Enhanced reliability of medical equipment operation in high-load environments	The model requires broader application testing
Ma et al. [12]	2021	BPNN + Historical Test Cases	Average prediction accuracy: 95.8%. Percentage of failed test cases: ≤5% Average recall rate: 78.9%.	Scalability of the model has not been fully validated
Shen [13]	2023	BPNN + Genetic Algorithm	Error rate reduced by 45.69%, Prediction accuracy：97.94%	The adaptability in dynamic environments requires further study

Combined with Table 1, although many scholars have organized massive research on BPNNs and tax RW, most existing tax RW models are based on a single machine learning algorithm and lack comprehensive application and optimization of multiple algorithms. Moreover, the existing tax RW models are mostly black box models, lacking transparency and interpretability, making it difficult to gain the trust and recognition of users. In view of this, the study will combine Auto Encoder and BP neural network, and introduce GA to improve their parameter selection, to construct an Auto Encoder-Backpropagation Neural Network (AEBP) SEM tax RW model. The study aims to raise the accuracy and reliability of tax RW for SEMs, and provide strong support for tax management of SEMs.

2 Methods and materials

2.1 Establishment of tax risk warning indicator system for SMEs

In the research of tax RW for SEMs, establishing a scientific RW indicator system is the most crucial step. The establishment of a warning indicator system can not only effectively reflect the financial health and tax compliance of enterprises, but also provide reliable data support for subsequent risk prediction models [14]. Through systematic analysis and screening, it can determine the most representative and sensitive risk indicators to ensure the accuracy and practicality of the early warning system. The choose of tax RW indicators is diverse and requires adherence to the three principles of accessibility, criticality, and sensitivity. The selected tax RW indicators for research are shown in Figure 1.

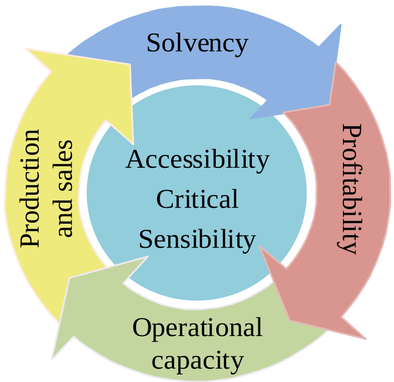


Figure 1: Tax risk early warning indicators

As shown in Figure 1, the study divides tax warning indicators into solvency indicators, profitability indicators, operational capability ratio indicators, and production and sales indicators. The solvency index is mainly used to assess the ability of a company to repay its debts, the profitability index refers to the ability of the company to obtain profits, the operational capability ratio index mainly reflects the efficiency of the company's daily operations, and finally the production and sales indicators are related to the company's production and sales activities. The tax warning system can evaluate the tax risks of enterprises based on different combinations and weights of these indicators. Firstly, in the solvency index, it is divided into current ratio, quick ratio, and asset liability ratio, and their expressions are shown in equation (1).

$$\begin{cases} R_{current} = A_{current} / L_{current} \\ R_{quick} = A_{\mathrm{quick}} / L_{current} \\ R_{A_L} = L_{total} / A_{total} \end{cases} \quad (1)$$

In equation (1), $R_{current}$ represents the current ratio, $A_{current}$ represents current assets, and $L_{current}$ represents current liabilities. R_{quick} represents quick ratio. A_{quick} represents quick assets. R_{A_L} represents the asset liability ratio, L_{total} represents total liabilities, and A_{total} represents total assets. An enterprise with a higher current ratio and quick ratio is considered to have stronger short-term solvency. The lower the asset liability ratio, the more stable the financial structure of the enterprise and the stronger its debt paying ability. Secondly, profitability indicators can be expressed in three forms: operating profit margin, total asset net profit margin, and equity net profit margin, as shown in equation (2).

$$\begin{cases} MP_{operating} = P_{operating} / R_{operating} \\ MN_{profit} = N_{profit} / A_{total} \\ ENM_{profit} = N_{profit} / SE \end{cases} \quad (2)$$

In equation (2), $MP_{operating}$ represents operating profit margin, $P_{operating}$ represents operating profit, and $R_{operating}$ represents operating revenue. MN_{profit} represents asset net profit margin, N_{profit} represents net profit, and A_{total} represents total assets. ENM_{profit} represents equity net profit margin, and SE represents shareholder equity. In the operational capability ratio indicator, the expressions for inventory turnover rate, accounts receivable turnover rate, and total asset turnover rate are shown in equation (3).

$$\begin{cases} R_{I_{turnover}} = R_{sales} / I \\ R_{AR_{turnover}} = R_{sales} / R_{accounts} \\ R_{TR_{turnover}} = R_{sales} / A_{T_{average}} \end{cases} \quad (3)$$

In equation (3), $R_{I_{turnover}}$ represents inventory turnover, R_{sales} is sales revenue, and I is inventory. $R_{AR_{turnover}}$ represents accounts receivable turnover rate. $R_{accounts}$ represents accounts receivable. $R_{TR_{turnover}}$ is the total asset turnover rate. $A_{T_{average}}$ is the average total assets. Finally, the expression for the production sales rate indicator is shown in equation (4).

$$\begin{cases} R_{O_{cost}} = C_{operating} / R_{operating} \\ R_{PE} = E_{period} / R_{sales} \\ R_{N_{OE}} = E_{N_{Operating}} / R_{sales} \\ R_{B_{Taxberben}} = T_S / R_B \end{cases} \quad (4)$$

In equation (4), $R_{O_{cost}}$ is the operating cost rate, $C_{operating}$ is the operating cost, and $R_{operating}$ is the operating revenue. R_{PE} is the period expense ratio, and E_{period} is the period expense sales revenue. $R_{N_{OE}}$ is the non operating expenditure rate. $E_{N_{Operating}}$ is the non operating expenditure. $R_{B_{Taxberben}}$ is the business tax burden rate, T_S is the business tax and surcharges, and R_B is the operating income. Based on the above calculations, the final established tax RW indicator system for SEMs is denoted in Table 2.

