Application of Financial Big Data in Systemic Risk Warning

Chen Qiu

Fordham University, New York, NY 11101, USA

Abstract

With the rapid development of the financial market, the volume of financial data has become increasingly large and complex, and the advantages of using big data methods to analyze systemic financial risks have become increasingly prominent. This paper deeply discusses the application of financial big data in systemic risk early warning, aiming to provide theoretical support and practical guidance for improving financial risk prevention and control ability.Traditional systemic risk analysis mainly focuses on individual financial institutions, which has limitations and cannot fully capture the correlation risk exposure and the potential domino effect threat to the stability of financial markets. It is therefore crucial to assess exposures at the level of the financial system as a whole. However, the complexity of the financial system is increasing, and the financial data faced by risk analysis is characterized by huge volume, complex types and strong correlation, which brings challenges to the prevention and control of financial risks.Big data methods, including machine learning, data visualization, information extraction and other technologies, provide new solutions for systemic financial risk early warning. Information extraction technology helps to extract valuable information from unstructured text data, such as investor behavior, sentiment changes, etc. These data can reflect changes in the financial market in real time, and is an important source of supplementary information for systemic financial risk analysis. Machine learning algorithms can integrate multi-source data, automatically learn and identify different credit risk levels, and improve the accuracy and efficiency of risk assessment. Data visualization technology can clearly and intuitively show financial micro-data in the form of graphs and charts, and help analysts quickly understand the relationship and trend between data.This paper analyzes the application cases of big data technology in financial contagion mechanism analysis, financial supervision and other aspects, and demonstrates the practical effect of big data methods in systemic risk early warning. By integrating financial micro-data of various financial sectors and even different countries, and analyzing financial risks as a whole based on the data, big data methods can more comprehensively capture the potential threat of systemic risks and provide timely and accurate early warning information for financial regulators.At the same time, this paper also points out the challenges faced by big data in the early warning of financial systemic risk, such as data security and privacy protection, technical threshold and capital investment, and puts forward corresponding solutions. It is emphasized that financial institutions need to strengthen data protection measures, establish a sound risk prevention and control system, increase scientific and technological investment and personnel training, and ensure that the application of big data technology in the financial industry achieves positive results.To sum up, financial big data plays an important role in the early warning of systemic risks and provides strong support for improving the ability to prevent and control financial risks. In the future, with the continuous development and improvement of big data technology, its application prospects in financial risk early warning will be broader.

Keywords

Financial big data, systemic risk, risk warning, machine learning, data mining.

1. Introduction

1.1. Research Background and Significance

Definition: Systemic risk, also known as market risk or macro risk, refers to the changes in the entire financial system or macroeconomic factors that lead to sharp fluctuations in the price of most assets in the market, thus causing extensive and far-reaching impacts on financial institutions and financial markets. Such risks are often contagious, and problems at one financial institution can quickly spread throughout the financial system.

Features:

Universality: The impact is broad, involving multiple financial institutions and markets.

Contagion: Risk events are prone to spread from one institution or market to another.

Non-diversification: It cannot be completely eliminated by portfolio diversification.

Macro correlation: closely related to the macroeconomic environment, susceptible to policy, economic cycle and other factors.

Damage:

Economic losses: resulting in the depreciation of financial institutions' assets and heavy losses for investors.

Confidence crisis: Undermining market confidence, triggering financial panic and capital outflows.

Credit crunch: Banks reduce loans to avoid risks, exacerbating the financing difficulties of the real economy.

Policy pressure: Forcing the government to take emergency measures may trigger an increased burden on public finances.

Figure 1. Big Data

1.1.1. Limitations of Traditional Systemic Risk Early Warning Methods

Traditional systematic risk warning methods mainly include statistical model based on historical data, expert scoring method, stress test and so on. These methods, while effective to a certain extent, have the following limitations:

Data limitations: Relying on limited historical data makes it difficult to capture new changes and complexities in the market.[1]

Model assumptions: Built based on specific assumptions that may not accurately reflect the dynamics of the real world.

Early warning lag: Often only after the risk exposure to discover the problem, lack of foresight.

