ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

# Comparative Study on Risk Early Warning Models of Capital Chain in Real Estate Industry

## -- Based on BP Neural Network Model and SVM

Jingxuan Liu, Yijun Chen and Yu Gan

Sichuan University of Science & Engineering, School of Management, Zigong 643000, Sichuan, China


## Abstract

**This paper takes the financial status of the capital chain of listed companies in the A-share real estate industry from 2016 to 2021 as the research object, constructs an early warning indicator system from four dimensions of capital chain investment chain, operation chain, return chain and non-financial indicators, using BP neural network , Support Vector Machine (SVM) to build early warning models respectively, and divide the early warning time into two groups of variables: 1-3 years short-term early warning and 3-5 years long-term early warning, and then choose a model that is more suitable for the real estate industry. . The experimental results show that the difference in prediction accuracy between the two models is small in short-term early warning, but in long-term early warning, the prediction accuracy of BP neural network model is significantly better than that of SVM.**



## Keywords

**BP Neural Network; Support Vector Machine; Capital Chain Risk Early Warning.**


## 1. Introduction

Since 2016, the real estate industry has been affected by policies such as 'three removals, one reduction and one supplement', 'no housing speculation', and 'three red lines'. . With the development of modern computer technology, it has become an important trend to introduce machine learning algorithms to build enterprise financial early warning models. In terms of enterprise financial management, machine learning algorithms can effectively avoid the linearization of data distribution requirements of traditional models, and the introduction of machine algorithms can avoid the judgment of financial crises by subjective opinions to a certain extent. Based on this, this paper will introduce machine learning algorithms - BP neural network and support vector machine (SVM) to build a capital chain risk early warning model for the real estate industry and compare the differences in the accuracy of the two algorithms in long-term and short-term early warning. The purpose is to find A machine learning algorithm model more suitable for real estate companies.

## 2. Literature Review

Initially, scholars used statistical methods to build financial early warning models, such as univariate discriminant, multivariate discriminant, and logistic regression. Fitzpatrick (1932) and Beaver (1966) first used the univariate discriminant to predict the company's financial status [1-2], but the univariate discriminant has higher requirements for the selected data, and the prediction results are unstable. Altman (1968) proposed the Z-score model [3], Zhou Shouhua (1996), Yang Shue (2003) improved the Z-score, and introduced the concept of cash flow to propose the F-score model and the Y-score model [4-5]. However, the discriminant

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

analysis method requires that the multivariate variables conform to the normal distribution, and the covariance must be known, which requires higher model construction. Ohlson (1980) constructed a Logistic model for bankrupt companies and non-bankrupt companies to discriminate financial status [6]. Dai Hongjun (2007), Luo Yi, Zheng Chunwei (2014), Wang Junping, Bai Qiongqiong (2015) adopted the logistic regression method for the research objects to build a financial early warning model [7-9].

With the development and popularization of computer technology, more scholars have tried to build enterprise financial risk early warning models through machine algorithms. Zhou Min and Wang Xinyu (2002) proposed an algorithm based on fuzzy optimization and neural network to measure financial crisis, avoiding the subjective judgment of financial crisis [10]. Adnan et al. (2006), Yang Shue, Wang Leping (2007), Lin (2009) made neural network model predictions for listed companies, and believed that neural network models would be more accurate in the medium and long term than statistical analysis methods [11-13]. Guan Xin and Wang Zheng (2016) found that the BP neural network has a higher prediction degree by comparing the accuracy of the logistic regression model and the neural network model for the first-class error and the second-class error [14].

