

## Research Article

# Research on the Construction of Financial Supervision Information System Based on Machine Learning

Ge Lin,<sup>1</sup> Ai Shangle ,<sup>2</sup> Zhao Haoxiang ,<sup>1</sup> Yu Jingyue ,<sup>1</sup> and Yang Junyao<sup>1</sup>

<sup>1</sup>School of Economics, Xinxiang University, Xinxiang 453003, China

<sup>2</sup>School of Public Administration, Guangdong University of Finance, Guangdong Guangzhou, 510521, China

Correspondence should be addressed to Ai Shangle; 47-118@gdudf.edu.cn

Received 8 March 2022; Revised 7 April 2022; Accepted 18 May 2022; Published 15 June 2022

Academic Editor: Kuruva Lakshmana

Copyright © 2022 Ge Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to fully implement systematic, continuous and effective supervision of financial institutions and promote the safe, steady, and efficient operation of China's financial system, this research needs to develop a fully intelligent financial supervision information system, so as to take measures to effectively prevent and resolve financial risks. In this paper, based on ML (machine learning), an LSTM (long short-term memory) model with good comprehensive performance is built. This model is different from the existing scorecard model which relies on statistical learning. It not only further reduces the dependence on financial experts, but also has the ability of rapid iteration. The post-loan risk early warning model based on RF (random forest) algorithm is designed. The model parameters are optimized, which makes the risk early warning model have higher accuracy. The results show that when the data amount is 10,000 to 50,000, the accuracy of the model is relatively high when it is input into the model. When the amount of data is low, the overall throughput of the model is still very high. It shows that the pre-warning model of post-lending risk constructed in this paper has a strong risk prediction ability.

## 1. Introduction

With the great strides from digitalization to intellectualization in various fields of economy and society, the traditional risk management system does not adapt to the new mode of financial innovation and development, and the problem of financial risk management has become increasingly prominent. To further improve the quality and efficiency of financial supervision, not only must an independent financial regulator be established, but also set up advanced financial supervision information system and risk early warning system [1]. When conducting financial supervision, strengthening consultation and communication, establishing and improving the institutional framework of information sharing and coordination among various departments are of vital importance to maintaining financial stability. The essence of financial electronization and networking is a profound change of emerging information technology to traditional financial industry [2]. Information system has become an indispensable platform for strategic decision-making, management, and business operation of financial industry, and

the financial industry is transforming from financial intermediary to financial information service industry to some extent.

Compared with developed countries, China's financial supervision has a big gap in the basic work of financial supervision, such as the construction of supervision index system and supervision information system, which seriously affects the effective exertion of financial supervision functions and the improvement of supervision quality and efficiency [3, 4]. At present, financial data centralization has become a trend, and in order to cooperate with business centralization, supervision centralization is needed. However, apart from public information, financial regulators usually cannot observe the global financial network [5]. The development and design of off-site financial supervision system supported by Internet information technology not only effectively improves the efficiency of data collection and realizes the sharing of information and data among various departments, but also improves the efficiency of risk identification, which enables financial regulators to enhance the identification of financial risks and lays a solid foundation

for further deepening financial services [6]. Literature [7, 8] found that many financial platforms are reluctant to make online loans because they do not have enough information about lenders and put forward that enough user credit data must be collected if online credit is implemented. Literature [9] found that online credit risk can be controlled, and all the above problems can be solved only by solving the information asymmetry between credit companies and credit users. In literature [10] from the analysis of the loan data of users on the credit platform, it is found that online credit is superior to traditional credit methods in cost control and the benefits it brings are higher than those of traditional credit business, so it is of practical significance to promote online credit. Literatures [11, 12] adopt linear support vector machine classifier, combine genetic algorithm and particle swarm optimization to select parameters, and use discriminant analysis to identify the bankruptcy grade of bank customers. The research shows that this method can better identify bankrupt customers than the original model. Literature [13] realizes parameter estimation and variable selection simultaneously by penalty method and proves that the network structure based on logical regression is better than other methods by Monte Carlo model. Literatures [14, 15] put forward an improved idea that only the wrong samples are artificially synthesized when dealing with unbalanced samples, which improves the problem of hyperplane deviation when classifying unbalanced samples by SVM (support vector machine), and applies the improvement to the case of credit risk assessment of customers of small loan companies, and its classification accuracy is higher than other algorithms.

With the development of ML (machine learning) technology, the traditional credit service model has been broken, and a risk model based on artificial intelligence and supplemented by big data has been gradually opened up. Based on the advantages of existing customers and data, Internet finance enterprises take the lead in opening the online credit model based on big data and ML [16]. This paper studies and realizes the off-site supervision system which is based on the characteristics of commercial banks and meets their supervision requirements. The system mainly realizes the functions of theme analysis, index analysis, data analysis, abnormal changes, etc. Especially for the index analysis function, it analyzes each index from the time dimension, organization dimension, and other dimensions so that commercial banks can timely grasp their own operating conditions from different angles, improve internal management, and improve risk monitoring level, which has a certain constructive significance for commercial banks.