Ignoring nonlinear relationships: It is difficult to capture the complex interactions and nonlinear relationships within the financial system.

1.1.2. Opportunities Brought by Financial Big Data for Systemic Risk Early Warning

Financial big data, including high-frequency trading data, social media information, macroeconomic indicators, etc., provides new perspectives and tools for systemic risk warning:

Real-time monitoring: Big data processing technology can realize real-time data collection and analysis, and improve the timeliness of early warning.

Multi-dimensional analysis: Combining multiple data sources allows for a more comprehensive understanding of risk sources and transmission paths.

Machine learning: Algorithms learn patterns in historical data to improve the accuracy and generalization of early warning models.

Predictive power: Big data analytics can reveal potential trends and provide a forward-looking perspective for risk warning.

1.2. Research Status at Home and Abroad

1.2.1. Foreign Research Status

Foreign research on the application of financial big data to systemic risk early warning started earlier, mainly focusing on the use of big data technology for risk identification, model construction and predictive analysis. For example, research institutions in the United States, Europe and other places have developed a variety of early-warning systems based on big data, such as using social media sentiment analysis to predict market fluctuations, and monitoring market anomalies through high-frequency trading data.

1.2.2. Domestic Research Status

In recent years, domestic scholars have also begun to actively explore the application of financial big data in systemic risk early warning. The research covers the construction of big data platform, optimization of data mining algorithm, establishment of early warning model, etc. Despite the late start, domestic research has made remarkable progress in combining the characteristics of the local market and integrating traditional financial theories with big data technology.

1.3. Research Contents and Methods

1.3.1. Research Content

This study aims to explore the application of financial big data in systemic risk warning, including:

Analyze the characteristics of financial big data and its impact on systemic risk warning.

Assess the applicability and challenges of existing big data technologies in systemic risk early warning.

The system risk early warning model based on big data is constructed and empirically tested.

Optimization strategies and suggestions are proposed to promote the effective application of financial big data in systemic risk early warning.

1.3.2. Research Methods

The following methods were used in this study:

Literature review: Sort out relevant researches at home and abroad, summarize the existing achievements and shortcomings.

Case study: Select typical financial events and analyze the actual effect of big data early warning.

Empirical research: Collect and process financial big data, build an early warning model, and verify its forecasting ability.

Comparative analysis: The advantages and disadvantages of traditional early warning methods and big data early warning methods are compared.

2. Theoretical Basis of Financial Big Data and Systemic Risk Early Warning

2.1. Overview of Financial Big Data

2.1.1. Definition, Characteristics and Sources of Financial Big Data

Definition: Financial big data refers to a collection of data that covers massive, diverse and rapidly generated financial related information. It includes not only the traditional financial transaction records, such as stock trading, bond issuance and trading, but also the emerging Internet financial activity data, such as mobile payment, online lending and so on.[2]

Features:

Massive: The scale of financial big data is huge, including a large number of transaction records, customer information, market data, etc.

Diversity: Data comes from a wide range of sources and types, including structured data (such as transaction records), semi-structured data (such as emails), and unstructured data (such as social media messages).

High speed: The generation speed of financial data is fast, and efficient data processing technology is needed for real-time analysis and processing.

Source:

Financial institutions: transaction records, customer information, financial statements, etc. generated by financial institutions such as banks, securities companies, and insurance companies.

Financial market: transaction data and price fluctuation information of financial markets such as stock market, bond market, foreign exchange market, etc.

Internet financial platform: data generated by Internet financial activities such as mobile payment, online lending, and crowdfunding.

Macroeconomic data: macroeconomic indicators such as GDP growth rate, inflation rate and unemployment rate released by the government.

2.1.2. Processing Technology of Financial Big Data

The processing technology of financial big data mainly includes data acquisition, pre-processing, storage, analysis and visualization.

Data acquisition: Data collection through various means, such as API interfaces, crawler technology, data exchange platforms, etc.

Data preprocessing: The original data is cleaned, reduplicated, filled with missing values, and converted into data formats to improve the quality and availability of data.

Data storage: Use distributed storage systems (such as Hadoop, Spark) and database technologies to efficiently store large-scale data.