## 3. Overview of the Theoretical Basis of The Model

### 3.1. Overview of BP Neural Network

BP neural network is a multi-layer feedforward neural network, which was proposed by D.E. Rumelhart, J.L. McCelland and PDP research group in 1986. BP neural network can be regarded as a nonlinear mapping between input and output sets, and this Mapping does not require complex internal relationship reactions, mainly through the signal in the forward transmission process, the input signal is processed layer by layer from the input layer to the hidden layer and then to the output layer. If the result of the output layer does not reach the expected value, the signal will be back-propagated to adjust the weight and threshold according to the error, so that the prediction result is constantly approaching the expected value [15]. The topology of the BP neural network is shown in Figure 1 below:

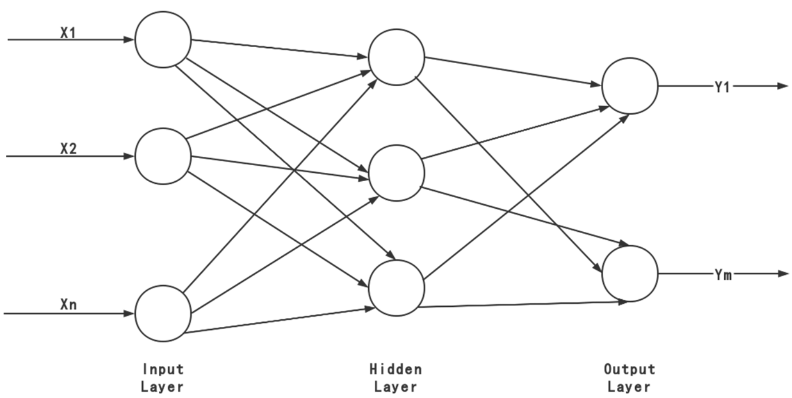


**Figure 1.** Topological structure of BP neural network

The basic algorithm of BP neural network: according to the sample T (x, y), determine the number of input layers n of the neural network, the number of hidden layer nodes l, and the number of output nodes m, and initialize the weights according to the input layer, hidden layer, and output layer, and The hidden layer threshold is calculated and the hidden layer output H is calculated.

$$H_j = f(\sum_{i=1}^{n} w_{ij}\, x_i - a_j)\; j=1,\ 2,\ 3......l \quad 3.1.1$$

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

According to the hidden layer output H, connect the weights $w_{jk}$ and the threshold b, calculate the predicted output O.

$$O_k=\sum_{j=1}^{l} H_j\, w_{jk} - b_k \quad k=1,\ 2,\ 3......m \qquad 3.1.2$$

According to the predicted value O and the expected output Y, calculate the error e and update the weight $w_{ij}$, $w_{jk}$

$$e_k = Y_k - O_k \qquad 3.1.3$$

$$w_{ij}=w_{ij}+\eta H_j(1-H_j)x(i)\sum_{k=1}^{m} w_{jk}\, e_k$$

$$i = 1,\ 2,\ 3 \ldots\ldots,\ n;\ j = 1,2,,\ \ldots\ldots,\ l \qquad 3.1.4$$

$$w_{jk}=w_{jk}+\eta H_j e_k \qquad j = 1,2,,\ \ldots\ldots,\ l\ k=1,\ 2,\ 3......m \qquad 3.1.5$$

In the formula, η is the learning efficiency

Update the network nodes and thresholds a, b according to the error e.

$$a_j = a_j++\eta H_j(1\text{-}H_j)\sum_{k=1}^{m} w_{jk}\sum_{k=1}^{m} w_{jk}\, e_k\ j = 1,2,,\ \ldots\ldots,\ l \qquad 3.1.6$$

$$b_k = b_k+e_k\ k=1,\ 2,\ 3......m \qquad 3.1.7$$

Finally, after each learning and training, it is judged whether the error accuracy meets the requirements. At the end of the algorithm iteration, the neural network learning ends to make predictions.

## 3.2. Overview of Nonlinear SVM

Support Vector Machine (SVM) is a machine algorithm proposed by Vapnik et al. Support vector machines are built on the basis of statistical learning theory and follow the principle of structural risk minimization. To a certain extent, it solves the problems of high-dimensional disaster and local convergence [15]. The support vector machine architecture diagram is shown in Figure 2:

