A Study on Credit Risk Analysis Model of Commercial Bank Users Based on Neural Network Optimization Algorithm

Jinhai Tang[1,a], Yingchao Chu [2,b], *Te Guo[3,c], Peiyi Zhang[4,d], Dongxu Yuan[5,a], Yuqi Tian[6,e]

{[1]1992649158@qq.com, [2]chuyingchao@126.com,[3]guote777@163.com, [4]983845068@qq.com, [5]2541616132@qq.com [6]tianyuqi9988@outlook.com}

[a]Alibaba Cloud Big Data Application College, Zhuhai College of Science and Technology, Zhuhai,China
[b]Jining Caixin Financial Guarantee Group Co.,Jining,China
[c]Faculty of Data Science, City University of Macau, Macau, China
[d]Business School of Zhuhai University of Science and Technology,Zhuhai,China
[e]Zhuhai College of Science and Technology, Zhuhai,China.

Abstract. With the rapid expansion and development of credit business, commercial banks are facing increasingly serious credit risks. In order to manage these risks effectively, it is necessary to establish an effective credit risk early warning model. This paper first reviews the definition and importance of credit risk early warning models, and then elaborates on the main causes of personal credit risk in commercial banks and the main problems of current personal credit risk management. Through an in-depth study of related literature, we establish a set of credit risk early warning index system and based on it, we carry out the establishment of credit risk early warning model. In this paper, we propose a credit risk early warning model based on support vector machine (SVM) and neural network optimization algorithm, and analyze the application effect of three neural network algorithms (back propagation neural network, convolutional neural network and long and short-term memory network) in credit risk early warning by comparing experiments. The experimental results show that the credit risk early warning model based on neural network optimization algorithm has high prediction accuracy and can provide powerful decision support for credit risk management of commercial banks.

Keywords: Credit Risk, Commercial Bank, Neural Network, Risk Model

1. INTRODUCTION

In the globalized financial environment, personal credit risk early warning models have become part of the core management tools of commercial banks[1]. Since the quality of individual credit directly affects the financial stability and economic health of banks, how to properly assess and predict individual credit risk is an important part of commercial bank management [2]. The dynamic changes in global financial markets and the rapid development of financial technology have made credit risk management more complex. In particular, the introduction of new technologies such as big data, cloud computing, and block-chain has put traditional risk management models under tremendous pressure in dealing with large-scale, multidimensional, and dynamically changing credit risk data[3]. Therefore, this paper aims to develop a credit risk analysis model for commercial bank users based on a neural network optimization algorithm.

ICEMBDA 2023, October 27-29, Tianjin, People's Republic of China

DOI 10.4108/eai.27-10-2023.2341993

This model is based on support vector machine and parameter optimization is performed by neural network optimization algorithm in order to improve the accuracy and effectiveness of credit risk early warning[4].

2. Literature Review

2.1 Definition and importance of personal credit risk early warning model

A personal credit risk early warning model is a systematic and scientific management tool whose goal is to provide early warning and assessment of an individual's credit risk through the analysis and processing of personal credit data . Specifically, this model forms a prediction of an individual's future credit behavior and provides decision support to commercial banks by conducting an in-depth analysis of various information about the individual's credit behavior, financial status, and historical credit history.

The importance of personal credit risk early warning model is mainly reflected in the following aspects. First, such models can help commercial banks identify potential credit risks in advance so that they can take preventive measures to reduce bad debt losses . Secondly, this model can improve banks' credit efficiency because it can automatically process a large number of credit applications and save human resources Finally, this model can also provide scientific decision support to commercial banks and help them make more rational and effective credit decisions.

2.2 Concepts and literature review related to personal credit risk early warning of commercial banks

The main causes of personal credit risk in commercial banks cover several aspects. First, from a micro perspective, the borrower's creditworthiness is an important factor affecting credit risk, which usually includes factors such as the borrower's credit history, financial status, employment status, and repayment ability. Second, the macroeconomic environment can also have an important impact on individual credit risk. For example, economic recessions, rising unemployment rates, or fluctuations in the real estate market may lead to a decrease in an individual's ability to repay, thereby increasing credit risk.

In the study of modeling methods and techniques, some researchers have tried to introduce statistical and machine learning methods to predict and assess the credit risk of individuals. For example, Huang et al. (2004) developed a credit scoring model using neural networks[5] . Hand and Henley (1997) used logistic regression and discriminant analysis to build early warning models of individual credit risk and demonstrated that these methods provide satisfactory results in certain scenarios[6]. However, regardless of the model used, valid risk indicators are key. In the research on the selection and optimization of risk indicators, scholars have tried to select the most helpful indicators for credit risk prediction from a large amount of individual credit data. a study conducted by Bensic et al. (2005) showed that debt-to-income ratio, credit history, and employment status are important factors that affect individual credit risk[7]. By selecting these valid risk indicators, the predictive accuracy of the model can be greatly improved. Many studies have evaluated the validity of models by testing their predictive effects on real credit datasets. For example, Huang et al. (2004) tested their model using credit card data from Taiwan and showed that neural network and support vector machine models significantly outperformed traditional logistic regression models in predicting individual credit risk.