Data analysis: Use data mining, machine learning and other technologies to conduct in-depth analysis of financial data and extract valuable information and insights. Analysis methods include descriptive analysis, predictive analysis, anomaly detection and model evaluation.

Data visualization: Display analysis results in graphs, tables, etc., to help people understand and analyze data more easily.[3]

2.2. Systemic Risk Early Warning Theory

2.2.1. Formation Mechanism of Systemic Risk

The formation mechanism of systemic risk involves many aspects, including the vulnerability of the real economy, the excessive correlation within the financial sector, and the volatility of the financial market. When there are vulnerabilities in the real economy, such as excessive debt and overcapacity, they will be transmitted to the financial system, leading to instability in the financial sector. At the same time, excessive linkages between financial institutions within the financial sector can also amplify risks, exposing the entire financial system to the risk of collapse.

2.2.2. Transmission Mechanism of Systemic Risk

The transmission mechanism of systemic risk is mainly realized through the transmission of credit risk among financial institutions, the transmission of price fluctuations among financial markets and the transmission of investor sentiment. When a financial institution goes wrong, its credit risk is passed on to other financial institutions that do business with it, setting off a chain reaction. At the same time, the price fluctuation of the financial market will also affect the sentiment and behavior of investors, thus intensifying the contagion and spread of risks.

2.2.3. Warning Index System of Systemic Risk

The index system of systemic risk warning includes macroeconomic index, financial market index, financial institution index and global factor index.

Macroeconomic indicators: such as GDP growth rate, inflation rate, unemployment rate, etc., reflect the overall state of economic operation.

Financial market indicators: such as the volatility of the stock market, the yield curve of the bond market, the exchange rate fluctuations of the foreign exchange market, etc., measure the stability and volatility of the financial market.

Financial institution indicators: such as the bank's capital adequacy ratio, liquidity coverage ratio, loan default rate, etc., to assess the capital adequacy, liquidity, risk exposure and other conditions of financial institutions.

Global factor indicators, such as global economic growth rates, international trade and investment flows, reflect the dynamics of the global economy and financial markets.

2.3. Application Advantages of Financial Big Data in Systemic Risk Early Warning

Financial big data has significant application advantages in systemic risk early warning, mainly reflected in the following aspects:

Real-time monitoring and early warning: Financial big data processing technology can achieve real-time data collection and analysis, timely detection of potential risks and early warning signals.

Multi-dimensional analysis: Combining multiple data sources and types of data can provide a more comprehensive understanding of risk sources and transmission paths, and improve the accuracy of early warning.

Intelligent decision support: The use of machine learning and artificial intelligence and other technologies, deep mining and analysis of financial big data, to provide decision-makers with intelligent risk prevention and control recommendations.

Improve the efficiency of risk management: Through the application of big data technology, financial institutions can respond to risk events more quickly, take effective measures to reduce losses, and improve the overall efficiency of risk management.

To sum up, financial big data plays an important role in systemic risk early warning, which can provide strong decision-making support for financial institutions and regulatory authorities, and ensure the stability and security of financial markets.[4]

3. Construction of Systemic Risk Early Warning Model based on Financial Big Data

3.1. Data Preprocessing

Data preprocessing is a key step to build financial risk early warning system, which involves data cleaning, data integration and data transformation.

3.1.1. Data Cleaning

Data cleaning refers to the pre-processing and cleaning of the original data to eliminate noise, missing values, outliers and other problems in order to conduct effective data analysis and model construction. The main steps of data cleaning include:

Data check: Perform a preliminary check of the data to detect whether there are missing values, outliers, duplicate values, etc.

Missing value detection: Correlation functions can be used to detect missing values.

Outlier detection: Methods such as Z-score and IQR (interquartile) can be used to detect outliers.

Duplicate value detection: Correlation functions can be used to detect duplicate values.

Data cleaning: Take appropriate measures to clean the detected problem.

Fill missing values: You can use correlation functions to fill missing values, including mean fill, median fill, mode fill, interpolation fill, etc.

Delete outliers: Deletes data that exceeds the normal range based on the result of the outlier detection.[5]

Deduplication: Use correlation functions to remove duplicate data records.