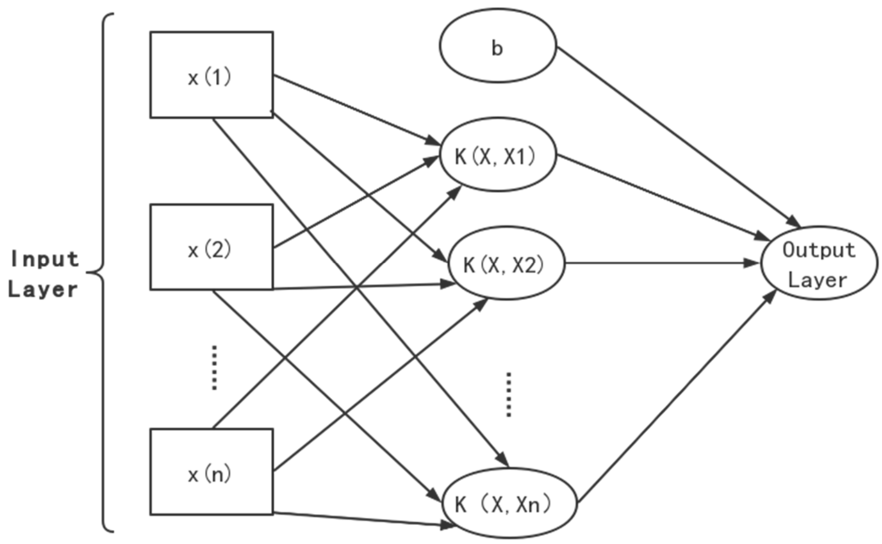


**Figure 2.** Support vector machine architecture diagram

The basic principle of SVM is based on the error rate of the learning machine during testing, with the error rate during testing and the number of terms of a VC dimension as the boundary, in the separable mode, the value of the previous term is made equal to zero, and the value of the second term is minimized. That is to construct an optimal classification hyperplane and use it as a decision surface to transform the sample into a secondary classification problem.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

Assuming that the sample data T={(xi,yi), i=1,2,3...,n}(xi∈$R_p$, yi∈$R_q$), there is an optimal hyperplane $w^T x+b=0$ to satisfy.

$$\min_{w,b}\frac{1}{2}||w|| = \min_{w,b}\frac{1}{2}W^TW \qquad 3.2.1$$

Introduce Lagrangian coefficients $\partial_i$ to find dual functions.

$$\min_{a}\frac{1}{2}\sum_{i=1}^{j}\sum_{j=1}^{i} y_i y_j \partial_i \partial_j K(x_i,x_j) - \sum_{j=1}^{i}\partial j \qquad 3.2.2$$

$$\text{s.t.} \sum_{j=1}^{i} y_i \partial_i = 0$$

Get the decision function.

$$f(x) = \text{sgn}(\sum_{j=1}^{i} y_i \partial_i K(x_i,x_j)+b^*) \qquad 3.2.3$$

$$s.t.\ b^*=y_i-\sum_{j=1}^{i} y_i \partial_i{}^* K(x_i - x_j) \qquad 3.2.4$$

# 4. Empirical Analysis

## 4.1. Sample Selection and Risk Level Division of Capital Chain

This paper selects 115 real estate companies listed on A-shares from 2016 to 2021, removes companies with missing data, and finally determines that the total number of sample data is 533. All data come from CSMAR database and Flush website. Considering that if the enterprise's capital chain status is judged based on whether the enterprise is marked with the 'ST' label, it may cause a large error in the prediction accuracy. ), Liao Ying and Wu Qinghe (2020) both took the net profit or adjusted net profit as an important indicator for judging the financial risk warning range of enterprises in the writing of the paper [16-17], and the 'Shenzhen Stock Exchange' The GEM Stock Listing Rules (Revised in 2020) point out that there is a risk of delisting when the net assets at the end of the most recent fiscal year are negative, so the net assets per share at the end of the period can also be used as a criterion [18].

When dividing the risk warning interval of enterprise capital chain, the risk status of capital chain can be divided into three categories: healthy capital chain, sub-healthy capital chain, and capital chain crisis through literature study and summary. Considering the particularity of the long operating cycle of the real estate industry, when dividing the early warning interval, the judgment period will be increased from two years to three years: if the company has three consecutive fiscal years in the past three years, net profit or net assets per share at the end of the period. In order to be regular, the company is defined as a healthy capital chain. If there is one year of net profit in the past three years or the net assets per share at the end of the period is negative, the company is defined as a sub-healthy capital chain. If the company has two years of net profit in the past three years. If the profit or net assets at the end of the period is negative, the enterprise is defined as a possible capital chain crisis.

