Applied Mathematics, Modeling and Computer Simulation
C.-H. Chen et al. (Eds.)

doi:10.3233/ATDE230950

Company Failure Prediction Model of CWOA-BP Neural Network

Yaocheng WANG[a,1]
[a]*Shanghai High School International Division, Shanghai, China*

Abstract. It is important to detect failure companies in order to protect the financial market and investors.This study introduces a back propagation(BP) failure detection model to identify failure firms. In terms of solution accuracy, the empirical results document the superior measurement of CWOA with BP compared with standard BP.

Keywords. Company failure identification, BP neural network, whale optimization algorithm(WOA)

1. Introduction

In current years, several models have been designed to detect firm failure [1][2][3]. Compared with other methods, BP neural network is a better learning algorithm, as it can learn and store various input-output mapping linkages without identifying the actual mathematical equation. The article designs a modified whale optimization algorithm (WOA) incorporated with a BP neural network to indicate failure companies. The remaining paper is listed as follows. Section briefly introduce the arithmetic statement used. Section 3 report and analyze the data and experimental results. The conclusion are given in Section 4.

2. Arithmetic Statement

2.1. BP neural network

The BP neural network is a multi-layer feed forward model and discovers the nonlinear relationship of multiple I/O issues.The weights of the network are adjusted to minimize the square error of the network by the steepest descent method through back propagation. The BP neural network model with significant adaptive learning capability can be used to various training datasets[4]. Figure 1 depicts the learning process of BP neural network. The number of inputs can be used to determine the number of hidden layers and nodes in the BP neural network model for firm failure

1 Corresponding Author,Yaocheng WANG, Shanghai High School International Division, Shanghai, China; E-mail: albert_wang123@sina.com.

identification. The number of hidden layers is set to one according to the Kolmogorov theorem.

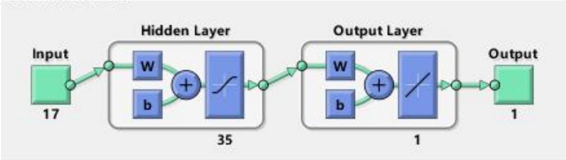


Figure 1. BP neural network structure

2.2. CWOA

2.2.1 Standard whale optimization algorithm(WOA)

The whale optimization algorithm (WOA), developed by Mirjalili and Lewis[5], primarily models the predation behavior of humpback whales.This optimization population-based algorithm generates a population of whales randomly at first. Then it simulates how these whales searching prey in the space, which can be used solving for optimal problems[6]. The whale uses Equations (1) and (2) to search randomly during the exploration stage.

$$\vec{D} = \left|\vec{C}.\vec{X}_{rand}{}^{*}(t) - \vec{X}(t)\right| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}_{rand}(t) - \vec{A}.\vec{D} \quad (2)$$

Encircling prey and bubble-net attacking methods are two types of predation behavior that occur during the exploitation phase. Equations (3) and (4) are mathematical function of shrinking encircling mechanism.

$$\vec{D} = \left|\vec{C}.\vec{X}^{*}(t) - \vec{X}(t)\right| \quad (3)$$

$$\vec{X}(t+1) = \vec{X}^{*}(t) - \vec{A}.\vec{D} \quad (4)$$

D is the distance between Whale $\vec{X}(\mathrm{t})$ and the best Whale $\vec{X}^{*}(\mathrm{t})$. A and C are coefficient vectors computed using Equations (5) and (6). In each iteration, a is linearly decreased from 2 to 0 and r is a random vector with a range of [0, 1].

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \quad (5)$$

$$\vec{C} = 2.\vec{r} \quad (6)$$

Equation (7) is mathematical function of bubble-net attacks mechanism. Spiral upgrade technique is used to compute the distance between the whales in position (X, Y) and the best whale position in (X^*, Y^*).

$$\vec{X}(t+1) = \vec{D}'.e^{bl}.\cos(2\pi l) + \vec{X}^{*}(t) \quad (7)$$

Where $\vec{D}' = \left|\vec{X}^{*}(\mathrm{t}) - \vec{X}(\mathrm{t})\right|$ is the distance between the whale to the best whale position. l is a random number with the range of [-1, 1] and b is a constant number.

The whales simultaneously employ the shrinking encircling and spiral mechanisms. The value of p can be a random number with the range of [0, 1] in Equation (8).

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A}.\vec{D} & \text{if p} < 0.5 \\ \vec{D}'.e^{bl}.\cos(2\pi l) + \vec{X}^*(t) & \text{if p} \geq 0.5 \end{cases} \tag{8}$$

2.2.2 Chaotic cubic mapping

The starting values usually have effects on the performance of most optimization algorithms. But the distribution of the starting values may be not good if random generation methods are used to generate the beginning values of the algorithms. Other than random generating, the initial values of WOA are set by the chaotic cubic mapping method[7].

$$z_{k+1} = \rho z_k (1 - z_k^2) \qquad \rho \in (0,1) \tag{9}$$

The cubic chaotic mapping methods is a type of nonlinear dynamic system with high complexity and random properties, which is sensitive to initial values and parameters and can generate seemingly irregular motion trajectories. The distribution of cubic mapping are shown in figure 2.

