Financial Risk Assessment and Early Warning System Based on Machine Learning Algorithm

Xu Zhang

Universiti Sains Malaysiae, Penang, 11600, Malaysia

Keywords: Financial risk assessment; ML algorithm; Early warning system; System architecture; Risk management

Abstract: Traditional financial risk assessment methods have limitations in the face of big data and nonlinear relationships. Therefore, this study introduces ML (Machine Learning) technology in order to improve the accuracy and timeliness of risk assessment. In terms of methods, this paper first combs the theoretical basis of financial risk assessment and makes clear the application advantages of ML algorithm in financial risk assessment. Then, a multi-level and multi-module system architecture including data collection, preprocessing, model training and evaluation, risk early warning and user interaction is designed, and financial data is deeply studied and analyzed by ML algorithm. Through the system test, the stability and accuracy of the system are verified, and its application potential in actual financial risk assessment is demonstrated. This study provides a new, more accurate and efficient risk assessment and early warning tool for financial institutions and investors, which is helpful to improve the stability and transparency of the whole financial market.

1. Introduction

In today's increasingly complex global economy, financial market fluctuates frequently and uncertainties increase, making financial risk assessment a key link to maintain financial stability and protect investors' interests [1]. Traditional financial risk assessment methods mostly rely on expert experience and historical data analysis, but in the face of big data, nonlinear relationship and rapidly changing market environment, its limitations and lag are increasingly prominent [2]. Therefore, it is particularly important to explore more efficient and accurate financial risk assessment methods [3-4]. This research is carried out under this background, aiming at using the cutting-edge technology of ML to provide new ideas and methods for financial risk assessment, so as to improve the accuracy and timeliness of assessment and provide more scientific and reliable decision support for financial institutions, regulators and investors.

The main purpose of this study is to build a financial risk assessment and early warning system based on ML algorithm, which can automatically learn the potential laws in market data, identify risk signals and issue early warning in time, and provide a powerful tool for financial risk management.

2. Theoretical basis and methodology

2.1. Theoretical basis of financial risk assessment

Financial risk assessment is one of the core topics in the financial field, and its theoretical basis covers many disciplines such as finance, statistics and economics [5]. Traditionally, financial risk assessment mainly relies on capital asset pricing model, credit rating model, Value at Risk model and other tools. These models provide a basis for financial decision-making by quantifying market risk and credit risk [6]. However, with the complexity of the financial market and the surge of data, the traditional model is unable to deal with the nonlinear relationship and capture the dynamic changes of the market. Therefore, exploring new risk assessment theories and methods, especially combining with modern information technology, has become an important direction of current financial risk assessment research.

  DOI: 10.25236/icemeet.2024.075

2.2. Overview of ML algorithms

As an important branch of artificial intelligence, ML has become a powerful tool to deal with complex data and solve nonlinear problems by allowing computers to automatically learn laws from data and then make predictions or decisions [7]. In financial risk assessment, ML algorithms, such as logistic regression, support vector machine, decision tree, random forest and neural network, are widely used in credit scoring, market trend prediction, fraud detection and other fields because of their strong data processing ability and pattern recognition ability [8]. These algorithms can automatically extract data features and discover complex relationships between data, thus improving the accuracy and efficiency of risk assessment. Based on these ML algorithms and the characteristics of financial data, this study will build a model suitable for financial risk assessment in order to achieve more accurate risk assessment and early warning.

3. Construction of financial risk assessment and early warning system based on ML algorithm

3.1. System architecture and functional modules

When building a financial risk assessment and early warning system based on ML algorithm, this paper first designs a clear and reasonable system architecture. The system architecture includes data acquisition layer, data preprocessing layer, model training and evaluation layer, risk early warning layer and user interaction layer. Each level undertakes different functions and works together to complete the whole process of risk assessment and early warning. Table 1 shows the five main levels of financial risk assessment and early warning system based on ML algorithm and their respective functions, which is helpful to understand and build the system.