Table 2: SME tax risk warning indicator system

Primary indicator	Secondary indicator	Tertiary indicator
SME tax RW indicators	Solvency indicators	Current ratio
		Quick ratio
		Asset liability ratio
	Profitability indicators	Operating profit margin
		Net profit margin on assets
		Equity net profit margin
	Operating capability indicators	Inventory turnover ratio
		Accounts receivable turnover ratio
		Total asset turnover ratio
	Production and sales indicators	Operating cost ratio
		Period expense ratio
		Non operating expense rate
		Business tax burden rate

In Table 2, the secondary indicators of the system include solvency, profitability, operational capability, and production and sales indicators. The tertiary indicators further elaborate on the specific measurement methods for each secondary indicator. After determining the indicator system, the next step is to determine how to quantify and compare the relative importance of these indicators to obtain more accurate and effective risk assessment results. AHP is a systematic and hierarchical decision analysis method, and it quantitatively compares the importance of these sub problems to obtain overall priority ranking and decision results. It has the advantages of clear structure, quantitative analysis, and strong flexibility, and is broadly applied with significant effects. Therefore, the study uses AHP to construct the weights of tax RW indicators, and obtains the weight distribution of each indicator through maximum eigenvalue calculation and normalization processing, as shown in Table 3.

Table 3: Distribution of weights for SME tax risk warning indicators

Primary indicator	Secondary indicator	Secondary weight	Tertiary indicator	Tertiary weight
SME tax risk warning indicators	Solvency indicators	0.1534	Current ratio	0.1383
			Quick ratio	0.6232
			Asset liability ratio	0.2385
	Profitability indicators	0.2252	Operating profit margin	0.6194
			Net profit margin on assets	0.0964
			Equity net profit margin	0.2842
	Operating capability indicators	0.0716	Inventory turnover ratio	0.6651
			Accounts receivable turnover ratio	0.2311
			Total asset turnover ratio	0.1038
	Production and sales indicators	0.5498	Operating cost ratio	0.5601
			Period expense ratio	0.2268
			Non operating expense rate	0.0835
			Business tax burden rate	0.1296

Table 3 shows the weight distribution of tax RW indicators for SEMs, including primary indicators, secondary indicators, and their corresponding tertiary indicators and weights. Through weight allocation, it can be seen which indicators are more important in tax RW, thus providing more scientific and accurate risk management guidance for enterprises.

2.2 Construction of a risk warning model integrating auto encoder and BP neural network

On the basis of the previous section, this section will further design an early warning model. BPNN can achieve high-precision prediction in complex nonlinear mappings through multi-layer perceptron structure and backpropagation algorithm, and has strong self-learning and adaptive capabilities. At the same time, in the face of partial data loss or large noise, BPNN can still maintain good prediction performance and has strong robustness [15]. Therefore, the study selects BPNN as the main architecture of the warning model, and its basic structure is shown in Figure 2.

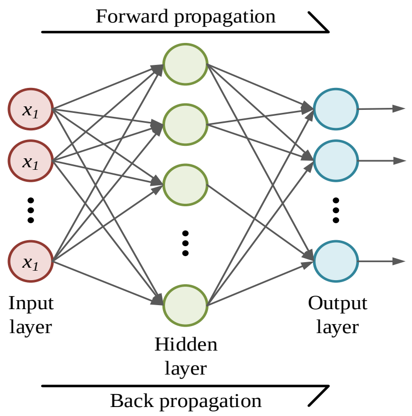


Figure 2: BPNN structure

As shown in Figure 2, the BPNN contains an input layer, a hidden layer, and an output layer. The input layer receives external data input, which is passed through multiple nodes to the hidden layer. The hidden layer contains one or more layers of neurons that perform nonlinear transformations on input data through activation functions to extract deep features of the data. The connection weights between hidden layers and between hidden layers and output layers will be continuously adjusted according to the error backpropagation algorithm to minimize prediction errors. During the training process, data is passed from the input layer to the output layer through various hidden layers, and the output value is calculated and compared with the true value to obtain the error. Then, the error propagates back layer by layer from the output layer through the backpropagation algorithm, adjusting the connection weights between each layer. This process is repeated until the error converges to a smaller value. Among them, the input functions of each output layer neuron are shown in equation (5) [16,17].

$$C_i = \sum_{j=1}^{q} V_{ji} S_j - \lambda_i \quad (5)$$

In equation (5), C_i represents the input value of the i th output layer neuron. V_{ji} means the weight from the j th input layer neuron to the i th output layer neuron, and S_j means the output signal of the j th input layer neuron. q means the number of neurons in the input layer, and λ_i represents the bias value. Subsequently, the sum function of the output layer neurons is shown in equation (6) [18].

$$L_i = \frac{1}{1+\exp(-\sum_{j=1}^{q} V_{ji} S_j + \lambda_i)} \quad (6)$$

In equation (6), $\exp$ is an exponential function, and $\sum_{j=1}^{q} V_{ji} S_j + \lambda_i$ represents the weighted input value of the output layer neurons. Finally, the Sigmoid activation function expression used is shown in equation (7).

$$\sigma(x) = \frac{1}{1+\exp(-x)} \quad (7)$$

In equation (7), the Sigmoid activation function can map any real number to $(0,1)$, and the processed output value L_i represents the final output result of the i th output layer neuron. However, BPNNs also face some challenges in practical applications, such as the need for a large amount of sample data, insufficient scientific selection of initial parameters, and stability issues in model prediction results. Therefore, the study introduces Auto Encoder for feature extraction and dimensionality reduction preprocessing of indicators in tax RW models, and then passes it to the BPNN model. The BPNN will train and predict based on these simplified data. Auto Encoder, as an unsupervised learning algorithm, can effectively extract features from high-dimensional data by encoding and decoding data, reducing data dimensionality while improving data quality [19]. Therefore, the AEBP structure that integrates Auto Encoder and BP is denoted in Figure 3.