## 4.2. Indicator Selection

This paper follows the principles of systematicness, comprehensiveness, effectiveness and operability in the selection of indicators. At the same time, considering that the indicators should fully reflect the characteristics of the capital chain of real estate enterprises during the selection process, the indicators are selected from the capital chain input chain, The selection of indicators from three angles of operation chain and return chain is shown in Table 1. At the same time, Jingxuan Liu (2022) found that adding non-financial indicators can improve the prediction accuracy of the model when building a Logistic financial early warning model for real estate companies [19]. Therefore, this paper adds a non-financial indicator 'auditor's opinion' to the original three financial perspectives. One of the advantages of introducing a machine learning algorithm is that the machine learning algorithm has a strong self-learning ability. It is not allowed to artificially spend a lot of time in finding and extracting eigenvalues

ISSN: 2637-6067
DOI: 10.6918/IJOSSER.202210_5(10).0012

and selecting indicators among many indicators. The above factors were subjected to principal component analysis.

**Table 1.** Early warning indicator system of capital chain

| Indicator classification | Indicator name | Formula |
|---|---|---|
| Capital chain investment chain | Cash ratio | (Monetary funds + trading financial assets)/current liabilities |
| | Interest coverage ratio | EBIT/expense |
| | Assets and liabilities | Total liabilities/total Assets |
| | Long-term capital demand guarantee rate | (Non-current liabilities + owners' equity)/non-current liabilities |
| | EBIT-Asset Ratio | EBIT/Assets |
| Capital chain operation chain | Roe | Net profit/average net assets |
| | Total asset growth rate | Growth rate of total assets = growth of total assets this year / total assets at the beginning of the year |
| | Net profit growth rate | (Total net profit for the reporting period-Total net profit for the base period)/total net profit for the base period |
| | Inventory turnover | Cost of goods sold / average inventory balance |
| | Total asset turnover | Total sales revenue/average total assets |
| | Operating profit margin | Operating profit/operating income |
| Capital chain return chain | Accounts receivable turnover | Net credit income/average accounts receivable |
| | Sales collection rate | (Operating Income + net accounts receivable) / operating Income |
| | Return on total assets | Net Profit/total Average assets |
| Non-financial indicators | Auditor opinion | Standard unqualified opinion = 0, otherwise = 1 |

## 4.3. Model Construction

The construction of the BP neural network model in this paper is mainly composed of an input layer, a hidden layer and an output layer. The weights of each layer are automatically assigned by the BP neural network model through learning, and there is currently no method for confirming the optimal number of nodes in the hidden layer. , so the optimal solution can only be found by trial and error every time. Finally, as shown in Figure 3, under the condition of 300 training times, the model accuracy is the highest and the training missing value is the lowest. The neural network structure is determined to be 15-60-3, and the output layer Y defines it as 3, where y=0 represents The company's capital chain is healthy, y=1 means the company's capital chain is sub-healthy, and y=2 means that the company has a capital chain crisis.