The specific innovative achievements of this paper mainly include the following contents:

- (1) The risk management of banks should not only stay at the current level, but also need to find out the risks of banks in time and analyze the long-term risks of banks. In this paper, a financial supervision information system is established, which establishes a unified off-site supervision data center to provide data basis for risk analysis

- (2) The results of the supervision system are used to strengthen on-site inspection of institutions with deteriorating financial situation, the areas that need to be audited most in the identification process, and assign more experienced supervisors to financial institutions in trouble. In this section, the risk early warning model after loan is constructed by RF. In the data set, each region is divided into two regions by inputting recursive data in space, and the corresponding binary tree is constructed

The paper is organized as follows: The first section introduces the research background and significance and then introduces the main work of this paper. The second section puts forward the concrete methods and implementation of this research. The third section verifies the superiority and feasibility of this research model. The fourth section is the summary of the full text.

## 2. Research Method

*2.1. Construction Scheme of Financial Supervision Information System.* In recent years, the head office of the People's Bank of China and branches at all levels have developed several versions of financial supervision information systems with different characteristics and covering different business functions. These systems cover all major aspects and links of financial supervision such as institutional management, off-site supervision, and on-site inspection. Although a lot of work has been done and some achievements have been made in the construction of China's banking financial supervision information system, there are still many obvious deficiencies and problems compared with the rapidly developing economic and financial situation and the standardized, scientific, and systematic requirements for financial supervision after China's accession to the WTO. Perfect data collection system is the important foundation and main component of financial supervision information system, and also the prerequisite for effective supervision. To strengthen the networking of business systems between CBRC and the People's Bank of China and the regulated departments, or direct data collection, every time a commercial bank innovates a new business, it is necessary to provide a regulatory interface for the regulatory authorities.

The risk management of banks cannot just stay at the current level. It is necessary not only to discover the bank risks in time, but also to analyze the long-term risks of banks and give early warning tips for bank risks. This is the purpose and significance of establishing off-site supervision system of commercial banks [17]. Off-site supervision system of commercial banks can analyze the risks faced by financial institutions and their branches and the possibility of risk assessment degradation and provide reference for decision makers of financial institutions and can make decisions to resolve risks and ensure the normal and stable operation of financial institutions.

The application service platform of off-site supervision system of commercial banks interacts with the working platform of financial institutions through the application server

and communicates with massive report data through the database server. This system is a huge and complex system, and it takes considerable manpower and material resources to realize it. The author of this paper is involved in it and is mainly responsible for developing more complex index analysis functions. The functional architecture diagram of the system is shown in Figure 1. We have reviewed the relevant works and sorted out the contributions and limitations of these studies (see Table 1 for details).

Thus, this system establishes a unified off-site supervision data center, which provides data basis for risk analysis. Secondly, learn from advanced regulatory analysis concepts and technologies, establish a data analysis platform, and assist business departments to conduct online analysis through business data, so as to achieve the purpose of effective supervision, risk prevention, and decision-making. Finally, on the basis of synchronizing the data of the Head Office, analyze the submitted data with its own analysis method.

Because financial institutions are involved in massive submitted data, and the data loaded into the system are mainly analyzed and compared with various purposes and angles, the data flow direction of the system is relatively simple; that is, after being loaded into the system, it flows into each module of the system for comparative analysis, and the corresponding results are obtained for users to view. The system can define the data source and calculation rules of normative analytical indicators according to the agreed indicator formula, and the calculation results are stored on schedule [18, 19]. Indicator calculation under multi-version basic reports is supported. Formulas under applicable versions can be obtained according to the number of periods of statistical analysis, and the historical continuity of the same indicator across versions can be maintained.

Some functions of data analysis are similar to those of index analysis, except that the index analysis is a comparative analysis of each index calculated according to the data submitted to the report, while data analysis is a comparison of individual cells of the report in different periods. The data analysis module includes two sub-functions-homogeneous analysis and time series analysis, and the general process is similar to that of index analysis, so I will not repeat them here.

In this study, the latest ML-based financial supervision information system is proposed. Through the data collected by regulators and the data stored by financial institutions, available and valuable information can be mined, which can provide information processing methods for the financial supervision structure to keep abreast of market trends, formulate appropriate supervision policies and improve the market supervision system. The overall construction scheme of financial supervision information system is shown in Figure 2. The whole system consists of data collection layer, system database layer and system function layer.

The data collection layer is mainly responsible for data collection, and its main functions include the following: financial information entry system; regulatory agency interaction platform, through which all regulatory agencies can share information; and the information audit layer audits

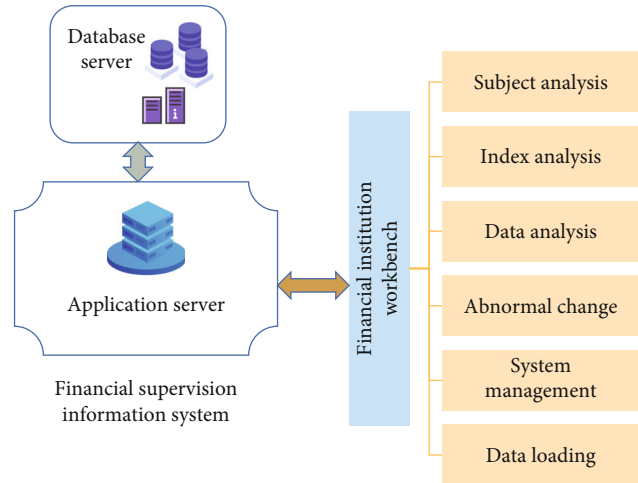


FIGURE 1: System function architecture diagram.

the format of the input data. The system database layer stores the data obtained from different channels in the database according to the regulatory tasks required by key regulatory indicators.