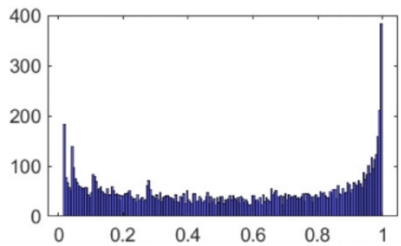


Figure 2. Chaotic cubic mapping

2.3 CWOA-BP Model

The chaotic whale optimization algorithm (CWOA), in addition to setting the random initial weights of BP neural networks, can address the issues of low accuracy.The major difference between CWOA algorithm and standard WOA is that it uses chaotic cubic mapping to generate random number sequences instead of the pseudo random number sequences used in standard WOA, in order to enhance the randomness and diversity of the algorithm.Cubic mapping utilizes the stochastic properties of chaotic mapping to improve the global search ability of WOA. One advantage of CWOA algorithm is that it may avoid the problems of premature convergence and fall into local optimal that occur in standard WOA, thereby improving the algorithm's global search ability and optimization accuracy[8]. Meanwhile, due to the use of chaotic mapping to generate number sequences, the CWOA algorithm can also increase the randomness and diversity of the algorithm, thereby better exploring the search space. Algorithm 1 describes the pseudo code of CWOA-BP. The process of CWOA-BP:

(1) Preprocess the data and identify the topological structure of the BP network.

(2)Generate a random number sequence using chaotic cubic mapping to alter the starting position of each whales.

(3)Update the historical optimal position of whales based on their current position and historical optimal position.

(4) Identify if the convergence condition is met or not before entering the BP neural network operation.

(5) Update the global optimal position based on the historical optimal positions of all whales.

(6) Based on the updated position, continue iteration until the preset stop conditions are reached.

Algorithm 1 The Chaotic whale optimization algorithm with BP (CWOA-BP)

Initialization {
initialize the whales population $\mathbf{X}_i$ by chaotic cubic mapping
(i = 1, 2, 3…, n) *%(the weights of BP neural networks)*
compute the fitness of each search whale
$\mathbf{X}^*$ = the best search whale}
Main loop{ *%(WOA)*
while (t < maximum number of iterations)
for each search whale
update the position of the current search whale by the Eq. (8)
% (the exploitation phase)
or
select a random search whale ($\mathbf{X}_{rand}$)
update the position of the current search whale by the Eq. (2)
% (The exploration phase)
or
update the position of the current search whale by the Eq. (4)
% (The encircling prey method)
or
update the position of the current search whale by the Eq. (7)
% (The bubble-net attacking method)
end for
update $\mathbf{X}^*$if there is a better solution for the weights of BP neural networks
$t = t + 1$
end while}
return **X**

3. Data And Empirical Results

In this research, the data include 814 samples of listed firms in Chinese stock market. The company will be identified as a failure firm if it is coded as ST company or *ST company. In Chinese stock market, Special treatment (ST) indicate the firms with abnormal financial numbers or other abnormal conditions which may have the risk of delisting in near future. Our sample include 265 ST firms and 549 non-ST firms. ST firms are coded as 1 and Non-ST firms are coded as 2.The China Stock Market & Accounting Research Database (CSMAR) is used to collect all of the financial and corporate governance variables used in the study.Table 1 tabulates 17 input variables used in this study. The 17 independent variables include 12 financial variables and 5 corporate governance variables[1][3][9]. The independent variables are at year t-1 to predict the future status of firms.

Table 1. The definition of independent variables

Category	Independent variables
Profitability features	ROA:net income divided by total assets RETURN: yearly stock turn EPS: earning per share
Leverage features	LEV:total debt to total assets
Growth features	GSALE: sale growth ASSET: natural log of total assets
Efficiency features	ATURN: sales to total assets
Corporate governance features	Age: firm age SOE: firm ownership type INTERNAL: the type of firm's internal system AUDITOR: the type of firm's auditor firm OPINION: auditor's opinion

Table 2 show the descriptive analysis of all independent variables. The mean value of ROA is about -0.013, while EPS is -0.192. 156 samples are state-owned enterprises. 230 samples have not clean auditor opinion at year t-1.

