Study on Early Warning on the Financial Risk of Project Venture Capital through a Neural Network Model

Xianjuan Li
Hunan City University
Yiyang, Hunan 413000
China

Abstract—This paper aims to effectively reduce the financial loss of enterprises by accurately and reasonably making early warning of investment project risks. This paper briefly introduced the index system used for investment project risk early warning. It constructed a project investment risk early-warning model with a back-propagation neural network (BPNN) algorithm, and improved it with a genetic algorithm (GA) to solve the defect that the traditional BPNN is easy to fall into over-fitting when reversely adjust parameters. An analysis was conducted on an electric power company in Hunan Province. Orthogonal experiments are performed to determine the population size and the number of hidden layers in the improved BPNN algorithm. The results showed that the improved BPNN algorithm had the best performance when the population size was set as 25 and the number of hidden layers was four; compared with support vector machine (SVM) and traditional BPNN algorithms, the GA-improved BPNN algorithm had better performance for early risk warning of investment projects. In conclusion, adjusting the parameters of a BPNN with a GA in the training stage can effectively avoid falling into over-fitting, thus improving the early warning performance of the algorithm; in addition, the improved BPNN has better early warning performance.

Keywords—Neural network; project investment; early risk warning; genetic algorithm

I. INTRODUCTION

The rapid development of the economy has led to the emergence of various new enterprises, which further promotes the development of the economic market. In the process of enterprise development, the scale of an enterprise will be expanded, and during expansion, in addition to the existing profitable projects, the enterprise will also invest in other profitable projects [1] to further expand revenue and accelerate its development. However, there are few projects in the market that can make steady earnings, and more often than not, the projects available for investment carry different degrees of risk. Generally speaking, the higher the risk of an investment project is, the higher the ultimate return is, but the high risk of an investment project also means that the project is more likely to fail and lead to losses. The risk level of a project depends not only on the success probability of the project but also on the ability of enterprises to bear the losses after project failure [2]. When faced with the same project with a probability of failure, large enterprises that have more financial support than small enterprises can still operate normally even if the project fails, while small enterprises may not be able to operate normally, i.e., small enterprises will take more risks when facing the project. Therefore, before investing in a risky project, an enterprise needs to make an early warning assessment of project risks in conjunction with its financial situation to minimize the loss of the venture investment. The early warning assessment of a venture usually requires the appropriate professional knowledge of the assessor, but the managers of enterprises generally do not have the relevant professional knowledge [3]. Relying excessively on expert experience and subjective judgment when making decisions on a venture will seriously affect early risk warning. Therefore, enterprises need a relatively perfect project venture capital model to objectively warn the risk of investment projects and guarantee the smooth operation of the projects. This paper briefly introduced the index system used for investment project risk early warning, constructed a project investment risk early-warning model with a back-propagation neural network (BPNN) algorithm, optimized it with a genetic algorithm (GA), and analyzed an electric power company in Hunan Province. The novelty of this paper lies in the use of the GA to adjust the parameters in the BPNN, avoiding falling into over-fitting when adjusting the parameters reversely. The organization of this paper is introduction, related works, the construction of the neural network-based early financial warning model, example analysis, discussion, and conclusion.