The function layer mainly realizes the ability of statistical analysis and data mining: It completes the statistical analysis of routine data in the process of financial supervision, generates various statistical reports, and calculates various risk exposures through the built model [20]. On the basis of classical data mining algorithms such as association rules, sequence patterns, Markov chains, and genetic algorithms, this paper puts forward mining algorithms suitable for financial supervision and can also make full use of the current mature data warehouse technology to deeply mine and utilize these information.

**2.2. Design of Financial Risk Control Model.** The development of modern information technology not only reduces the cost of financial supervision and improves the efficiency of supervision, but also increases the complexity of supervision. Data warehouse, OLAP, and data mining technologies can help financial supervision departments to adapt to this changing trend. By analyzing the data integrated in the data warehouse, the information related to supervision can be obtained, which can help to establish the supervision system more quickly and perfectly, thus scientifically and effectively identifying, measuring, monitoring, and controlling risks.

At present, the regulatory agencies require banks to implement the reporting system; that is, the reports of monthly, quarterly, and annual reports reflect the information at a certain point in time, rather than real-time data. Moreover, due to the imperfect self-restraint and incentive mechanism of commercial banks, the authenticity of the report information can hardly be guaranteed. Therefore, regulators should actively carry out spot checks on real-time information of banks while requiring banks to submit statements regularly, which can not only strengthen supervision and ensure the accuracy of information, but also obtain real-time information and more effectively assist regulators to supervise effectively.

TABLE 1

Reference/title	Contribution	Research gap/limitations
Construction of legal system for maritime international trade financial supervision cooperation	It is found that many financial platforms are unwilling to make online loans because they do not have enough lender information and it is proposed that if online loans are implemented, enough user credit data must be collected	Less consideration of internal management
Risk disclosure and performance of Malaysian construction public listed companies	It is found that the network credit risk can be controlled. All the above problems can be solved only by solving the information asymmetry between credit companies and credit users	Time dimension and organization dimension are not considered
Podda et al.'s ensembling and dynamic asset selection for risk-controlled statistical arbitrage	From the analysis of users' loan data on the credit platform, it is found that online credit is better than traditional credit in cost control and its income is higher than traditional credit business. Therefore, it is of practical significance to promote online credit	Less consideration of internal management
Research on financial early warning of mining listed companies based on BP neural network model	Particle swarm optimization classifier and support vector machine are used to identify bankruptcies	Unable to provide real-time feedback on abnormal changes of commercial banks
Preferences and beliefs about financial risk taking mediate the association between anterior insula activation and self-reported real-life stock trading	The penalty function method is used to realize the parameter estimation and variable selection at the same time, and the Monte Carlo model is used to prove that the network structure based on logical regression is better than other methods	Time dimension and organization dimension are not considered
Optimal investment–reinsurance strategies with state dependent risk aversion and VaR constraints in correlated markets	This paper puts forward the improvement idea of only synthesizing wrong samples when dealing with unbalanced samples, improves the problem of hyperplane deviation in the classification of unbalanced samples by support vector machine, and applies the improvement to the case of customer credit risk assessment of small loan companies, and its classification accuracy is higher than other algorithms	Less consideration of internal management

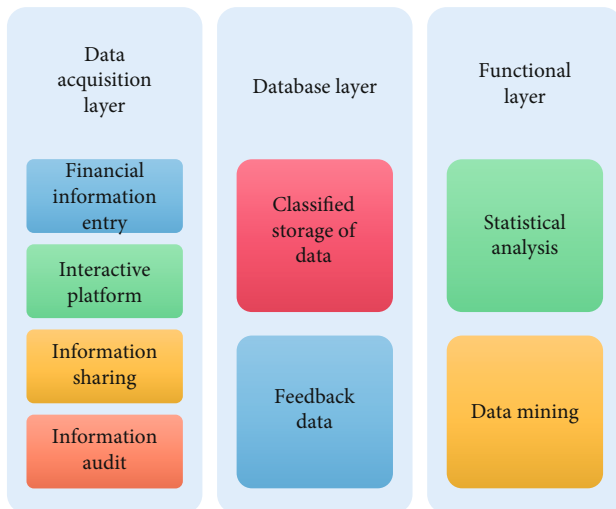


FIGURE 2: Overall construction scheme of financial information system.

Financial information collection is between the financial business department and the financial supervision department. It can be independent to form a separate organization,

which is responsible for providing financial information. Because the organization is independent, detached, and neutral, and there is no conflict of interest with the financial business department and financial supervision department, the accuracy of financial information can be guaranteed to the greatest extent [21]. Therefore, financial regulators must further artificially analyze and evaluate the financial supervision report generated by the technology-driven financial supervision system according to the ideas, values, principles and even policy choices behind the financial supervision rules, which constitutes the complementarity between artificial principle supervision and technical rule supervision.

This paper establishes a model with one input layer, three hidden layers, and one output layer [22]. Next, taking the model with only one hidden layer as an example, the algorithm flow of LSTM (long short-term memory) model is described in detail. The model is shown in Figure 3.