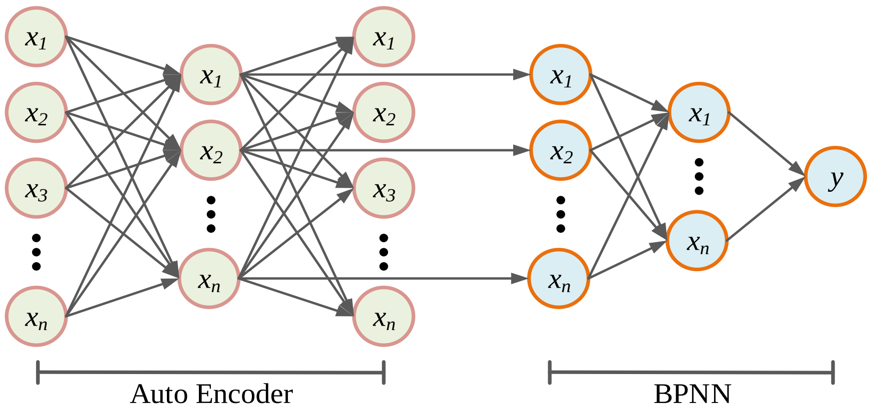


Figure 3: Network structure of AEBP

As shown in Figure 3, AEBP consists of two parts: Auto Encoder and BPNN. Firstly, the input layer receives n input variables $(x_1, x_2, ..., x_n)$, which are processed through Auto Encoder. Auto Encoder performs dimensionality reduction and feature extraction on input data, thereby reducing data redundancy and noise and improving data effectiveness. Secondly, the encoder compresses high-dimensional data into a low dimensional feature space $(z_1, z_2, ..., z_m)$ to extract key features of the data. Subsequently, these features are passed on to the BPNN for further processing. When using Auto Encoder for dimensionality reduction and feature extraction, first, the selection of input features is based on the previously established indicator system, and the most representative features are screened out after expert consultation and correlation analysis. The input layer of the Auto Encoder accepts these high-dimensional features and compresses them into a low-dimensional space through the encoder, extracting key features to reduce data redundancy. During the dimensionality reduction process, the input layer of the Auto Encoder processes high-dimensional data containing 100 original features and finally compresses them into 20 key features, successfully achieving a dimensionality reduction effect from 100 dimensions to 20 dimensions. This dimensionality reduction not only effectively reduces the computational complexity, but also improves the generalization ability and prediction accuracy of the model.

The BPNN receives features output from the Auto Encoder as input, and trains and predicts them through a multi-layer perceptron structure and backpropagation algorithm. However, the initial parameters of BPNNs, such as network layers, amount of neurons in each layer, weights, and learning rate, are usually selected based on experience or automatically assigned initial values, lacking scientific and targeted approaches, which may lead to model instability and further affect prediction results. GA is a global search optimization algorithm that can effectively optimize the parameter settings of BPNNs, such as weights, biases, learning rates, etc. [20]. Through iterative optimization of GAs, better parameter combinations can be found to raise the predictive effectiveness and accuracy of the model. The process of introducing GA to optimize the parameters of AEBP model is shown in Figure 4.

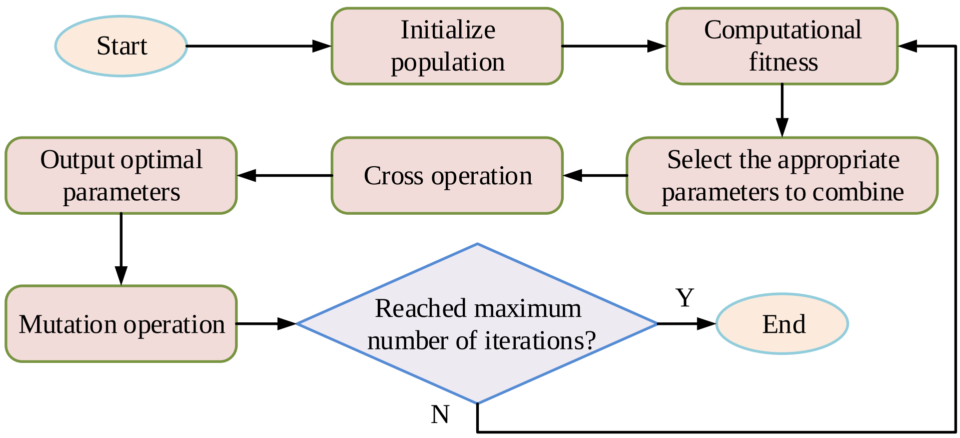


Figure 4: Process of introducing GA to optimize AEBP model parameters

In Figure 4, first is to initialize the population, randomly generate an initial parameter set, and set the max amount of iterations. Next is to calculate the fitness of the current AEBP model, that is, evaluate the performance of the current parameter combination in model prediction. Subsequently, it selects parameter combinations with higher fitness as the basis for generating the next generation parameter set. Then it performs a crossover operation on the selected parameter set to generate a new parameter combination. The next step is to perform mutation operations, which involves randomly adjusting certain parameters in the parameter set to introduce diversity and explore new parameter spaces. Then, if the max amount of iterations is reached or the area under curve (AUC) of the model changes less than the set threshold, the optimization is stopped and the current optimal AEBP model parameters are output. Through this parameter selection strategy, GA can continuously optimize the parameters of AEBP to raise the predictive performance of the model.

3 Results

To verify the performance of the AEBP-based tax RW model for SEMs, a suitable experimental environment was established, and financial and tax data of SEMs were preprocessed. The first section conducted performance testing on the AEBP warning model. The second section conducted simulation tests in practical application scenarios to verify the actual effectiveness of the system in different economic environments and diverse enterprise data.