ISSN: 2637-6067
DOI: 10.6918/IJOSSER.202210_5(10).0012

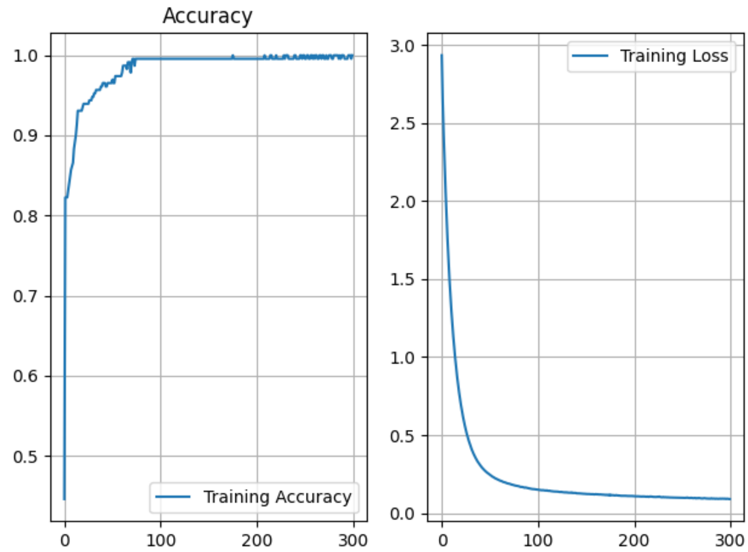


**Figure 3.** The relationship between the number of BP neural network training and the training missing values

In the process of building a model for SVM, this paper defines the dimension x as 15 in the sample T(x, y) according to the results of principal component analysis, and y is defined as an attribute category variable as 3. The assignment of y is the same as that of the BP neural network. The gray wolf algorithm is used to find the optimal solution penalty coefficient C and gamma value.

## 4.4. Validation Analysis

In order to ensure the accuracy of the experimental results, in the long-term early warning accuracy test, the two groups of models selected 2016-2018 financial data and non-financial data as the training group data in terms of data selection, and the test group data were 2019, 2020 and 2021. Years of data, the prediction accuracy of the model in the long-term is obtained by comparing the prediction results with the true value of the test group.

In the short-term early warning accuracy test, the two groups of models select 2019-2020 financial data and non-financial data as the training group data in terms of data selection, and the test group data is 2021. Rate.

The experimental results are shown in Table 2 Short-term early warning BP neural network and SVM prediction accuracy comparison and Table 3 short-term early warning BP neural network and SVM prediction accuracy comparison:

**Table 2.** Comparison of prediction accuracy between short-term early warning BP neural network and SVM

| | BP neural network | | | SVM | | |
|---|---|---|---|---|---|---|
| Year | Actual number | Correctly judged number | Accuracy rate | Actual number | Correctly judged number | Accuracy rate |
| 2019 | 115 | 101 | 88.2% | 115 | 99 | 86.1% |
| 2020 | 115 | 100 | 86.9% | 115 | 94 | 81.7% |
| 2021 | 115 | 90 | 78.2% | 115 | 84 | 73.0% |

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

**Table 3.** Comparison of prediction accuracy between short-term early warning BP neural network and SVM

| | BP neural network | | | SVM | | |
|---|---|---|---|---|---|---|
| Year | Actual number | Correctly judged number | Accuracy rate | Actual number | Correctly judged number | Accuracy rate |
| 2021 | 115 | 104 | 91.2% | 115 | 105 | 91.3% |

## 5. Conclusion

According to the empirical analysis results, comparing the early warning effects of the two models in the capital chain risk of real estate enterprises, it can be seen from Table 4-2 that in the short-term early warning accuracy test, the early warning effect of the BP neural network model is almost the same as that of the SVM, so it is used for short-term funds. For chain risk warning, you can choose one of the two models. But in the long-term warning accuracy test, Table 4-3 reports the judgment accuracy of the two models. From the results, the judgment accuracy rates of the two models show a downward trend year by year, but in comparison, the BP neural network model has a slower downward trend, and its correct judgment rate is also better than that of SVM in the same period. There are two possible reasons: First, in terms of the real estate policy environment, before 2020, the policy environment for the real estate industry is relatively loose, and corporate financial indicators change linearly every year. In terms of capital chain risk early warning, with the help of past financial indicators, a more accurate judgment can be made on the risk status of the capital chain of the enterprise in the next 2-5 years. Nowadays, the external market environment and policies are constantly changing, and the central supervision is intensifying, and most real estate companies' development strategies cannot achieve their expected goals, which weakens the ability of early warning models in long-term early warning. 2. In terms of the two machine learning algorithms, the BP neural network function relationship is more complex than the SVM. The BP neural network can automatically find the relationship between the input and output and optimize the function relationship according to the results of each training. In this paper, this paper finds that during 2014-2019, the overall development trend of the real estate industry is good, and only a few enterprises have major capital chain risks. 2' is less, and in 2019-2021, there will be capital chain risks in the real estate industry, and the number of eigenvalues 'y=2' will increase sharply. SVM is inferior to BP neural network in terms of self-learning ability and algorithm optimization, resulting in SVM in the long-term. The prediction performance is not as good as the BP neural network.