Each LSTM layer contains three parts: forgetting gate, input gate, and output gate. The goal of LSTM is to control the transmission of information through these three control gates, so as to solve the possible gradient disappearance phenomenon in neural network. Its activation function also uses sigmoid, which actually determines the retention degree of long-term memory flow. Finally, the output gate reads the output content of the forgetting gate and uses sigmoid

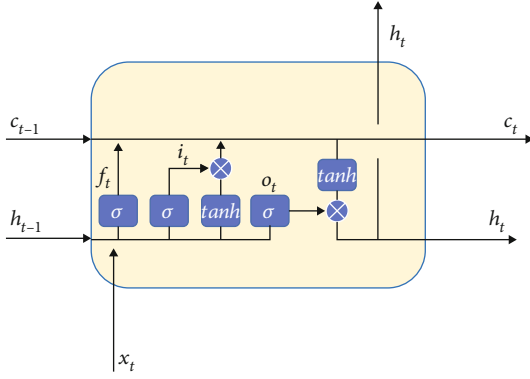


FIGURE 3: LSTM model.

function to control the output content. Its actual effect is to determine the degree of memory flow in a short time.

The input value of the input gate consists of two parts: the weighted sum of the input vector and the weighted sum of the state value at the last moment of the hidden layer. Formula (1) is shown as follows:

$$x_p^t = \sum_{i=1}^I w_{ip} x_i^t + \sum_{h=1}^H w_{hp} y_h^{t-1}. \quad (1)$$

The output of the input gate is the calculation result of the function:

$$y_p^t = f(x_p^t). \quad (2)$$

The state value fed back to the input gate by the hidden layer memory unit is

$$s_c^t = s_c^{t-1} + y_p^t g(x_c^t). \quad (3)$$

The output formula of the forgetting gate is shown

$$y_\phi^t = f(x_\phi^t). \quad (4)$$

The output result of the hidden layer memory cell is

$$y_c^t = y_\pi^t h(s_c^t). \quad (5)$$

Next, let us look at the back propagation. The purpose of back propagation is to adjust the weight of the model through the existing calculated errors so that the model can iterate in a good direction.

The used user credit data set includes the user's collection data, the user's credit score data, other credit data, overdue data, etc. By collecting these credit user information, the corresponding repayment rate model is established. Through the input data, in the study, the model can be trained, and then the weights of each connection point can be obtained. That is to say, when determining the learning rate, the LSTM model will adjust the weight according to the prediction error until the maximum number of iterations is reached [23]. Finally, the activation function used in this

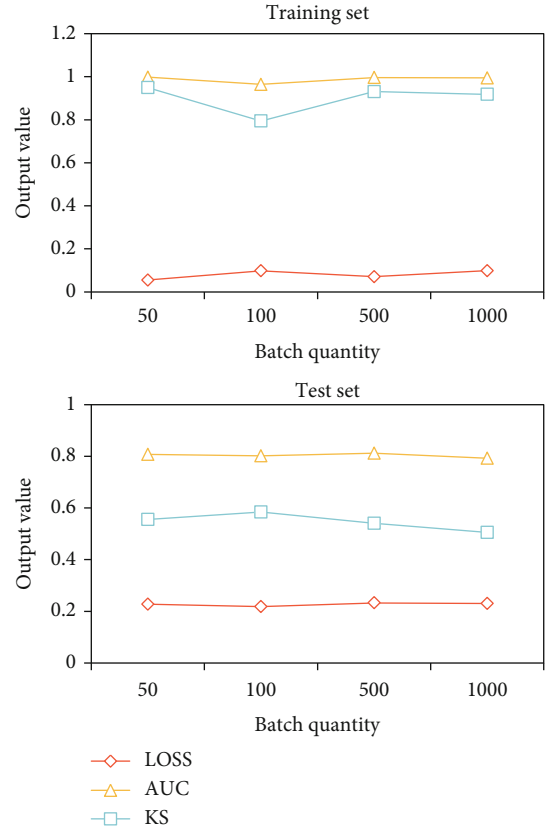


FIGURE 4: LOSS, AUC, and KS values under different error functions and optimization functions.

paper is Sigmoid commonly used in LSTM network, the error function is the mean square error function, and the optimization function is Adam optimization function which has the best effect at present. The hidden layer is set to 3 layers, and the number of hidden layer nodes is 64, 32, and 16 in turn. Batch number is 100, and iteration number is set to 1000.

When the batch number is 100, the loss value reaches a relatively low value, and the loss reaches a minimum value when the number of iterations is about 17 times. After all the four models are trained, the final values of loss, AUC, and KS are shown in Figure 4.

It can be seen from Figure 4 that different batch processing speeds have little effect on AUC and KS values and smaller batch processing numbers have better calculation speed. To sum up, the batch number selected in this paper is 100 times.

#### Construction of Financial Risk Early Warning Model

In recent years, with the vigorous development of ML-related technologies, many business giants have created their own online financial products and service models by using related technologies. With the advent of the artificial intelligence era, credit business has gradually become one of the core businesses of banks and new Internet banks. Online lending not only meets the needs of users, but also breaks the complicated manual investigation and approval mode of traditional credit process.

Post-loan risk is the last barrier in the whole life of loan. The research goal is to better help credit companies manage the post-loan funds by analyzing the factors that affect the post-loan repayment rate; identify the risks of credit users and calculate the repayment rate; and give an early warning to users who may have risks, so as to facilitate later collection. In the past two decades, various supervision systems have been developed, but their purpose is generally the same—to identify financial problems in financial institutions—so as to establish standards for limited inspections and other supervision means. The output of the supervision system is used to strengthen the on-site inspection of institutions whose financial situation deteriorates; identify the areas that need to be reviewed most in the inspection process; and assign more experienced supervisors to financial institutions in trouble.