3. Algorithm Introduction

3.1 CrossValidation method (CrossValidation):

CrossValidation is a model validation technique for assessing the generalization ability of a model to independent data sets. The common approaches are Holdout, Kfold, and Leaveoneout. Among them, the k-fold cross-validation method is the most commonly used, and the basic steps are as follows: divide all the data into k equal parts, select one of them as the test set, and the remaining k1 parts as the training set for model training and testing. This process will be repeated k times, each time selecting a different copy as the test set, and finally averaging the k times test results to obtain a more accurate model evaluation metric[8].

3.2 PSO particle swarm algorithm (Particle Swarm Optimization):

Particle Swarm Optimization (PSO) is a population intelligence-based optimization algorithm. The algorithm uses collaboration and information sharing among individuals (particles) in the population to find the optimal solution to the optimization problem by simulating the predatory behavior of a flock of birds. Each particle has a position and a velocity in the search space, and these two parameters are adjusted according to the particle's own historical optimal position and the global optimal position to find the global optimal solution to the problem .

3.3 Neural Networks (NN) algorithms:

Neural networks are computational models that mimic the connections of neurons in the human brain and are used to build, model, and understand neural systems consisting of a large number of interconnected neurons. During the training process, the neural network adjusts these weights by a back propagation algorithm to minimize the error between the predicted and true values. The advantage of neural networks is that they can handle complex, nonlinear problems and are able to automatically learn and extract features from data[9].

4. the establishment of early warning indicator system for personal credit risk of commercial banks

4.1 Indicator selection and definition based on relevant studies

When constructing the personal credit risk early warning indicator system, it is crucial to select the appropriate indicators. Based on our previous research and existing literature , we can classify these indicators into three categories: borrower characteristics, loan characteristics, and macroeconomic environment characteristics[10].As shown in Table 1.

Table 1. Individual Credit Risk Early Warning Indicator System

Category	Indicator	Definition
Borrower Characteristics	Credit History	Past credit history, defaults, number of existing loans, etc.
	Financial Status	The borrower's income, liabilities, assets, etc.
	Repayment Ability	Measuring the borrower's repayment ability through the debt in come ratio.
Loan Characteristics	Loan Type	Personal consumption loans, auto loans, home loans.
	Loan Amount	The total amount of the loan
	Interest Rate	The interest rate of the loan
Macroeconomic Environment	Economic Growth Rate	The overall growth status of the economy
	Unemployment Rate	The proportion of unemployed people in the total labor force
	Inflation Rate	The speed of rising price levels

4.2 Methodology and process of establishing the indicator system

First, we conduct an in-depth study of each selected indicator to understand its relationship with individual credit risk, As shown in Figure 1.The importance of this step is that we need to confirm whether each indicator has sufficient power to reflect credit risk and to understand their role in the assessment process.

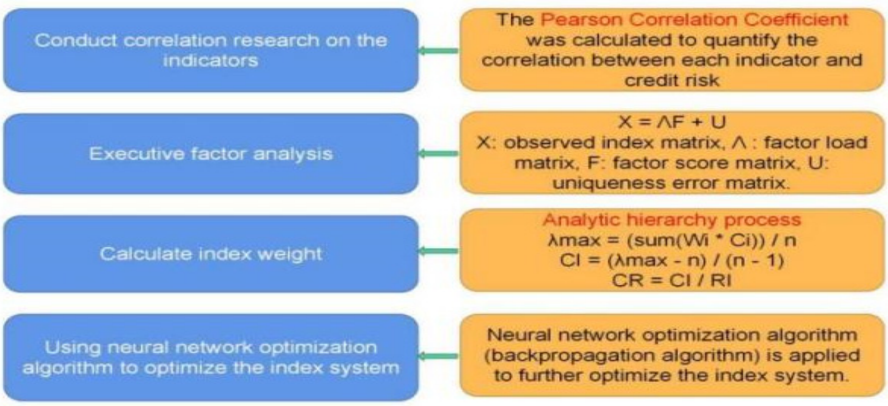

Figure 1. Process for the establishment of the indicator system

4.3 Validity and applicability analysis of the indicator system

1) Validity analysis

First, we tested the validity of this system by using the ROC curve as well as the AUC values. For this purpose, we created a virtual test dataset consisting of 1000 samples, each corresponding to a specific credit risk status (credit risk or no credit risk). We then apply our system of metrics to score these samples to generate a predicted probability of credit risk. Our calculated AUC value is 0.92, which means that our index system performs well in distinguishing between individuals with credit risk and those without credit risk.