Data transformation: The data is transformed to fit the requirements of the model.

Data type conversion: You can use related functions to convert data types, such as a string type to a numeric type.

Data format conversion: Related functions can be used to convert data formats, such as from a wide format to a long format, or from a long format to a wide format.

Data consolidation: Combining data from different sources to create a complete data set. The main methods of data consolidation include merging data and joining data.

Merge data: You can use correlation functions to merge data, such as combining two or more datasets by row or column.

Join data: You can use correlation functions to join data, such as joining two data sets based on one or more common fields.

Figure 2. Fintech

3.1.2. Data Integration

Data integration is the process of combining and integrating data from different data sources to form a unified view of the data. In the financial field, data integration usually involves the integration of multiple data types such as market data, transaction data, financial data, economic data, social media data, customer data and risk management data. The main challenges of data integration include data diversity, real-time requirements, and data privacy concerns. To address these challenges, financial institutions can adopt standardized data formats and interfaces, use data transformation tools for format conversion and cleansing, and establish a data governance framework to ensure data quality and consistency.

3.1.3. Data Conversion

Data transformation is the use of linear or nonlinear mathematical transformation method to compress multidimensional data into less dimensional data, in order to eliminate the differences in the performance of data in time, space, attributes and accuracy. The main methods of data transformation include data smoothing, data generalization, data reduction and so on.

Data smoothing: Used to remove noisy data, common methods include Bin method, clustering method and regression method.

Data generalization: A more abstract way to represent data, such as using category labels or ranges instead of specific numerical values.

Data reduction: Looking for useful features that depend on finding the target data to reduce the size of the data. Data reduction can minimize the amount of data on the premise of keeping the original appearance of data as much as possible, so as to improve the efficiency and effect of data processing.

3.2. Feature Engineering

Feature engineering refers to the transformation and combination of raw data to create new features to better represent and predict problems. Feature engineering is a key step in machine learning and has a significant impact on the performance and accuracy of models.

3.2.1. Feature Selection

Feature selection is the selection of features that are valuable to model prediction based on the relevance and importance of data. The main methods of feature selection include filtering, wrapping and embedded.

Filter: Sort the features first, and then select the top features. The common ranking basis includes correlation coefficient, Chi-square statistics, mutual information and so on.

Wrap: The feature subset is taken as input, the performance of the feature subset is evaluated by training the model, and the feature subset with the best performance is selected. Common wrapping methods include recursive feature elimination (RFE) and forward/backward search.

Embedded: Features are automatically selected during model training. Some machine learning algorithms (such as Lasso regression, decision trees, etc.) have embedded feature selection capabilities that automatically select important features during training.

3.2.2. Feature Extraction

Feature extraction is the process of extracting useful information from the original data and transforming it into a new feature representation. The main methods of feature extraction include principal component analysis (PCA), linear discriminant analysis (LDA), independent component analysis (ICA), etc.

Principal component analysis (PCA) : The original data is transformed into a new coordinate system by a linear transformation so that the variance on the first axis in the new coordinate system is the largest (i.e., the first principal component), the variance on the second axis is the second largest (i.e., the second principal component), and so on. PCA can be used to reduce and remove noise.

Linear discriminant Analysis (LDA) : is a supervised dimensionality reduction technique that attempts to find a linear combination (i.e., the projection direction) that maximizes the ratio of the inter-class variance to the intra-class variance. LDA can be used for classification and dimensionality reduction.

Independent Component Analysis (ICA) is a statistical and computational technique used to extract latent, non-Gaussian, mutually independent components from multivariate data. ICA can be used for signal separation and feature extraction.[6]

3.3. Early Warning Model Selection

Early warning model is the core component of financial risk early warning system, which is used to predict the future risk situation according to the input characteristic data. Choosing a suitable early warning model is very important to improve the accuracy and reliability of the system.

3.3.1. Introduction to Common Machine Learning Models

Logistic regression: is a widely used linear classification model that classifies input features by mapping them to probability values in the interval (0,1). Logistic regression has the advantages of simple and easy to understand, fast calculation speed, and easy to implement, but it may not be effective for nonlinear data.