II. RELATED WORKS

Zhu et al. [4] constructed a financial risk early-warning model based on the K-means clustering algorithm and found that the K-means clustering algorithm effectively avoided the negative subjective impact brought by artificially divided thresholds. Sun et al. [5] constructed a back-propagation neural network (BPNN)-based financial early-warning model, took mining listed companies as the research object, and found that the constructed early-warning model had a high prediction accuracy. Li et al. [6] introduced the L1 regularized support vector machine (L1-SVM) into the modeling of financial early warning systems as an effective feature selection technique and verified the feasibility of the technique in practical applications. Ouyang et al. [7] proposed a long short-term memory (LSTM) neural network under an attention mechanism for early warning of financial market risks. The final experimental...
results showed that this neural network had good generalization ability and higher prediction accuracy compared with BPNN, support vector regression (SVR), and autoregressive integrated moving average (ARIMA) models. Qu et al. [8] put forward an improved kernel principle component analysis-based financial risk prewarning model for public hospitals, conducted experiments on the financial data of multiple public hospitals and listed companies, and verified the feasibility and effectiveness of the method. Ding [9] proposed to establish a fuzzy theory-based early risk warning management and intelligent real-time monitoring model system for financial enterprises and analyzed a listed company engaged in automobile sales. His study found that the use of fuzzy theory and modern network technology provided more accurate early warning and assessment of potential and apparent risks of financial enterprises. Feng et al. [10] constructed a BPNN-based enterprise financial risk prewarning model to predict financial crises and verified that the constructed prediction model could accurately predict the financial crises of the enterprises through predicting the finance of 200 manufacturing enterprises in 2018 and 2019. Qi et al. [11] proposed a variable precision rough set weighted k-nearest neighbor (KNN) network-based financial risk control algorithm and verified the algorithm's algorithm through experiments. Zhang [12] used fuzzy neural networks to warn the credit risk of financing platform loans, verified the effectiveness of the algorithm by example analysis, and gave relevant suggestions.

III. NEURAL NETWORK-BASED EARLY FINANCIAL RISK WARNING

A. Constructing Early Risk Warning Indicators

Before warning financial risks of an investment project, it is necessary to build an indicator system that can determine the investment risks of the project, and these indicators will be used as input parameters of the project risk early-warning model [13]. Selecting early risk warning indicators generally follows the criteria of indicator criticality, data source accuracy, indicator relativity, indicator validity, and indicator simplicity. Indicator criticality means that the selected indicators are related to capital flow. Data source accuracy means that the selected indicators can be obtained from the data sources. Indicator relativity means that the selected indicators are relative indicators, ignoring the influence of company size as much as possible. Indicator validity means that the selected indicators need to be universal and valid. Indicator simplicity means that the selected indicators should be easy to calculate.

Different companies will invest in different types of projects due to their different positioning, and the risks possessed by different types of investment projects are also different. Therefore, only the general classification of indicators is given here, and the specific indicators will be given in the example analysis below. As shown in Fig. 1, the risk variables of project investment can be broadly classified into four types of risks: economic, technical, policy, and environmental risks. Project economic risk refers to the crisis faced by the investment project at the economic level. Project technical risk refers to the technical risk that arises during enterprise operation that can affect the project. Project policy risk refers to the degree of influence that can be caused by laws and regulations in the project investment process. Project environmental risk refers to the degree of influence of the local environment during the operation of the investment project, and outdoor projects are more likely to be affected by the environment, which requires specific analysis in different cases [14].

B. Neural Network-Based Project Risk Early-Warning Model

The evaluation indicators of investment projects will be input into the input layer of the BPNN algorithm [15]; therefore, the number of nodes in the input layer depends on the number of indicators used to evaluate the risk of the investment project. The output layer outputs the evaluation results of the risk level of the investment project. The risk level is represented by 1, 2, 3, 4, and 5. “1” represents a low risk, and the larger the value is, the higher the risk is. The hidden layer is the core structure of the BPNN algorithm, and its number is decided according to the demand. Usually, the more the layers and nodes are, the deeper the law can be mined, and the more accurate the model prediction is, but it will increase the amount of computation [16]. The basic process of constructing an early warning model for investment project risk is as follows.

1) The indicator data of different investment projects are collected according to the constructed financial risk early warning indicator system, as well as the risk assessment results of corresponding projects.

2) The collected data are pre-processed to eliminate the abnormal data. The indicator data of investment projects are input into the input layer of the BPNN algorithm.

3) The indicator data in the input layer are calculated layer by layer in the hidden layer [17]:

\[
\alpha = f \left( \sum_{i=1}^{n} \omega x_i - \beta \right),
\]

where \( \alpha \) is the output of every layer, \( \beta \) is the adjustment term of every layer, \( f(\bullet) \) is the activation function [18], and \( \omega \) is the weight between layers.