Before studying the post-loan repayment and overdue situation, this paper briefly expounds the relationship between the types of overdue users and risk levels. Users' overdue risk levels after lending can be divided into four categories [24]. The first category is mild overdue, which refers to users with strong repayment ability and willingness. The second category is moderate and slight overdue, which refers to the willingness to repay, but its repayment ability is insufficient, and the risk of default is moderately low. The third category is medium overdue, which means that users have repayment ability without repayment willingness, and the risk of default is medium. The fourth category is serious overdue, which means that users have no willingness to repay, and their repayment ability has basically been lost.

In order to facilitate the performance measurement of the model, the data set should be divided before the feature engineering of the sample data. Divide the credit user data according to the ratio of 6:4. The training set is used for model building, and the test set is used for model testing and performance measurement.

This section uses RF (random forest) to build the risk early warning model after lending. In the data set, each region is divided into two regions by inputting spatially recursive data, and the corresponding binary tree is constructed. Select segmentation point  $s$  and segmentation variable  $k$ , and solve through the optimal segmentation point:

$$\min_{k,s} \left[ \sum_{x_i \in R_1(k,s)} (y_i - c_1)^2 + \sum_{x_i \in R_2(k,s)} (y_i - c_2)^2 \right]. \quad (6)$$

Traverse the segmentation variable  $k$  to find the segmentation point, so as to minimize the objective function.

There are several important parameters in RF model, which are leaf nodes, minimum number of samples, number of trees, maximum depth of trees, maximum number of features, and minimum number of samples divided by internal nodes of trees. The parameter adjustment basis is used to measure whether the parameter values selected by the above grid search method are good or bad. The mean square error

is selected as the parameter adjustment basis. The mean square error (MSE) is a simple way to measure “mean error.” The formula of average error is

$$\text{MSE} = \sum_{i=1}^n \frac{1}{n} (f(x_i) - y_i)^2. \quad (7)$$

According to the above formula, the essence of mean square error is another manifestation of mean error, and mean error is the superposition of all error values as follows:

$$D(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (8)$$

According to the above formula, the average error describes the average of the sum of squares of  $x_i$ ,  $\bar{x}$  difference.

On the one hand, the increasing data mining ability of ML has seriously damaged the privacy protection of consumers, and the phenomenon of personal information disclosure is common. On the other hand, consumers should improve their understanding of risk management-related businesses. With the active innovation of new financial products and services, consumers should know the relevant risk points in time and improve their risk identification and prevention ability.

Relevant financial institutions should provide consumers with real-time information and continuous education, improve the transparency of business processes, be responsible for the authenticity and security of financial products, and guide and help consumers to understand risk factors and correctly report security issues.

### 3. Results Analysis and Discussion

One of the main responsibilities of the banking supervision institution is to minimize the losses caused by the failure of the banking operation. In order to fulfill this responsibility, the banking supervision institution evaluates the financial and operating conditions of the bank and takes prompt measures to correct the problems after discovering them. In the process of evaluating banks, the supervisory authorities use the combination of on-site supervision and off-site supervision system. At the same time of on-site inspection, supervisors use computer-based supervision system to supervise financial institutions off-site. These regulatory systems generally analyze the financial information submitted to regulators by various financial institutions every quarter.

In this paper, the time span method commonly used in the financial field is adopted to divide the data set. The data spanning half a year is divided into training set, verification set, and test set according to the month, and then the data set is used for modeling. The 50% cross-validation method is used to reduce the randomness of the algorithm and make the detection result more stable. After all the models are trained, the comparison results are shown in Figure 5.

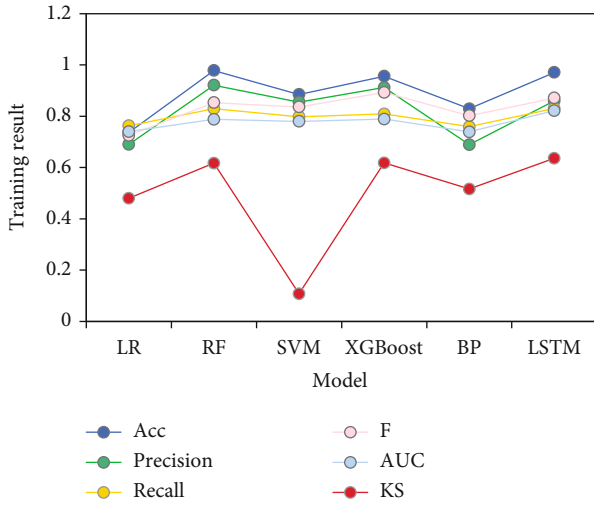


FIGURE 5: Comparison of final results of different models.

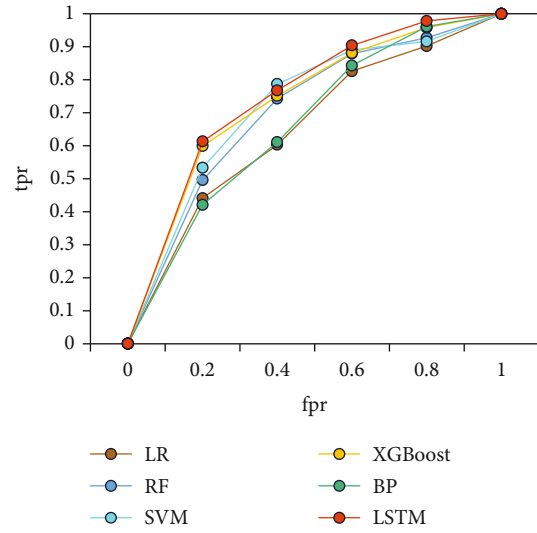


FIGURE 7: ROC curve.