Support Vector machine (SVM) : is a binary classification model based on the principle of maximum margins, which tries to find a hyperplane to separate different classes of data. SVM has a good classification effect for high-dimensional data and small sample data, but the computational complexity is high, and some strategies need to be used to extend the multi-classification problem.

Random forest is an ensemble learning method based on decision trees, which improves the accuracy and stability of the model by constructing multiple decision trees and synthesizing their prediction results. Random forest has the advantages of strong anti-overfitting ability, processing high-dimensional data, and being insensitive to missing data, but it has high computational complexity and may not be effective for some noisy data.

In addition to the above models, there are machine learning models such as neural networks, naive Bayes, and K-nearest neighbors that can also be used for financial risk early warning.

3.3.2. Basis for Model Selection

When choosing an early warning model, the following factors need to be considered:

Data characteristics: including the type of data (such as numerical type, category type, etc.), scale (such as data volume size), distribution (such as normal distribution, skewed distribution, etc.), etc. Different data characteristics may suit different models.

Model performance: including accuracy, stability, computational complexity, etc. Accuracy is an important index to measure the predictive ability of the model. Stability refers to whether the model's performance is consistent across different data sets. The computational complexity affects the speed of model training and prediction.

Business requirements: including the timeliness, accuracy and interpretability of early warning. Different business requirements may suit different models. For example, for an early warning system that requires a fast response, a model with a faster computing speed can be selected. For the highly accurate early warning system, the model with better performance can be selected. For scenarios where the warning results need to be interpreted, an interpretable model can be chosen.

3.4. Model Training and Optimization

Model training and optimization are the key steps to improve the performance and accuracy of early warning models. By training the model and adjusting its parameters, it can be better adapted to the data and predict future risk scenarios.

3.4.1. Model Training

Model training refers to the process of using training data sets to train a model and adjust its parameters. In the training process, it is necessary to select appropriate optimization algorithms (such as gradient descent, random gradient descent, Adam, etc.) and loss functions (such as cross entropy loss, mean square error, etc.) to guide the training of the model. At the same time, it is also necessary to set the appropriate training rounds, learning rate and other hyperparameters to control the training process.

3.4.2. Model Evaluation Index

Model evaluation index is an important standard to measure model performance. Commonly used model evaluation indicators include accuracy, recall rate, F1 score, OC-ROC curve, etc. These indicators can reflect the performance of the model from different angles. For example, the accuracy rate measures the proportion of correct predictions of the model. The recall rate measures the proportion of real positive examples covered by the model; F1 score is the harmonic average of accuracy and recall rate, which is used to measure the performance of the model comprehensively. The OC-ROC curve reflects the performance of the model under different thresholds.

3.4.3. Model Optimization Method

Model optimization methods include parameter optimization, feature engineering optimization, ensemble learning and so on.

Parameter tuning: By adjusting the parameters of the model (such as learning rate, regularization coefficient, etc.) to improve the performance of the model. The commonly used parameter tuning methods include grid search, random search and Bayesian optimization.

Feature engineering optimization: improving feature selection, feature extraction and other methods to improve the quality of features, so as to improve the performance of the model. Feature engineering optimization can include adding new features, deleting redundant features, transforming features and so on.

Ensemble learning: Improves overall performance and accuracy by combining predictions from multiple models. Common ensemble learning methods include Bagging, Boosting, and Stacking. These methods can improve the overall performance by reducing the variance or bias of the model.

To sum up, data preprocessing, feature engineering, early warning model selection and model training and optimization are important steps to build a financial risk early warning system. Through reasonable data preprocessing and feature engineering, the quality of data and the effectiveness of features can be improved. The performance and accuracy of the model can be improved by selecting a suitable early warning model and optimizing it. These steps together constitute the core part of the financial risk early warning system, which provides effective risk early warning and decision support for financial institutions.

4. Empirical Analysis

4.1. Data Sources and Description

The financial Risk warning system is based on a wide range of data sources and aims to provide comprehensive coverage of various factors that may affect financial stability. These data mainly include:

Macroeconomic data: such as GDP growth rate, inflation rate, interest rate, exchange rate, etc., these data reflect the state of the overall economic environment and have a direct impact on the stability of the financial system.[7]

Financial market data: including stock market data, bond market data, real estate market data, etc., these data provide direct evidence of the operation of the financial market and help identify abnormal market fluctuations and potential risks.