4) The output result of the last hidden layer is passed to the output layer. The softmax function calculates in the output layer. The risk level of the investment project is output according to the calculation result.
5) The risk level of investment projects calculated by the BPNN algorithm is compared with the actual risk level obtained when collecting data, and the error between them is calculated. The cross-entropy [19] is used as the calculation error. The calculation formula of the error is:

$$E = -\sum_{i} y_i \cdot \ln p_i,$$

where \( E \) is the error, \( i \) is the label serial number, which is the risk level, \( y_i \) is the judgment parameter [20], whose value is 1 when the actual risk level of the project is \( i \) and 0 otherwise, and 0 otherwise, \( p_i \) is the probability that the project risk level is \( i \) in the calculation result.

6) Whether the BPNN algorithm reaches the termination condition is determined. If it does, then the training ends, and the construction of the early risk warning model ends; if not, then the weight parameter in the hidden layer is reversely adjusted. The termination conditions for training the model are that the number of iterations reaches the preset maximum value or the calculation error converges to the preset threshold. The training is stopped when either of the above two termination conditions are met.

The adjustment of the weight parameter in the hidden layer is based on the calculation error and learning rate. The output result converges in the direction of minimum error through the calculation error and learning rate. In this reverse adjustment process, the learning rate is crucial as it controls the convergence speed of the algorithmic model, and it is usually a fixed value; however, in the practical application process, there are a large number of local minima in the nonlinear error surface, and once the model training falls into the local minima, it is difficult to get out, which will seriously slow down the convergence speed [21]. Therefore, the GA is introduced to adjust the weight parameter of the BPNN algorithm.

The process of training the improved BPNN algorithm using the GA [22] is as follows. Firstly, the chromosome population is generated for the GA. Every chromosome represents a parameter scheme of the BPNN algorithm, and every gene in the chromosome represents a parameter to be adjusted. Then, the parameter schemes represented by the chromosomes are substituted into the BPNN algorithm for forward computation according to steps (1)–(5) described previously to obtain the error. Then, whether the training should be terminated is determined. If not, the genetic operation is performed on the chromosome population, including crossover and mutation [23]. The crossover operation refers to exchanging the data on the same gene locus of two chromosomes according to the crossover probability, and the mutation operation refers to changing data at a single chromosome locus according to the mutation probability. The genetically manipulated population is substituted into the BPNN algorithm again to repeat steps ①–⑤ are repeated until the model reaches the termination condition.

IV. EXAMPLE ANALYSIS

A. Analysis Object

An electric power company in Hunan Province was taken as an example. The work that this power company can undertake includes power engineering survey, manufacturing, design, and sales. The company has a relatively good organizational structure. The shareholders’ meeting is the highest authority of the company and appoints other departments. The board of directors is the representative department elected by the shareholders’ meeting to manage the company’s business operations, and the operating management layer established under the board of directors manages 11 departments.

The basic process for project investment is to bid and process the winning project. The company’s management department does not have a set of strict evaluation procedures. The department manager expresses his investment intention first, and then the financial department decides whether the project can be invested in after a simple qualitative analysis. The whole process is highly subjective and nonstandard.

In addition to qualitative analysis that can determine the presence or absence of project risks, quantitative analysis is also needed to determine the level of project risks to help the management layer make more scientific and rigorous judgments [24].

B. Project Risk Early Warning Indicator System

The project data required for the case study were collected from the information disclosed on the official website of the company. A field survey was conducted on the company to collect project analysis information such as project investment plans and cost reports that are publicly available.

These data were pre-processed before formal use, which was because the data volume that the company could provide was limited and some projects were suspected of fraud. Pre-processing supplemented some data and eliminated the part that could not be supplemented. After pre-processing, the total number of samples was 500, of which 300 were randomly selected as the training set and the remaining 200 as the test set.

Before establishing the improved BPNN-based early risk warning model, the corresponding early risk warning indicator system was established. This paper analyzed the early warning indicators of this electric power company by referring to the general classification of the indicators given in the previous section. There were thirty-four second-level indicators under the four first-level indicators of economy, technology, policy, and environment. Although more early warning indicators were good for prediction accuracy, the calculation volume was also larger. Some indicators had low relevance and would not affect the prediction even if ignored. Thus, the 34 second-level indicators were screened to eliminate those with low relevance to reduce the computational effort.