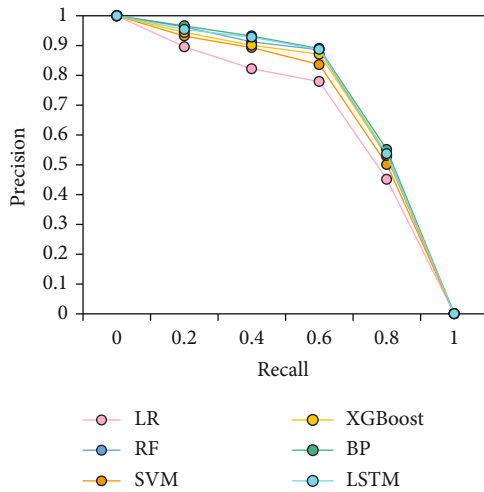


FIGURE 6: P-R curve.

According to the analysis in Figure 5, as a simple classifier, logistic regression is more popular than RF, SVM, BP neural network, etc., and it is proved that the more effective classification algorithm is less effective in research. Then, XGBoost has the highest  $F$  value, while LSTM model has higher AUC and KS indexes. Compared with the two models, LSTM model is better than XGBoost model currently used in production environment.

To sum up, this paper adopts LSTM model as the classification model of financial risk control pre-loan approval.

In this case, if the demand is not clear, whether the built data warehouse can meet the demand and can be applied often has to wait until the upper application and management information system is established before it can be discovered. The data warehouse whose requirements are not clear will be accepted, but it cannot be further utilized. This may be that the data warehouse is only suitable for one application mode and does not support other application modes.

TABLE 2: Analysis results of key variables.

Key variable	Importance of variable
Bank card utilization	0.0387
Borrower APR	0.0312
Borrower rate	0.0288
Borrower state	0.0549
Credit grade	0.0716

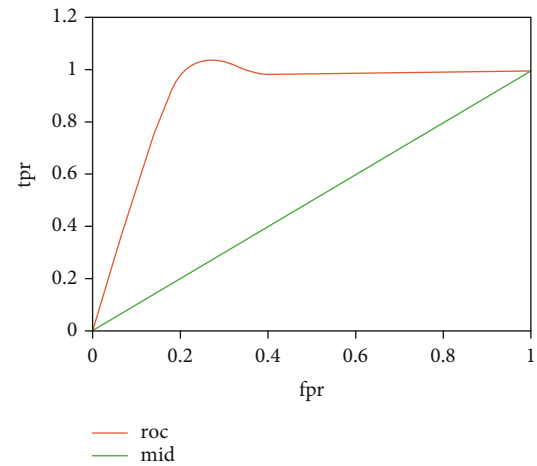


FIGURE 8: ROC curve of post-loan model.

In order to further verify the effectiveness of LSTM as a classification model, the P-R diagram and ROC diagram are drawn for reference, as shown in Figures 6 and 7.

From the P-R diagram, it can be seen that all the models except the logistic regression model perform well and the RF algorithm performs even better. However, from the ROC diagram, it can be seen that LSTM has the largest area under the ROC curve, which shows that LSTM model has a better classification effect. Therefore, the LSTM model in this paper

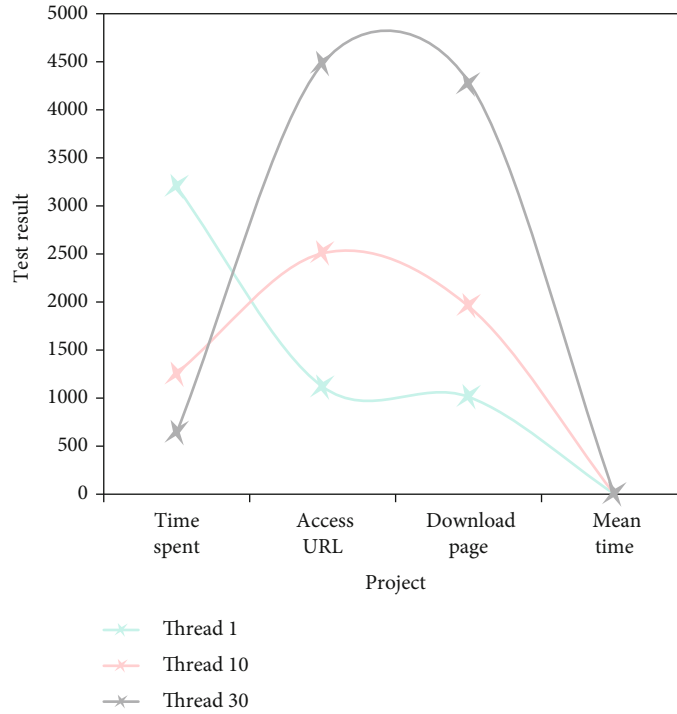


FIGURE 9: Test results of crawler performance under different number of threads.

has the best comprehensive effect and is suitable for use in production environment.

Based on the above theory, the data set can be constructed according to the model algorithm mentioned above. For the training set imported into the corresponding RF library, the model is built with RF default parameters, and the risk predictability of the model is 0.88. The test set selects the following five key variables as shown in Table 2.