Table 1: Architecture of a Financial Risk Assessment and Early Warning System Based on ML Algorithms

Hierarchical name	Functional description
Data acquisition layer	Responsible for collecting original data from different financial data sources, such as transaction records, user information, market data, etc.
Data preprocessing layer	The collected original data are cleaned, transformed and normalized to eliminate noise and dimensional differences and improve data quality.
Model training and evaluation layer	The preprocessed data are used to train ML models, such as DNN (Deep Neural Network), and the performance of the model is evaluated by cross-validation and accuracy evaluation.
Risk early warning layer	According to the trained model, the risk of new financial data is evaluated, and risk warning signals are generated, such as high-risk trading tips, credit score decline warning and so on.
User interaction layer	Provide user-friendly interface and interactive mode, so that users can conveniently view risk assessment results, early warning signals and system reports, and support users to set parameters and adjust models.

In addition to the hierarchical structure, the system also includes several functional modules, including data processing module, model building module, early warning generation module and so on. These modules are independent and interrelated, which together constitute a complete risk assessment and early warning system.

3.2. The realization of ML algorithm

In the financial risk assessment and early warning system, the model training and assessment layer is the core part of the system. It uses ML algorithm to learn the preprocessed data, constructs a risk assessment model, and verifies and optimizes the performance of the model through cross-validation and accuracy evaluation. Based on the DNN in ML algorithm and the characteristics of financial data, this section constructs a model suitable for financial risk assessment in order to achieve more accurate risk assessment and early warning. The construction of DNN model first involves data preprocessing and feature engineering in financial risk assessment. This step is very important, which includes cleaning, normalizing or standardizing financial data to

eliminate data noise and dimensional differences. At the same time, it uses domain knowledge to extract or construct features with business meaning from the original data, including transaction frequency, account balance change rate, etc., so as to enhance the model's ability to capture risk factors. The loss function of the DNN model is as follows:

$$L = -\sum_{c=1}^{M} y_{o,c} \log\left(p_{o,c}\right) \tag{1}$$

Where M is the number of categories, $y_{o,c}$ is a binary indicator, and $p_{o,c}$ is the prediction probability of observing that o belongs to category c.

Based on the preprocessed data, a DNN model is constructed. The model consists of multi-layer neurons, including input layer, multiple hidden layers and output layer. The input layer receives the processed feature vectors, while the hidden layer transforms the input data through nonlinear activation function, and extracts advanced abstract features layer by layer. Finally, the output layer gives the risk assessment result, credit score or fraud probability. In the process of model training, this paper uses back propagation algorithm and optimizer to constantly adjust the network weight to minimize the loss function and improve the performance of the model on the verification set. At the same time, in order to prevent over-fitting, L2 regularization measures are also adopted. L2 regularization refers to adding L2 norm of model parameters as a penalty term to the loss function, and the loss function becomes:

$$L(w) = L_0(w) + \lambda |w|_2^2 \tag{2}$$

Where $L_0(w)$ is the original loss function, λ is the regularization parameter, and $|w|_2^2$ represents the square of L2 norm of vector w. In order to ensure that the model has good generalization ability and can accurately evaluate the risk level of unknown samples.

3.3. System test

After the system is built, this section conducts a comprehensive test of the system to ensure its stability and accuracy. Testing covers all aspects of the system, including functional testing, performance testing and safety testing.

Functional testing is mainly to verify whether each functional module of the system can work normally. The results are shown in Table 2:

Table 2: Functional Testing Results

Functional Module	Test Content	Test Result	Remarks
User Login Module	Can log in with correct username and password	Pass	No anomalies
	Can log in with incorrect username or password	Fail	Error message
Data Entry Module	Can data be entered correctly into the system	Pass	No anomalies
	Handling of illegal character input	Fail	Error message
Data Query Module	Can correct query conditions return accurate results	Pass	No anomalies
	Handling of incorrect query conditions	Fail	No results prompt
Report Generation Module	Can accurate reports be generated based on data	Pass	Report content accurate
System Settings Module	Can system settings be modified correctly	Pass	Settings take effect

From the functional test results, all functional modules of the system are normal. Performance test is to evaluate the response speed and stability of the system when dealing with a large amount

of data, so as to ensure that the system can meet the needs of users in practical applications. The results are shown in Table 3:

Table 3: Performance Testing Results

Test Scenario	Data Volume	Response Time (seconds)	Stability Assessment	Remarks
Data Entry Performance Test	1,000 entries	<1	Stable	No significant delay
	10,000 entries	<5	Stable	Slight delay, acceptable
Data Query Performance Test	Single query	<0.5	Stable	Quick response
	Batch query	<3	Stable	Fast response
Report Generation Performance Test	Small report	<2	Stable	Generated quickly
	Large report	<10	Stable	Longer generation time, but stable