Table I shows the early warning indicators screened after the KMO test, Bartlett test [25], and regression analysis, the KMO test statistic was 0.736, which exceeded 0.7, and the Bartlett test statistic was 276.35. In addition, the p-values of all 13 indicators in Table I were less than 0.01.

www.ijacsa.thesai.org
TABLE I.  THE SCREENED-PROJECT RISK WARNING INDICATOR SYSTEM

<table>
<thead>
<tr>
<th>The first-level indicators</th>
<th>The second-level indicators</th>
<th>P-value</th>
<th>Kaiser-Meyer-Olkin test</th>
<th>Bartlett’s test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project economic risks</td>
<td>Project cost-income ratio</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Net income ratio</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asset turnover ratio</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Turnover of account</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>receivable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asset-liability ratio</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liquidity ratio</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sales growth rate</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preservation and</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>appreciation ratio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project technical risks</td>
<td>Payback period</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Construction technology</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Construction management</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project environmental risks</td>
<td>Project approval method</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social conduct</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Parameter Setting

A BPNN algorithm modified by the GA was used to construct a project investment risk early-warning model. The number of nodes in the input layer of the BPNN algorithm was set as 13 according to the feature number of indexes that need to be input, and the sigmoid function was used as the activation function in the hidden layer for fitting nonlinear laws. For the GA-improved BPNN algorithm, factors that affected its prediction performance also included the population size of the GA used for adjusting the weight parameter in addition to the number of hidden layers in its structure. Therefore, orthogonal experiments were used to test the performance of the improved BPNN algorithm with one, two, three, four, and five hidden layers under the genetic population size of 10, 15, 20, 25, and 30. The population size and the number of hidden layers with the best performance were selected and used for the subsequent comparison experiments.

To further test the early warning performance of the improved BPNN algorithm for project risks, it was compared with SVM and traditional BPNN algorithms. The parameters of the SVM algorithm are as follows. The sigmoid function was used as the kernel function for mapping features to a high-dimensional space to linearize nonlinear features as much as possible, and the penalty factor was set as one. The number of nodes of input and output layers and the number of hidden layers of the traditional BPNN algorithm were the same as those of the improved BPNN algorithm, the activation function was used as the activation function, and the learning step length was 0.02.

D. Experimental Results

Table II and Fig. 2 show the results of the orthogonal experiments for the population size and the number of hidden layers of the improved BPNN algorithm. First, it was noticed from Fig. 2 that the overall accuracy of the improved BPNN algorithm for investment project risk prediction increased as the population size and the number of hidden layers increased, but the overall accuracy of the improved BPNN algorithm tended to be constant after the population size reached 25, and the overall accuracy of the improved BPNN algorithm also tended to be constant after the number of hidden layers reached four. However, comparing the average single-project time consumption under different population sizes and hidden layers in Table II, it was found that the average single-project time of the algorithm always increased as the population size and the number of hidden layers increased. In other words, increasing the population size and the number of hidden layers could increase the overall accuracy of the improved BPNN algorithm and also increase the prediction time, but after they increased to certain levels, the prediction time still increased, but the overall accuracy of the prediction tended to be constant. Therefore, the population size of the improved BPNN algorithm was 25, and the number of hidden layers was four.

Fig. 3 shows the test results of the early risk warning performance of SVM, traditional BPNN, and improved BPNN algorithms for investment projects. It was seen that the accuracy, recall rate, and F-value of the SVM algorithm for early risk warning of investment projects were 75.3%, 70.1%, and 72.6%, respectively; the accuracy, recall rate, and F-value of the traditional BPNN algorithm for early risk warning of investment projects was 91.2%, 82.1%, and 86.4%, respectively; the accuracy, recall rate, and F-value of the improved BPNN algorithm for risk warning of investment projects was 97.5%, 96.6%, and 97.0%, respectively. It was seen from Fig. 3 that the accuracy, recall rate, and F-value of the SVM-based early-warning model were the lowest, and those of the GA-based BPNN model were the highest.