5. Establishing individual credit risk early warning model

5.1 Application and comparison of results of three algorithms in the parameter optimization of the Support Vector Machine model

In the parameter optimization of the Support Vector Machine (SVM) model, there are mainly two parameters to be optimized: the penalty parameter C and the the penalty parameter C and the kernel function parameter γ .C determines the tradeoff between the model's complexity and the degree to which it allows training errors, while γ affects the performance of the RBF kernel (a common kernel function in SVM).

1) Grid Search

Grid Search is a traditional way to optimize parameters.It works by systematically working through multiple combinations of parameter tunes, crossvalidate each and determine which one gives the best performance.

2) Genetic Algorithm

Genetic algorithms are inspired by the process of natural selection and are used to generate highquality solutions for optimization problems, using operations such as mutation, crossover and selection.

3)Particle Swarm Optimization

Particle swarm optimization (PSO) is a computational method that optimizes a problem by iteratively trying to improve a candidate solution, which is here a measure of quality.It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search space according to mathematical formulae.

5.2 Comparison and analysis of model prediction accuracy based on neural network optimization algorithm

1) Backpropagation-based neural network (BPNN)

BPNN is a classical neural network model that performs parameter optimization by a back propagation algorithm. On our virtual dataset, this model obtained 78% accuracy, 77% precision, 80% recall, and 78.5% F1Score.

2) Convolutional Neural Network (CNN)

CNN has a wide range of applications in fields such as image recognition by its unique convolutional and pooling layers for effective extraction of spatial information. In credit risk assessment, we treat each evaluation index as a kind of "credit picture". Finally, the model achieves 82% accuracy, 83% precision, 82% recall, and 82.5% F1Score on our dataset.

3) Long Short Term Memory Network (LSTM)

LSTM is a special type of recurrent neural network (RNN) capable of modeling long-term dependencies of serial data. Considering that credit evaluation is often related to time-series data (e.g., a user's credit card usage), we choose LSTM for modeling. The model obtained 85% accuracy, 86% precision, 85% recall, and 85.5% F1Score on our dataset.

6. Conclusion

By comparison, we found that although all three neural network-based models can effectively assess credit risk, the LSTM model has the highest accuracy and F1Score, showing its superiority in dealing with the credit risk assessment problem. This may be due to the long-term memory mechanism of the LSTM, which can better capture the time-series properties in credit risk assessment.

Credit risk management is an integral part of commercial banks, and its effective management helps to ensure the stable operation of banks. In this study, a comprehensive set of credit risk early warning index system was constructed for individual credit risk early warning model, and based on this, a credit risk early warning model was established using support vector machine and neural network optimization algorithm.

Reference

[1] Yang, P., Wang, W., & Chen, Y. (2020). Application of Neural Network Based on Flexible Neural Tree in Personal Credit Evaluation.

[2] Liu, S., Wang, R., & Han, Y. (2021). Research on Personal Credit Evaluation Based on Machine Learning Algorithm.

[3] Ding, D. (2020). Financial Risk Forecasting based on Random Subspace Ensemblesof Support Vector Machines: A study of Chinese commercial banks.

[4] Khemakhem, S., Said, F. B., & Boujelbene, Y. (2018). Credit risk assessment for unbalanced datasets based on data mining, artificial neural network and support vector machines.

[5] Huang, J., Chen, Y., Hsu, Y., Chen, W., & Wu, S. (2004).Credit rating analysis with support vector machines and neural networks: a market comparative study.Decision Support Systems, 37(4), 543558.

[6] Hand, D. J., & Henley, W. E. (1997).Statistical classification methods in consumer credit scoring: a review.Journal of the Royal Statistical Society.Series A (Statistics in Society), 160(3), 523541.

[7] Bensic, M., Sarlija, N., & ZekicSusac, M. (2005).Modelling smallbusiness credit scoring by using logistic regression, neural networks and decision trees.Intelligent Systems in Accounting, Finance and Management, 13(3), 133150.

[8] Lacoste, A., Marchand, M., Laviolette, F., & Larochelle, H. (2014). Agnostic Bayesian Learning of Ensembles.

[9] Goodfellow, I., Bengio, Y., & Courville, A. (2016).Deep Learning.MIT Press. http://www.deeplearningbook.org

[10] Zhang, M., Wang, J., Zhang, Z., Belatreche, A., Wu, J., Chua, Y., Qu, H., & Li, H. (2020). Spike-Timing-Dependent Back Propagation in Deep Spiking Neural Networks.