3.1 Performance testing of AEBP model

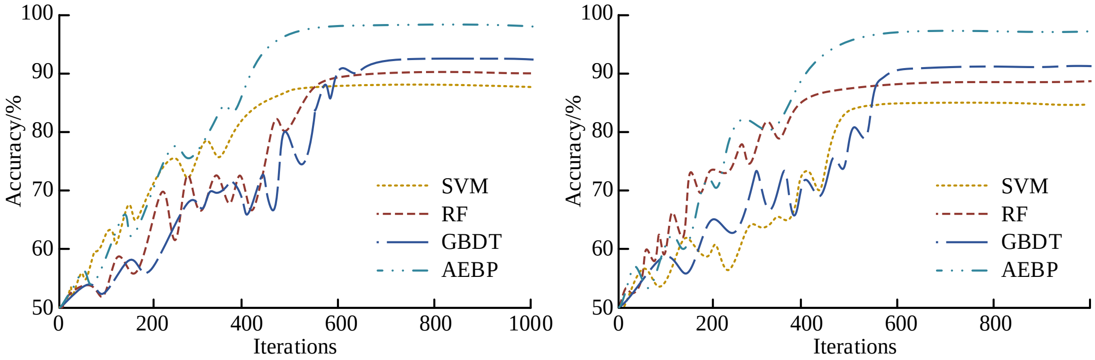
The operating system used in the study is Windows 10, Intel Core i7-4710MQ processor CPU, and 8.00GB of memory. The SME Credit Scoring dataset in Kaggle was used, which contained credit score data for SEMs, including financial status, credit scores, loan records, and other information, with a total of approximately 280000 samples and 31 feature numbers. After cleaning and preprocessing the dataset, it was broken into a test set and a training set in an 8:2 ratio. SVM, Random Forest (RF), and Gradient Boosting Decision Tree (GBDT) were selected as comparison models. The main reason is that they are widely used in machine learning and are recognized as benchmark models, especially in classification tasks, and are suitable as comparison benchmarks for new models. These three models represent different learning algorithm frameworks and can comprehensively evaluate the performance of the AEBP model. Compared with more complex models such as XGBoost or deep learning models (LSTMs), SVM, RF, and GBDT have advantages in computational efficiency and interpretability, and are more suitable for rapid experiments and verification on large-scale data sets. In addition, these models have shown good performance in classification problems similar to tax risks and have practical reference value. Taking into account the complexity of the model, the demand for computing resources, and interpretability, the study selected models that are easier to implement and can provide reliable results, laying a more efficient comparison foundation for future research.

In the study, GA was used to optimize the hyperparameter settings of the AEBP model, including the number of hidden layer neurons, learning rate, and dropout rate. To effectively promote the transmission of excellent genes and improve the convergence speed of the algorithm, GA adopted a crossover rate of 0.8. In addition, to increase the diversity of the population and prevent the algorithm from falling into a local optimal solution, the mutation rate was set to 0.1, and the mutation operation was achieved by randomly adjusting the genes of individuals. Finally, the population size of GA was set to 50. Firstly, the optimal parameters for AEBP need to be determined based on GA, and the test outcomes are indicated in Table 4 below.

Table 4: The optimal parameters of AEBP model

Parameter	Description
Amount of neurons in Auto Encoder hidden layer	8
Amount of neurons in BP hidden layer	11
Dropout rate of Auto Encoder	0.048
Learning rate of BPNN	0.024
Activation function of BPNN	sigmoid
Activation function of Auto Encoder	tanh
Optimizer	SGD

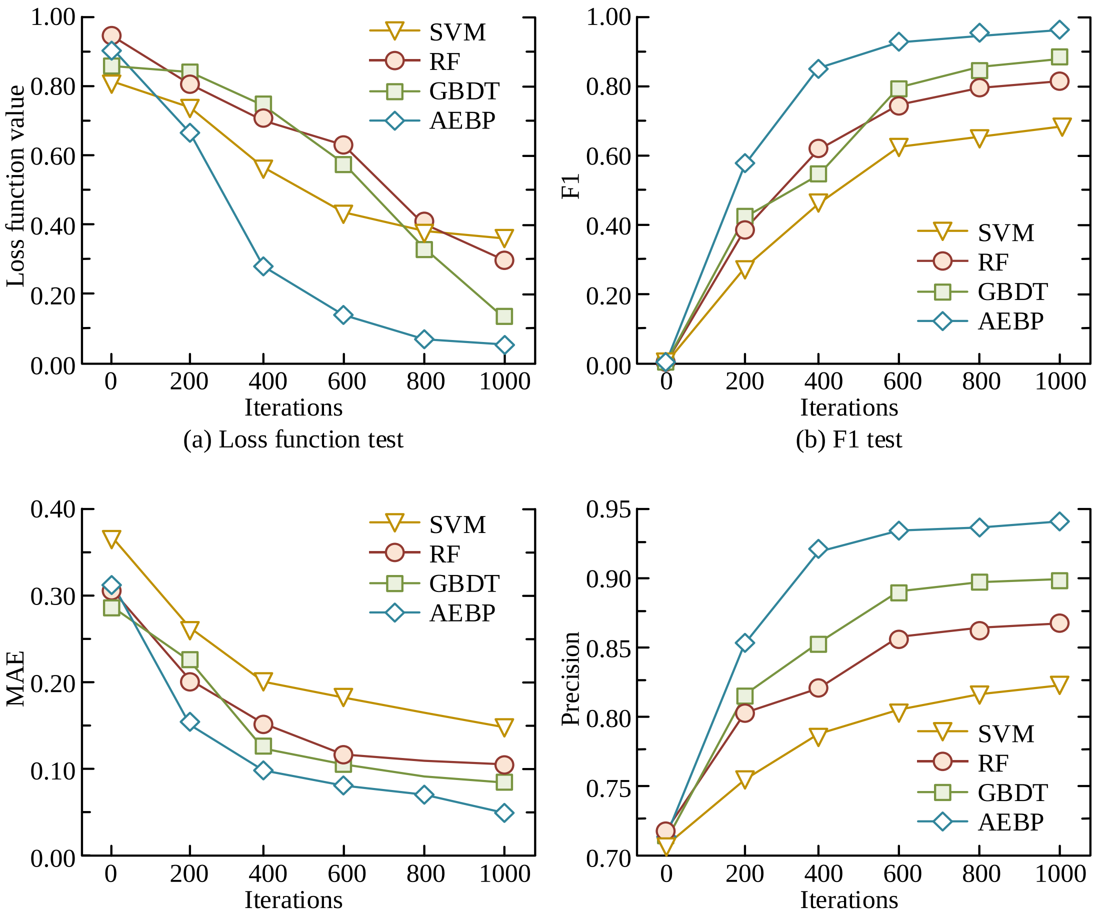
In Table 4, the hidden layer of the Auto Encoder was configured with 8 neurons, which was slightly increased compared to the original setting, to better capture the underlying features of the input data. The dropout rate was set to 0.048 to ensure a balance between preventing overfitting and preserving important information. The learning rate of the BPNN was adjusted to 0.024, making the convergence during training more stable and gradual. The hidden layer contained 11 neurons, which could enhance the model's ability to learn complex patterns while avoiding overfitting after optimization. The choice of sigmoid activation function was due to its simplicity and significant performance in binary classification tasks. The choice of tanh activation function was due to its ability to handle nonlinear relationships and prevent saturation problems. Therefore, the study used the above parameters in the AEBP algorithm for subsequent testing. Subsequently, with accuracy as the indicator, the test findings of each model are denoted in Figure 5.