Table 2. Summary statistics

Panel A: continuous variables

	Mean	Median	SD	N
$ROA_{(t-1)}$	-0.013	0.013	0.130	814
$RETURN_{(t-1)}$	0.078	-0.089	0.679	814
$EPS_{(t-1)}$	-0.192	0.071	0.159	814
LEV	0.251	0.228	0.185	814
GSALE	0.196	0.044	2.248	814
ASSET	21.890	21.801	0.929	814
ATURN	0.621	0.532	0.437	814
Age	11.296	10	6.156	814

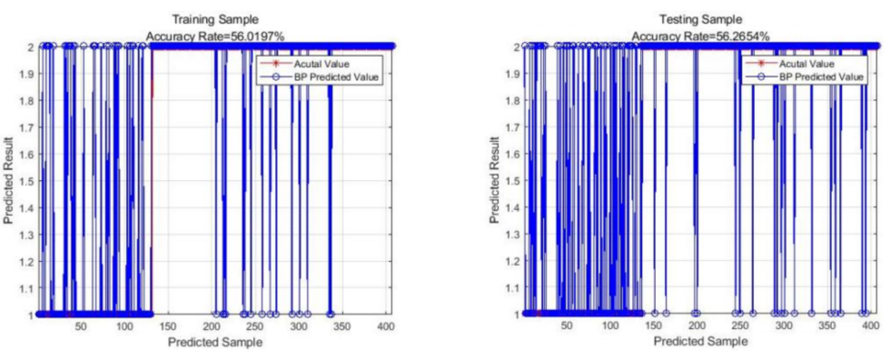
Panel B: Dummy variables

	Dummy variable=0	**Dummy variable=1**	**N**
SOE	658	156	814
INTERNAL	584	230	814
AUDITOR	794	20	814
OPINION	650	164	814

We split data equally into training and testing set after random disrupting full samples. Then the sample normalization pre-processing methods are used to correct scaling differences. The same data are used in models of BP and CWOA-BP to contract two methods. Figure 3 shows the performances of two methods in identifying BP parameters. The accuracy rate of CWOA-BP in training sample and testing sample are 58.4767% and 60.688% respectively, while the accuracy rate of BP in training sample and testing sample are 56.0197% and 56.2654%. It can be found that the CWOA algorithm outperforms the standard BP regarding to the average training and testing accuracy.

4. Conclusion

The identification of failure companies is a critical and challenging research topic in current years. Many different kinds of technology have been introduced to detect these firms in advance. This study presents a CWOA-BP model for analyzing firm's financial numbers and corporate information in previous years. The experimental findings indicate that the the CWOA-BP model is a reliable and efficient alternative in detecting failure companies. In the future, other input features such as board composition, compensation incentive and investment behavior might be included in the CWOA-BP model to further improve its effectiveness in analyzing firm's data.



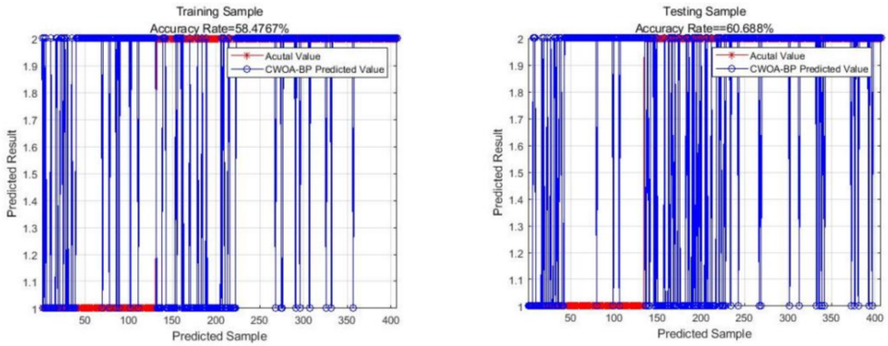


Figure 3. Algorithm prediction accuracy comparison

References

[1] Altman EI. Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy[J]. Journal of Finance, 1968, 23: 589- 609.

[2] Alam N, Junbin Gao J, Stewart Jones S.Corporate failure prediction: An evaluation of deep learning vs discrete hazard models[J]. Journal of International Financial Markets, Institutions and Money,2021,75, 101455.

[3] Chen J, Marshall BR, Zhang J, Ganesh S. Financial distress prediction in China[J]. Review of Pacific Basin Financial Markets and Policies, 2006, 9:317-336.

[4] Yan C, Li M, Wei L. Improved adaptive genetic algorithm for the vehicle insurance fraud identification model based on a BP neural network [J]. Theoretical computer science, 2019,7: 783-792.

[5] Mirjalili S Lewis A.The whale optimization algorithm[J].Advances in Engineering Software, 2016, 95: 51-67.

[6] Sun Y, Wang X, Chen Y, Liu J. A modified whale optimization algorithm for large-scale global optimization problems[J].Expert Systems with Applications,2018, 114: 563-577.

[7] Thomas D, Rogers D, Whitley C. Chaos in the cubic mapping[J]. Mathematical Modelling,1983, 4(1): 9-25.

[8] Huaijun Deng H, Liu L , Fang J , Qu B , Huang Q. A novel improved whale optimization algorithm for optimization problems with multi-strategy and hybrid algorithm[J].Mathematics and Computers in Simulation, 2023, 205:794-817.

[9] Zhou L, Lai KK, Yen J. Empirical models based on features ranking techniques for corporate financial distress prediction [J]. Computers & Mathematics with Applications, 2012, 64:2484-2496.