In order to better describe the predictive ability of the model, the ROC curve of the post-loan model can be drawn to describe its predictive ability to the post-loan training set data. The ROC curve of post-loan model is shown in Figure 8.

According to the above ROC curve, the prediction ability of the model in the test set can be obtained. The post-loan risk early warning model constructed in this paper has strong risk prediction ability.

The design of the system is based on the idea of software engineering, so the coupling between the whole system modules is very low, and even a model can be run independently. In order to verify the performance of the financial risk control early warning model, it is necessary to test the key modules. This section tests the crawler module from the aspects of web page importance and multithreading. Figure 9 shows the performance results of the crawler tested separately by changing the number of threads.

As can be seen from Figure 9, the performance of crawler has been greatly improved under multithreading. The crawler module has the effect of quick response, and at the same time, it has the characteristics of low proportion of processor and memory.

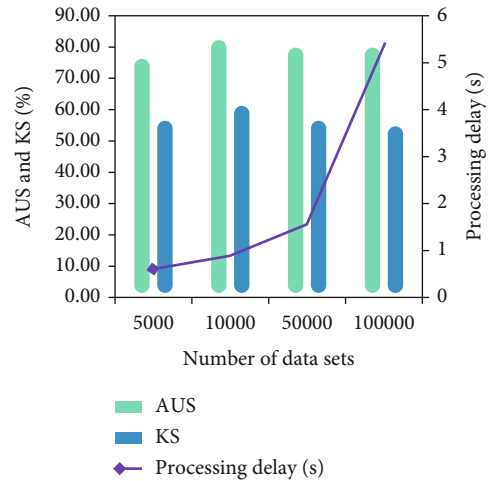


FIGURE 10: Test results of the model.

In the early stage of data warehouse system project start-up, managers at all levels need to be fully trained. First, they can strengthen the awareness of management change and technological innovation, make them have a comprehensive understanding of enterprise informatization, and fully understand and prepare for the problems that are easy or possible to encounter in the project process. In fact, the process of investigation is a process of communicating with users and reaching an agreement on ideas, a process of exchanging the experience of consulting companies and enterprises, and a process of diagnosing users and evaluating the original management methods.



The test of the model mainly tests the accuracy of the model under large samples. This section tests the credit risk control early warning system based on logistic regression algorithm. Considering the performance of the whole credit risk control model, use spark to read files from the cache module, process and analyze the read data, and then access the processed data into the model. The test results of the model are shown in Figure 10.

According to the figure above, it can be seen that when the amount of data is 10,000 to 50,000, the accuracy of the model is relatively high when it is input into the model. When the amount of data is relatively low, the overall throughput of the model is still very high. When the data set reaches 100,000, the processing time delay of the model is longer, and the accuracy of the model also decreases. Due to the limitation of the experimental environment, there will be a certain time delay when the cluster transmits and processes the data set. Because the real-time requirement of this system is not very high, it can run the system with a certain time delay.

#### 4. Conclusion

In the face of the rapidly developing financial industry, the methods and means of financial supervision in China are already at an obvious disadvantage, and the regulatory authorities are already in a rather awkward position. In order to change this situation, the financial supervision information system should be built on the basis of the overall plan of the information system of the People's Bank of China. Applying artificial intelligence and knowledge engineering technology to the financial supervision information system can greatly improve the processing capacity and flexibility of the system, enable the system to adapt to the ever-changing economic and financial environment, and improve the financial supervision level of China's central bank. After deeply understanding the characteristics of ML, a financial risk control model based on LSTM is established. Compared with the traditional ML model Xgboost, the final effect of LSTM model proposed in this paper has higher AUC and KS indexes, and the actual cost of adjusting parameters is shorter, and the model update iteration is more convenient, which is suitable for use in production environment. This paper studies the application of RF in post-loan credit early warning model and adjusts its parameters to make the model achieve the optimal effect. Finally, the accuracy of predicting post-loan risk through the model is about 90%.

Analyze different data sources submitted by users in multiple dimensions. The users can use images and words as the input of the model and can use deep learning related technologies for analysis. Therefore, a cross-dimensional and multi-angle credit risk control early warning model can be constructed.

#### Data Availability

The data used to support the findings of this study are included within the article.

#### Conflicts of Interest

The authors declare that they have no competing interest.

#### Acknowledgments

This paper is supported by the Discipline Co-Construction Research Project of Philosophy and Social Sciences of the 14th Five-Year Plan of Guangzhou in 2021: "Research on collaborative governance mechanism of digital economy in Guangdong-Hong Kong-Macao Greater Bay Area under the horizon of intergovernmental cooperation", Project No: 2021GZGJ67; the Humanities and Social Sciences Project of Henan Education Department: "Research on the Identification, Measurement and Control of Internet Financial Risk in the '14th Five-Year' Period", Project No. 2022-ZZJH-247; the Research Project of Henan Science and Technology Think Tank: "Research on the Development Strategy of Science and Technology Finance in Henan Province, Project No. HNKJZK-2022-22B; the Xinxiang Soft Science Research Project: Research on the Long-term Mechanism of Science and Technology Finance to Help the Development of Enterprises in Xinxiang City, Project No. RKX2021010; and the Philosophical and Social Science Planning Project of Henan Province: Long-term Mechanism of Green Finance Promoting the Development of Energy Conservation and Environmental Protection Industry in Henan Province, Project No. 2018CJJ085.