(a) Training set (b) Test set

Figure 5: Comparison of performance of different algorithms

Figures 5 (a) and 5 (b) show the classification accuracy test results of SVM, RF, GBDT, and AEBP models on the training and testing sets, respectively. As the amount of iterations increased, the accuracy of each model gradually increased and eventually stabilized at a certain level. In the training set of Figure 5 (a), when the amount of iterations was 1000, the accuracy of SVM, RF, GBDT, and AEBP models was 87.6%, 90.01%, 92.4%, and 98.3%, respectively. In the test set of Figure 5 (b), when the amount of iterations was 1000, the accuracy of the four models was 84.3%, 98.8%, 91.1%, and 97.6%, respectively. The curve of AEBP not only exhibited small oscillations in the early stages of iteration, but also quickly converged to find the global optimal value. Finally, the comprehensive indicator test results are shown in Figure 6.



(a) Loss function test (b) F1 test

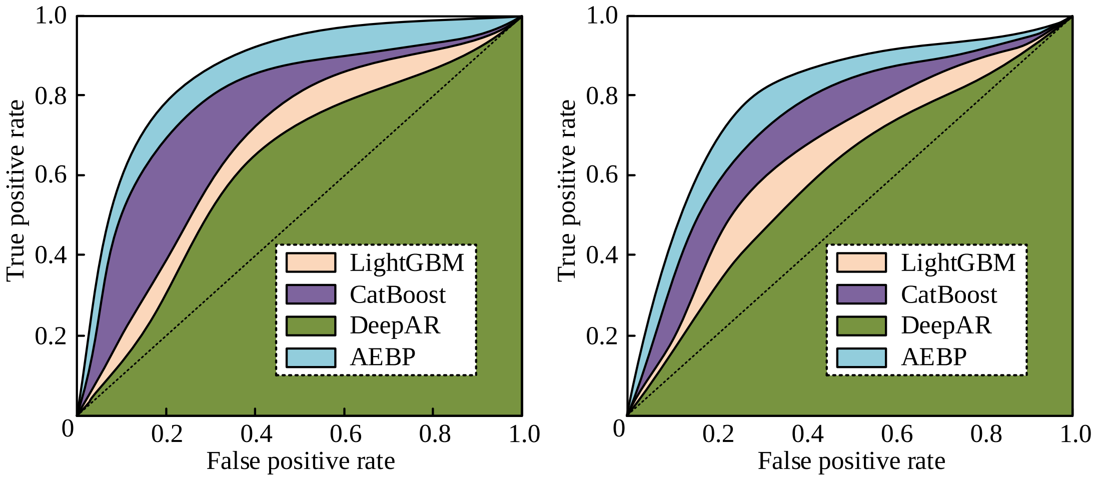
(c) MAE test (d) Precision test

Figure 6: Multi-index performance testing

Figures 6 (a), 6 (b), 6 (c), and 6 (d) showcase the loss function values, F1 scores, Mean Absolute Error (MAE), and precision test results of SVM, RF, GBDT, and AEBP models on the training set, respectively. In Figure 6 (a), when the amount of iterations reached 1000, the loss function values of the four models were 0.37, 0.30, 0.14, and 0.05, respectively. In Figure 6 (b), when the number of iterations reached 1000, the F1 scores of the four models were 0.69, 0.81, 0.88, and 0.96, respectively. In Figure 6 (c), when the amount of iterations reached 1000, the MAE values of the four models were 0.15, 0.11, 0.08, and 0.05, respectively. In Figure 6 (d), when the number of iterations reached 1000, the precisions of the four models were 0.82, 0.87, 0.90, and 0.94, respectively. From this, the AEBP built by the research performed well in terms of loss function value, F1 score, MAE, and precision, outperforming other comparative models.