TABLE II.  AVERAGE TIME SPENT BY THE IMPROVED BPNN ALGORITHM WITH DIFFERENT POPULATION SIZES AND NUMBER OF HIDDEN LAYERS ON A SINGLE PROJECT

<table>
<thead>
<tr>
<th>Population size/unit</th>
<th>One hidden layer</th>
<th>Two hidden layers</th>
<th>Three hidden layers</th>
<th>Four hidden layers</th>
<th>Five hidden layers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size 10</td>
<td>98 ms</td>
<td>110 ms</td>
<td>131 ms</td>
<td>163 ms</td>
<td>205 ms</td>
</tr>
<tr>
<td>Population size 15</td>
<td>121 ms</td>
<td>140 ms</td>
<td>162 ms</td>
<td>184 ms</td>
<td>213 ms</td>
</tr>
<tr>
<td>Population size 20</td>
<td>176 ms</td>
<td>195 ms</td>
<td>211 ms</td>
<td>234 ms</td>
<td>252 ms</td>
</tr>
<tr>
<td>Population size 25</td>
<td>223 ms</td>
<td>241 ms</td>
<td>264 ms</td>
<td>287 ms</td>
<td>303 ms</td>
</tr>
<tr>
<td>Population size 30</td>
<td>251 ms</td>
<td>272 ms</td>
<td>295 ms</td>
<td>316 ms</td>
<td>339 ms</td>
</tr>
</tbody>
</table>

Fig. 2. Overall Accuracy of the Improved BPNN Algorithm under different Population Sizes and Number of Hidden Layers.
V. DISCUSSION

Conventional enterprises will continue to expand their scales in the process of development, and they will make investments in different projects to obtain returns in the process of expansion. However, all project investments are risky. Usually, the greater the risk, the higher the return, but high risks also means that the probability of project failure. Once a project fails, the enterprise will suffer from huge losses, which is not conducive to its development. Thus, before investing in a new project, companies need to assess the risk to assist them in making decisions about the project investment. Traditional risk assessment is done manually, which is inefficient and subjective. In order to improve the efficiency of risk assessment and also to enhance the objectivity of the assessment results, intelligent algorithms are introduced into the early warning of project investment risks. This paper selects the BPNN algorithm, which can make a good fit to the nonlinear law, to warn the project investment risk and used the BPNN model to construct the project investment risk early-warning model. The BPNN algorithm was improved by the GA. An electric power company in Hunan Province was taken as a subject for analysis. The population size and the number of hidden layers in the improved BPNN algorithm were determined by orthogonal experiments. The results are as follows. (1) The increase in population size and the number of hidden layers in the improved BPNN algorithm improved the early warning accuracy; when the accuracy tended to be constant, the average time spent on early warning for a single project increased, so the final population size was set as 25, and the number of hidden layers was set as 4. (2) The accuracy, recall rate, and F-value of the SVM-based early-warning model were the lowest, those of the traditional BPNN algorithm-based model were higher, and those of the GA-improved BPNN algorithm-based model were the highest.

VI. CONCLUSION

This paper briefly introduced the indicator system used for investment project risk early-warning and used the BPNN algorithm to construct the project investment risk early-warning model. The BPNN algorithm was improved by the GA. An electric power company in Hunan Province was taken as a subject for analysis. The population size and the number of hidden layers in the improved BPNN algorithm were determined by orthogonal experiments. The results are as follows. (1) The increase in population size and the number of hidden layers in the improved BPNN algorithm improved the early warning accuracy; when the accuracy tended to be constant, the average time spent on early warning for a single project increased, so the final population size was set as 25, and the number of hidden layers was set as 4. (2) The accuracy, recall rate, and F-value of the SVM-based early-warning model were the lowest, those of the traditional BPNN algorithm-based model were higher, and those of the GA-improved BPNN algorithm-based model were the highest.

REFERENCES