The performance test results show that the system has good response speed and stability when dealing with scenes with different data volumes. Security testing is to check whether there are security loopholes and hidden dangers in the system to ensure the security and privacy of user data. The results are shown in Table 4:

Table 4: Security Testing Results

Test Item	Test Content	Test Result	Remarks
SQL Injection Test	Attempt to obtain data through SQL injection	Fail	System effectively defends against SQL injection
XSS Attack Test	Attempt to inject malicious scripts through XSS attack	Fail	System effectively defends against XSS attack
Password Security Test	Attempt with weak passwords (123456)	Fail	System prompts password strength insufficiency
	Attempt at brute force attack	Fail	System limits login attempts
Data Encryption Test	Check if data transmission and storage are encrypted	Pass	Data encrypted, secure
Access Control Test	Low-privilege user attempts to access high-privilege functions	Fail	System denies access, access control effective

The security test results show that the system has high security.

To sum up, the system performs well in function test, performance test and safety test, and meets the requirements of users for the stability and accuracy of the system. The functional test verifies the normal work of each functional module of the system. The performance test proves the rapid response and stability of the system when dealing with a large number of data. Security testing ensures the security and privacy of user data. Therefore, this paper thinks that the system is reliable, efficient and safe.

4. Conclusions

This study focuses on the construction of financial risk assessment and early warning system based on ML algorithm. By deeply exploring the theoretical basis of financial risk assessment, this paper clarifies the limitations of traditional assessment methods, and expounds the advantages of ML technology in dealing with complex financial data and capturing market dynamic changes. On this basis, a multi-level and multi-module system architecture including data acquisition, preprocessing, model training and evaluation, risk early warning and user interaction is designed and implemented. Through the system test, the stability and accuracy of the system are verified, and its application potential in actual financial risk assessment is demonstrated. The construction and

optimization of early warning mechanism further enhances the practicability and timeliness of the system, and provides more accurate and timely risk early warning services for financial institutions and investors.

The practical significance of this study lies in providing a new, more accurate and efficient risk assessment and early warning tool for financial institutions and investors. Through the application of this system, financial institutions can identify and control risks more accurately and improve the level and efficiency of risk management. Investors can obtain risk warning information more timely and make more wise investment decisions. The popularization and application of the system will also help to improve the stability and transparency of the whole financial market and promote the healthy development of the financial industry. From the perspective of social value, this study is helpful to maintain the order and stability of the financial market, protect the legitimate rights and interests of investors, promote the rational allocation and effective utilization of financial resources, and make contributions to social economic development and prosperity.

References

[1] Tang Chun, Liu Xiaoxing. Research on early warning of systemic financial risks based on integrated classification algorithm [J]. Logistics and Management, 2023, 32(6):186-191.

[2] Chen Yu, Cao Xinyi, Jin Shuyue, et al. Risk measurement of financial time series system: dynamic binary Dvine model [J]. Journal of China University of Science and Technology, 2023, 53(11):1-11.

[3] Jin Bo, You Yiling, Li Zhongfei. Generalized risk and return prediction based on forward-looking information [J]. journal of management sciences in china, 2024, 27(3):91-111.

[4] Xu Kun, Li Ying, Bao Xinzhong. Credit risk early warning of e-commerce small and micro enterprises considering the emotional characteristics of the text [J]. Logistics and Management, 2023, 32(12):195-201.

[5] Chen Shaoling, Li Jie, Tan Liming, et al. Study on high-dimensional time-varying measurement and transmission mechanism of systemic financial risk in China [J]. World Economy, 2021, 044(012):28-54.

[6] Liu Min, Ren Zhongyuan, Zhou Decai, et al. Research on Internet financial risk measurement and prediction based on mixed data [J]. Journal of Management, 2022, 19(12):1847-1854.

[7] Zhao Zhihong, Ying Du. Research on credit fraud risk prediction model based on machine learning [J]. Modern Electronic Technology, 2021, 044(024):177-181.

[8] Chen Zhenlong, Jin Shang. Bilateral Burr distribution and its application in financial market risk prediction and optimization [J]. System Science and Mathematics, 2023,43(02):505-530.