Financial institution data: including the bank's balance sheet, income statement, credit data, etc., these data reflect the operating conditions and risk exposure of financial institutions.

Big Data Sources:

E-commerce big data: Transaction data such as Alibaba, Taobao, Tmall and other platforms can be used to assess the credit status of individuals or businesses.

Credit card big data: credit card application, use, repayment and other data, as a reference for credit rating.

Social network Big Data: Data based on social platforms can be used to analyze user behavior patterns and credit risk.

Third-party payment big data: payment direction, amount, frequency and other data can be used for credit analysis and risk assessment.

Life service website big data: such as water, electricity, gas, property fees and other payment data, reflecting the basic living conditions and credit level of individuals or enterprises.

4.2. Analysis of Empirical Results

4.2.1. Model Prediction Results

By constructing appropriate early warning models (such as logistic regression model, machine learning model, etc.), and using historical data for training and verification, we can get the model's prediction results for future financial risks. The forecast results are usually presented in the form of risk scores, probability values, or risk grades (such as "high risk," "medium risk," "low risk").

4.2.2. Model Performance Evaluation

Model performance evaluation is a key step to ensure the effectiveness of early warning system. By comparing the model prediction results with the actual occurrence of risk events, the accuracy, stability and reliability of the model can be assessed. Common evaluation metrics include accuracy, recall, F1 scores, etc. In addition, the performance of the model can be further verified by cross-validation, A/B testing and other methods.

4.2.3. Analysis of Key Risk Factors

Through the analysis of the model prediction results, we can identify the key factors that have the greatest impact on financial risks. These factors may include changes in macroeconomic indicators, fluctuations in financial markets, operating conditions of financial institutions, etc. Understanding these key factors can help develop targeted risk responses.

4.3. Comparative Analysis with Traditional Early Warning Methods

Wide data coverage: While traditional early warning methods rely on limited financial data and market indicators, big data early warning systems are able to integrate data from multiple sources to provide a more comprehensive view of risk.

Strong real-time: The big data early warning system can monitor the market dynamics and the operating conditions of financial institutions in real time, and discover potential risks in time.

High prediction accuracy: Through advanced models and algorithms, big data early warning systems can more accurately predict the probability and impact of financial risks.

High degree of automation: The big data early warning system can realize the automation and intelligence of risk early warning, reducing the possibility of manual intervention and misjudgment.

However, big data early warning systems also have some challenges and limitations, such as data quality control and optimization of model parameters. Therefore, in practical applications, it is necessary to flexibly adjust and optimize according to specific situations.

To sum up, the financial risk early warning system based on big data and advanced technology has significant advantages in data coverage, real-time, forecasting accuracy and automation degree, providing financial institutions and regulators with more effective risk management tools.

5. Conclusion and Prospect

5.1. Research Conclusions

Effectiveness of big data in risk early warning: Big data analysis technology can significantly improve the accuracy and timeliness of financial risk early warning. By integrating data from multiple sources, such as macroeconomic data, financial market data, financial institution data, and big data sources (such as e-commerce, credit cards, social networking sites, etc.), a more comprehensive and accurate risk warning model can be built.

Performance of early warning models: Trained and validated early warning models perform well in identifying financial risks. The model can accurately predict the probability of risk events and effectively distinguish different risk levels. In addition, the model also has high stability and robustness, and can maintain a stable early warning effect in different market environments and economic cycles.

Importance of key risk factors: Through analysis of the model's predictions, we identified a series of key factors that have a significant impact on financial risk. These factors include changes in macroeconomic indicators, fluctuations in financial markets, the operating conditions of financial institutions, and the credit status of individuals or enterprises reflected in big data. Understanding these key factors helps to develop targeted risk responses and improve risk management efficiency.