3.2 Application experiment of AEBP-based tax risk warning model for SMEs

Study applied the AEBP in tax warning scenarios for SEMs to verify its practical effectiveness. SEMs listed on the board were selected, and the quarterly indicator data of SEMs from 2004 to 2020 was used as the research sample, including about 2500 samples with tax risk issues and 12000 samples without tax risk issues. The selected data were all from the Wind database. After extracting the tax reports of each enterprise, oversampling technology was introduced to balance the sample size. The balanced data consisted of 8000 non tax risk samples and 5000 tax risk samples, divided into training and testing sets in an 8:2 ratio. Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Deep Autoregressive (DeepAR) were selected as comparison models. Firstly, the receiver operating characteristic curve (ROC) results of each algorithm are shown in Figure 7.



(a) Training set

(b) Test set

Figure 7: ROC curve test results

Figures 7 (a) and 7 (b) show the ROC of LightGBM, CatBoost, DeepAR, and AEBP models on the training and testing sets, respectively. The horizontal and vertical axes mean false positive rate and true positive rate, respectively. The larger the value of the AUC enclosed by the ROC and the horizontal and vertical axes, the stronger the model's ability to distinguish between samples with and without tax risk. In the training set of Figure 7 (a), the AUC values of the four models were 0.73, 0.82, 0.62, and 0.89, respectively. On the test set in Figure 7 (b), the AUC values of the four models were 0.71, 0.80, 0.61, and 0.87, respectively. Subsequently, the accuracy of each model in warning tax risks of SEMs in different industries is shown in Table 5.

Table 5: Tax risk early warning success rate for different industries by various models

Industry category	Detection accuracy/%			
	LightGBM	CatBoost	DeepAR	AEBP
Leasing and business services	78.45	80.35	75.89	85.47
Accommodation and catering	72.58	74.92	70.45	82.34
Manufacturing	85.36	88.29	83.56	90.12
Information transmission, software, and information technology services	81.47	84.33	80.21	86.79
Culture, sports, and entertainment	74.36	77.85	72.41	80.67
Health and social work	80.28	83.19	79.56	85.43
Agriculture, forestry, animal husbandry, and fishery	78.22	81.45	76.98	84.15
Transportation, storage, and postal services	79.34	82.78	78.11	85.02

Education	82.19	85.23	81.46	89.34
Scientific research and technical services	83.47	86.45	82.98	89.76
Electricity, heat, gas, and water production and supply	80.56	83.67	79.23	86.58
Construction	75.68	78.34	73.89	89.23
Real estate	76.34	79.56	75.02	87.45
Water conservancy, environment, and public facilities management	78.89	82.12	77.33	84.78
Wholesale and retail trade	79.45	82.34	78.19	85.56
Average detection accuracy	79.11	82.06	77.68	86.18

Table 5 shows the results of tax RW for different industries using four models: LightGBM, CatBoost, DeepAR, and AEBP. In the table, the AEBP model consistently outperformed other models in various industries. For example, in the manufacturing industry with the largest market share in China, the warning success rate of AEBP reached 90.12%, significantly higher than LightGBM's 85.36%, CatBoost's 88.29%, and DeepAR's 83.56%. The average accuracy of tax RW for all industries by each model was 79.11%, 82.06%, 77.68%, and 86.18%, respectively. The comprehensive performance of AEBP was the best among all compared models. Finally, the importance scores of the indicators that affect the tax risk prediction results output by AEBP are shown in Figure 8.

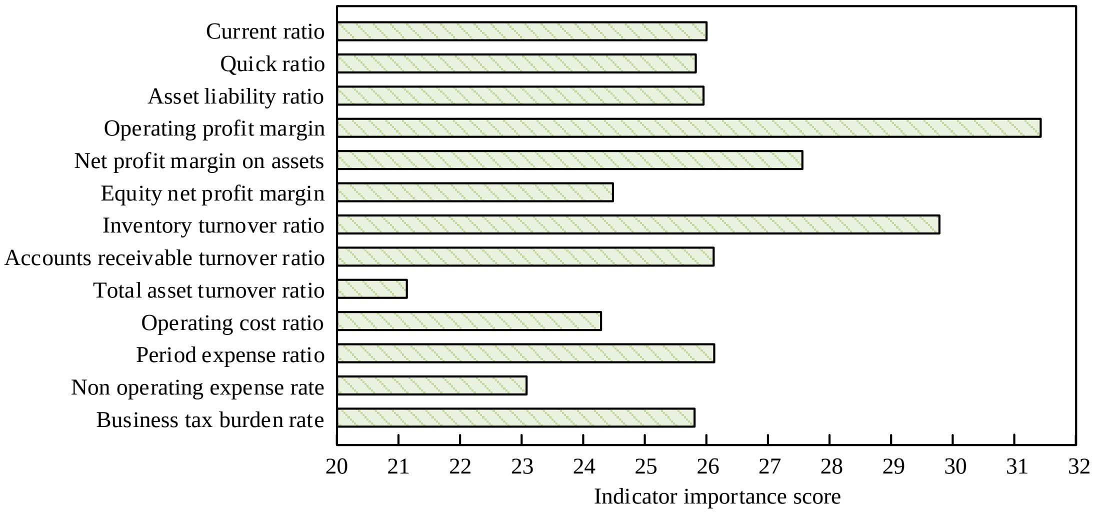


Figure 8: Index importance score test results

Figure 8 showcases the importance scores of each indicator for the tax risk prediction results. As shown in the figure, the operating profit margin scored the highest among all indicators, reaching 31.47, indicating the most significant impact on predicting tax risks. The inventory turnover rate and net profit margin of assets closely followed, with values of 29.74 and 27.61 respectively, indicating the important role of inventory turnover ability and market expansion ability in forecasting for enterprises. The total asset turnover rate and non operating expenses scored relatively low, at 21.14 and 23.12, respectively. Overall, profitability and growth indicators have a more significant impact on tax risk prediction, providing reference for corporate tax management. From the figure, specific financial indicators, such as cash flow and accounts receivable, had a higher weight in the model's decision-making process. This showed that the model was not only effective in risk prediction, but also provided insights into which features were most critical in prediction. This visualization of feature importance greatly enhanced the interpretability of the model. Compared with traditional "black box" models, the AEBP model enabled users to understand the model's prediction logic by showing the contribution of key features. This transparency not only improved the trustworthiness of the model, but also provided a substantive basis for corporate managers to make decisions in tax risk control. Through these analyses, the AEBP model not only provided high-precision prediction results, but also overcame the limitations of traditional models in interpretability, providing strong support for more transparent and insightful risk management. Finally, to evaluate the scalability and computational complexity of the AEBP model in practical applications, the study will further expand the data set size. The results of the comparative tests of each model under different data scales are shown in Table 6 below.