## References

[1] P.J Fitzpatrick.A Comparison of Ratio of Successful lndustrial Enterprises with those Failed Firms[J].Certified Public Accountant,1932:598-605

[2] Beaver Financial Rations as Predictors of FailureL] Supplement to Journal Accounting Research, 1966, (1): 71-111.

[3] Altman E.I.Financial Ratios Discriminate Analysis and Prediction of Corporate Bankruptcy[J]. The Journal of Finance,1968,(4):589-609

[4] Zhou Shouhua, Yang Jihua, Wang Ping. On the Early Warning Analysis of Financial Crisis—F Score Model [J]. Accounting Research, 1996 (08): 8-11.

[5] Yang Shu'e, Xu Weigang. Financial Early Warning Model of Listed Companies--An Empirical Study of the Y Score Model [J]. China Soft Science, 2003 (01): 56-60.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202210_5(10).0012

[6] Ohlson,J.S.Financial Ratios and the Probabilistic Prediction of Bankruptcy [J].Journal of Accounting Research,1980,(19):109-131.

[7] Dai Hongjun. Empirical research on early warning model of financial distress based on ternary Logistic [J]. Journal of Dalian University of Technology (Social Science Edition) 2007 (02): 60-66

[8] Luo Yi, Zheng Chunwei. Empirical Analysis of Financial Early Warning of Chinese Enterprises——Taking 23 Listed Companies of Financial Equipment in 2012 as an Example [J]. Finance and Economics, 2014 (02): 88-95.

[9] Wang Junping, Bai Qiongqiong. Research on financial crisis early warning of listed energy companies in my country [J]. Economic Issues, 2015 (01): 109-113.

[10] Zhou Min, Wang Xinyu. Early warning of enterprise financial crisis based on fuzzy optimization and neural network [J]. Journal of Management Science, 2002 (03): 86-90.

[11] Adnan, Aziz, Humayon. Predicting corporate bankruptcy: where we stand?[J].Corporate Governance The international journal ofbusiness in society,2006,6(1):18-33.

[12] Yang Shue, Wang Leping. Early warning of financial crisis of listed companies based on BP neural network and panel data [J]. System Engineering Theory and Practice, 2007 (02): 61-67.

[13] Lin T H. A cross model study of corporate financial distress prediction in Taiwan: Multiple di scriminant analysis, logit, probit and neural networks models[J]. Neurocomputing, 2009, 72(16) 3507-351.

[14] Guan Xin, Wang Zheng. Comparison of Financial Early Warning Models Based on Logistic Regression and BP Neural Network [J]. Statistics and Decision, 2016 (17): 179-181.

[15] Wang Xiaochuan, Shi Feng. et al. Analysis of 43 cases of matlab neural network [M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 2011: 18-25+128-136.

[16] Cen Hui. Research on early warning model of financial crisis of listed companies on ChiNext [D]. Jimei University, 2018.

[17] Liao Ying, Wu Qinghe. Internal governance, capital structure and early warning of financial crisis in small and medium-sized enterprises [J]. Cooperative Economy and Technology, 2021(02):90-94.DOI:10.13665/j.cnki.hzjjykj.2021.02. 036.

[18] 'Rules for the Listing of Stocks on the Growth Enterprise Market of the Shenzhen Stock Exchange (Revised in 2020)'[EB/OL].2020-6-12[2022-6-13].http://www.szse.cn/disclosure/notice/general/t20200612_578380.html

[19] Jingxuan Liu, Yijun Chen. An Empirical Study of Financial Early Warning in the Real Estate Industry-- Based on the Principal Component Logistic Regression Model[J]. World Scientific Research Journal,2022,8(1).