Comparison with traditional early warning methods: Compared with traditional early warning methods, the financial risk early warning system based on big data and advanced technology has significant advantages in terms of data coverage, real-time, prediction accuracy and automation degree. These advantages enable the new system to better adapt to the complex and changeable financial market environment and provide more effective risk management tools for financial institutions and regulators.

5.2. Policy Recommendations

Strengthen the construction of big data infrastructure: The government should increase investment in big data infrastructure and improve data collection, storage, processing and analysis capabilities. At the same time, promote the cooperation between financial institutions and big data enterprises, and jointly build a big data platform for financial risk early warning.

Improve the financial risk early warning mechanism: Establish and improve the financial risk early warning mechanism, and clarify the early warning indicators, early warning thresholds and early warning processes. The big data early warning system will be incorporated into the financial supervision system to realize the automation and intelligence of risk early warning.[8]

Strengthen risk education and training: raise the risk awareness of financial institutions and investors and strengthen risk education and training. Through the popularization of financial risk knowledge, improve the public's awareness of financial risks and prevention ability.

Promote fintech innovation: Encourage financial institutions and technology enterprises to strengthen cooperation and promote fintech innovation. We will use advanced technologies such as big data and artificial intelligence to enhance the intelligence level of financial services and improve the efficiency and accuracy of financial risk management.

5.3. Research Deficiencies and Prospects

Although this study has achieved certain results, there are still some shortcomings:

Data quality issues: Data quality varies due to the diversity of data sources. Some data may have problems such as missing, error or inconsistency, which may affect the accuracy and stability of the early warning model. In the future, the control and verification of data quality should be strengthened to improve the reliability and accuracy of data.

Model optimization problem: Parameter setting and algorithm selection of early warning model have important influence on prediction results. Although this study uses a variety of models and algorithms for comparison and analysis, there is still room for model optimization. In the future, more advanced models and algorithms can be further explored to improve the performance and accuracy of the early warning system.

Lack of interdisciplinary research: Financial risk early warning involves knowledge from multiple disciplinary fields, such as finance, economics, computer science, etc. Although this study combines the knowledge and methods of these disciplines, it still needs to be strengthened in terms of interdisciplinary research. In the future, cooperation and exchanges with other disciplines can be strengthened to promote interdisciplinary research and development in the field of financial risk early warning.

Looking to the future, with the continuous development of big data technology and the increasing complexity of financial markets, financial risk early warning systems will face more challenges and opportunities. We will continue to pay attention to the research dynamics and technological progress in this field, and constantly improve and optimize the early warning system to provide more effective risk management tools and support for financial institutions and regulators.

References

[1] Joachim W ,Nils M .Big Data in the financial industry: Applications, potential and regulatory approaches[J].Journal of Digital Banking,2024,9(1):6-18.

[2] Wang T .Internet financial risk prevention strategy under the background of big data[J].Computer Informatization and Mechanical System,2023,6(5):26-28.

[3] Wangsong X ,Jianjun C .Regulatory Mechanism of Financial Market Resource Management Driven by Big Data[J].Mobile Information Systems,2022,2022

[4] Wu C ,Liu J,Zhang H.Data Ecology and Accurate Portrait: Optimization of Credit Risk System for SMEs in Supply Chain Finance Based on Big Data Technology[J].Journal of Risk Analysis and Crisis Response,2022,11(4):

[5] Chen Z ,Dai Y ,Chen Y , et al.Prediction of mortality risk in critically ill patients with systemic lupus erythematosus: a machine learning approach using the MIMIC-IV database.[J].Lupus science & medicine,2025,12(1):

[6] Rokhideh M ,Fearnley C ,Budimir M .Multi-Hazard Early Warning Systems in the Sendai Framework for Disaster Risk Reduction: Achievements, Gaps, and Future Directions[J]. International Journal of Disaster Risk Science,2025,(prepublish):1-14.

[7] Biswas U ,Mahato S ,Joshi K P .Spatial prediction of forest fires in India: a machine learning approach for improved risk assessment and early warning systems[J].Environmental Science and Pollution Research,2025,(prepublish):1-23.

[8] Zhang X .Research on Financial Risk Early Warning System Based on Data Mining Technology[J].International Journal of High Speed Electronics and Systems,2025,(prepublish):