Table 6: Test results under different data scales

Dataset size	Model	Training time/s	Prediction time (seconds/1000 samples)	Memory usage/MB	Accuracy/%	F1 score
10,000	AEBP	123	0.83	135	92.51	0.91
	SVM	81	0.52	119	88.32	0.87
	RF	102	0.64	143	90.25	0.89
	GBDT	149	0.91	158	91.03	0.9
50,000	AEBP	597	3.48	623	93.02	0.92
	SVM	401	2.53	607	89.04	0.88
	RF	506	2.77	705	91.09	0.9
	GBDT	799	4.07	803	91.52	0.91
100,000	AEBP	1203	7.08	1221	93.54	0.93
	SVM	901	5.03	1202	89.52	0.89
	RF	1003	5.48	1394	91.51	0.91
	GBDT	1598	7.97	1593	92.03	0.92

As shown in Table 6, the AEBP model exhibited high computational efficiency and prediction performance under different data scales. On a dataset of 10,000 samples, the training time of the AEBP model was 123s, the prediction time was 0.83s/1000 samples, the memory usage was 135MB, the accuracy was 92.51%, and the F1 score was 0.91. In contrast, the accuracy of the SVM model under the same conditions was 88.32%, the F1 score was 0.87, and its training time was 81 s. This showed that the AEBP model achieved an effective balance between resource consumption and performance. When the data scale increased to 100,000 samples, the training time of the AEBP model increased to 1203s, the memory usage was 1221MB, but the accuracy remained at 93.54%, and the F1 score was 0.93, that is, the model had good scalability when processing large-scale datasets. In contrast, the GBDT model had the highest memory usage of 1593MB, but its accuracy and F1 score were slightly lower than those of the AEBP model. The results show that the AEBP model can still provide high prediction performance and low resource consumption under the condition of limited computing resources, and is an ideal choice for SEMs to conduct real-time tax risk monitoring.

4 Discussion

For the current tax risk early warning system for SEMs, traditional models such as SVM, RF and GBDT usually face problems such as limited model generalization capabilities and insufficient feature extraction accuracy in complex economic environments. This study proposed a model that combines Auto Encoder and BPNN, and optimized the parameters through GA to improve the prediction accuracy and stability of the model. The performance test results showed that the best model after GA optimization had a number of hidden layer neurons of 8, a hidden layer of 11, a dropout rate of 00.48, and a learning rate of 0.024. When the number of iterations reached 1000, the accuracy of AEBP on the training set was 98.3%, and the loss function value, F1, MAE, and precision were 0.05, 0.96, 0.05, and 0.94 respectively, which was significantly better than the traditional model. This was because the Auto Encoder performed well in data dimensionality reduction and feature extraction, effectively reducing data redundancy and noise, thus providing more refined input data for BPNN. The optimized BPNN showed more performance in complex nonlinear mapping. Strong predictive ability.

In the performance test of the AEBP model, the slight oscillation phenomenon in the initial iteration process reflected the sensitivity of BPNN in the initial parameter settings, which may lead to instability in the early stages of training. However, as the number of iterations increased, the model gradually converged to the global optimal value, the oscillation phenomenon gradually weakened, and the model showed good convergence and stability. In the simulation test, the AUC of AEBP in the training set was 0.89. In the manufacturing industry with the largest market share, AEBP's tax risk early warning success rate was 90.12%, and the average early warning success rate was 86.18%. In terms of indicator importance score, operating profit margin had the highest score of 31.47. Finally, the model had excellent early warning accuracy for tax risks in different industries, and could effectively help companies identify potential risks in advance.

It is worth noting that the AUC curve and the detection success rates of different industries are of great significance in the research. The AUC curve reflects the model's ability to distinguish positive and negative samples. The AEBP model maintained high AUC values in various industries, indicating its robustness and consistency in different scenarios. Especially compared with the "black box" model, the AEBP model was not only superior in performance, but also showed significant advantages in interpretability and reliability. By showing the detection success rate in various industries, this study proved that the AEBP model was highly adaptable and practical in different industries, which is crucial for decision support in tax risk management. The AEBP model not only provided high-precision prediction results, but also explained its decision-making process through a transparent model structure, avoiding common interpretive flaws in "black box" models. This interpretability and reliability made the AEBP model more trustworthy and operable in practical applications. Especially in scenarios that require in-depth analysis and understanding of the results, the advantages of the AEBP model were more obvious.

In summary, this study proposed an efficient and accurate tax risk early warning model for SEMs by combining Auto Encoders and BPNN and introducing GAs to optimize parameters. Compared with existing traditional methods, the AEBP model showed significant advantages in many aspects, especially in terms of classification accuracy, resource utilization, and computational efficiency. This not only provided reliable technical support for the tax management of SEMs, but also provided new ideas and directions for future tax risk control and decision-making optimization in complex economic environments.