#### References

- [1] A. Jain, Y. Sharma, and K. Kishor, "Solid state technology financial supervision and management system using ml algorithm," *Solid State Technology*, vol. 63, no. 6, pp. 18974–18982, 2020.
- [2] Y. Cao, "Internet financial supervision based on machine learning and improved neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 47, pp. 1–12, 2020.
- [3] B. Viala, M. Villiet, A. Redor et al., "Using the clinical information system and self-supervision to rationalize the need for antibiotic stewardship: an interventional study in a 2000-bed university hospital," *International Journal of Antimicrobial Agents*, vol. 57, no. 1, article 106233, 2021.
- [4] X. Chen and N. Metawa, "Enterprise financial management information system based on cloud computing in big data environment," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5223–5232, 2020.
- [5] M. Matthias, "The gold standard, fiscal dominance and financial supervision in Greece and South-East Europe, 1841–1939," *European Review of Economic History*, vol. 1, p. 1, 2020.
- [6] J. Aparicio, M. A. Duran, A. Lozano-Vivas, and J. T. Pastor, "Are charter value and supervision aligned? A segmentation analysis," *Journal of Financial Stability*, vol. 37, pp. 60–73, 2018.
- [7] W. Zhuo, C. Ding, Y. Xiong, and J. Peng, "Construction of legal system for maritime international trade financial supervision cooperation," *Journal of Coastal Research*, vol. 103, no. -sp1, p. 143, 2020.
- [8] R. Wang, C. Yu, and J. Wang, "Construction of supply chain financial risk management mode based on internet of things," *IEEE Access*, vol. 7, no. 99, pp. 110323–110332, 2019.

- [9] W. C. Ching, A. M. Alashwal, and F. Rahim, "Risk disclosure and performance of Malaysian construction public listed companies," *MATEC Web of Conferences*, vol. 266, no. 1, p. 3012, 2019.
- [10] S. M. Carta, S. Consoli, A. S. Podda, D. R. Recupero, and M. M. Stanciu, "Ensembling and dynamic asset selection for risk-controlled statistical arbitrage[J]," *IEEE Access*, vol. 9, pp. 29942–29959, 2021.
- [11] X. Sun and Y. Lei, "Research on financial early warning of mining listed companies based on BP neural network model[J]," *Resources Policy*, vol. 73, no. 2, article 102223, 2021.
- [12] H. Ma, L. Lin, H. Sun, and Y. Qu, "Research on the dependence structure and risk spillover of internet money funds based on C-vine copula and time-varying t-copula," *Complexity*, vol. 2021, 11 pages, 2021.
- [13] A. N. Häusler, C. M. Kuhnen, S. Rudolf, and B. Weber, "Preferences and beliefs about financial risk taking mediate the association between anterior insula activation and self-reported real-life stock trading," *Scientific Reports*, vol. 8, no. 1, pp. 1–13, 2018.
- [14] J. Bi and J. Cai, "Optimal investment–reinsurance strategies with state dependent risk aversion and VaR constraints in correlated markets[J]," *Insurance: Mathematics and Economics*, vol. 85, pp. 1–14, 2019.
- [15] E. Tang and S. Xu, "Pricing of embedded options in bank deposits and loans based on jump-diffusion interest rate model," *Complexity*, vol. 2021, Article ID 9975536, 15 pages, 2021.
- [16] L. Li and H. Li, "Analysis of financing risk and innovation motivation mechanism of financial service industry based on Internet of Things[J]," *Complexity*, vol. 2021, Article ID 5523290, 9 pages, 2021.
- [17] D. Chen, "Risk assessment of government debt based on machine learning algorithm," *Complexity*, vol. 2021, Article ID 3686692, 12 pages, 2021.
- [18] X. Huang, J. Sun, and X. Zhao, "Credit risk assessment of supply chain financing with a grey correlation model: an empirical study on China's home appliance industry[J]," *Complexity*, vol. 2021, Article ID 9981019, 12 pages, 2021.
- [19] A. Bensoussan, S. R. Celine, and Z. Yan, "A mean-variance approach to capital investment optimization[J]," *SIAM Journal on Financial Mathematics*, vol. 10, no. 1, pp. 156–180, 2019.
- [20] F. O. Catak and A. F. Mustacoglu, "CPP-ELM: cryptographically privacy-preserving extreme learning machine for cloud systems," *International Journal of Computational Intelligence Systems*, vol. 1, no. 1, pp. 33–44, 2018.
- [21] J. Wyrobek, "Application of machine learning models and artificial intelligence to analyze annual financial statements to identify companies with unfair corporate culture," *Procedia Computer Science*, vol. 176, pp. 3037–3046, 2020.
- [22] A. Sharma and S. K. Sahay, "Group-wise classification approach to improve android malicious apps detection accuracy," *International Journal of Network Security*, vol. 21, no. 3, pp. 409–417, 2019.
- [23] W. Li, S. Ding, Y. Chen, H. Wang, and S. Yang, "Transfer learning-based default prediction model for consumer credit in China[J]," *Journal of Supercomputing*, vol. 75, no. 2, pp. 862–884, 2019.
- [24] Y. Song and Y. Peng, "A MCDM-based evaluation approach for imbalanced classification methods in financial risk prediction," *IEEE Access*, vol. 7, pp. 84897–84906, 2019.