5 Conclusion

With the in-depth development of the global economy and the rapid advancement of information technology, SEMs play a key role in the national economy. However, the complexity and uncertainty in tax management bring many risks and challenges to SMEs. To deal with these challenges, this study constructed a tax risk indicator system suitable for SEMs through AHP. On this basis, it combined Auto Encoder with BPNN, introduced GA to optimize model parameters, and proposed a tax risk early warning model based on AEBP. Experimental results showed that the AEBP model performed well in terms of classification accuracy and model stability, and significantly improved the ability to identify and predict tax risks of SEMs. At the same time, through reasonable structural design, the model took into account the interpretability of the algorithm and the practicality of application, providing effective technical support for business managers and tax decision-makers. The research not only verified the potential of the AEBP model in practical applications, but also pointed out the direction for subsequent research. Future research can further expand the indicator system, explore more complex tax risk factors, and combine advanced optimization algorithms and computing technologies to improve the adaptability and prediction accuracy of the model to better serve tax risks in a dynamic economic environment manage.

Reference

[1] Pang Z, Xiao Y. Inventory control under corporate income tax and accrual accounting. IISE Transactions. 2023, 55(8):833-845. DOI:10.1080/24725854.2022.2128235.

[2] Qi Q, Li S, Zhang RQ. Optimal joint decisions of production and emission reduction considering firms' risk aversion and carbon tax rate. International Journal of Production Research. 2024, 62(4):1189-1205. DOI:10.1080/00207543.2023.2178833.

[3] Liu J, Yan L, Wang D. A hybrid blockchain model for trusted data of supply chain finance. Wireless personal communications. 202, 127(2): 919-943. DOI: 10.1007/s11277-021-08451-x.

[4] Li N, Feng J, Zhang C. Macro tax incentives and corporate sustainable innovation: Evidence from Chinese Enterprises. Environmental Science and Pollution Research. 2023, 30(45):101546-101564. DOI: 10.1007/s11356-023-29268-0.

[5] Sakalauskas E, Bendoraitis A, Lukšaitė D, Butkus G, Vitkutė-Adžgauskienė D. Tax declaration scheme using blockchain confidential transactions. Informatica. 2023, 34(3):603-616. DOI: 10.15388/23-INFOR531.

[6] Ding Q. Risk early warning management and intelligent real-time system of financial enterprises based on fuzzy theory. Journal of Intelligent & Fuzzy Systems. 2021, 40(4):6017-6027. DOI: 10.3233/JIFS-189441.

[7] Zeng H. Influences of mobile edge computing-based service preloading on the early-warning of financial risks. The Journal of Supercomputing. 2022, 78(9):11621-11639. DOI: 10.1007/s11227-022-04329-2.

[8] Li X, Wang J, Yang C. Risk prediction in financial management of listed companies based on optimized BP neural network under digital economy. Neural Computing and Applications. 2023, 35(3):2045-2058. DOI: 10.1007/s00521-022-07377-0.

[9] Yi X. Application of data mining in enterprise financial risk prediction based on genetic algorithm and linear adaptive optimization. Soft Computing. 2023, 27(14):10305-10315. DOI: 10.1007/s00500-023-08308-4.

[10] He J, Yang J. Network security situational level prediction based on a double-feedback Elman model. Informatica. 2022, 46(1):87-93. DOI: 10.31449/inf.v46i1.3775.

[11] Liu X, Liu Z, Ling M, Lin K, Chen P, Huang X, Zhai Y. Ergonomic Reliability Assessment of VDT Systemfor Operation Design Based on Improved BPNN and HCR under Special Circumstances. CMES-Computer Modeling in Engineering & Sciences. 2023, 136(1):685-707. DOI: 10.32604/cmes.2023.025058.

[12] Ma C, Liu S, Fu J, Zhang T. Test Oracle Generation Based on BPNN by Using the Values of Variables at Different Breakpoints for Programs. International Journal of Software Engineering and Knowledge Engineering. 2021, 31(10):1469-1494. DOI: 10.1142/S0218194021500492.

[13] Shen H. Assessment of financial risk pre-alarm mechanism based on financial ecosystem using BPNN and genetic algorithm. Soft Computing. 2023, 27(24):19265-19279. DOI: 10.1007/s00500-023-09317-z.

[14] Nguyen HH. Factors Affecting the Efficiency of Corporate Income Tax Management in Vietnam: Evidence from Mekong Delta Provinces. The Journal

of Asian Finance, Economics and Business. 2021, 8(2):175-182. DOI: 10.13106/jafeb.2021.vol8.no2.0175.

[15] Guo L, Yuan Y, Tang W, Xue H. Steady-state characteristics prediction of marine towing cable with BPNN. Ships and Offshore Structures. 2024, 19(2):208-222. DOI: 10.1080/17445302.2022.2157138.

[16] Chen Q, Zhao H, Qiu H, Wang Q, Zeng D, Ye M. Time series analysis of rubella incidence in Chongqing, China using SARIMA and BPNN mathematical models. The Journal of Infection in Developing Countries. 2022, 6(8):1343-1350. DOI: 10.3855/jidc.16475.

[17] Aibinu MO, Moyo S. On Approximate Analytic Techniques for the Construction and Analysis Of Solutions Of Mathematical Models. Matrix Science Mathematic. 2023; 7(1): 58-62. DOI: 10.26480/msmk.01.2023.58.62

[18] Ma Z, Liu J, Li Y, Zhang H, Fang L. A BPNN-based ecologically extended input–output model for virtual water metabolism network management of Kazakhstan. Environmental Science and Pollution Research. 2023, 30(15):43752-43767. DOI: 10.1007/s11356-023-25280-6.

[19] Um KS, Kim NJ, Heo SW. Surrogate-based model using auto-encoder for optimising multi-band antennas. IET Microwaves, Antennas & Propagation. 2022, 16(11):725-732. DOI: 10.1049/mia2.12288.

[20] Fadel IA, Alsanabani H, Öz C, Kamal T, Iskefiyeli M, Abdien F. Hybrid fuzzy-genetic algorithm to automated discovery of prediction rules. Journal of Intelligent & Fuzzy Systems. 2021, 40(1):43-52. DOI: 10.3233/JIFS-182729.