

Design of Investment Risk Control Model in Securities Market

Chang Liu^{1, +}, Rongjie Cai^{1, +, *}, Yuting Deng², Guanliang Chen¹, Yuqi Shan³, Min Luo⁴, Xin Guo⁵, Chuxiao Ye¹

¹Guangdong Ocean University, St. Petersburg Institute of Shipbuilding and Marine Technology, Zhanjiang 524088, Guangdong, China

²Guangdong Ocean University, College of Literature and News Communication, Zhanjiang 524088, Guangdong, China

³Guangdong Ocean University, School of Economics, Zhanjiang 524088, Guangdong, China

⁴Guangdong Ocean University, School of Management, Zhanjiang 524088, Guangdong, China

⁵Guangdong Ocean University, College of electronics and information engineering, Zhanjiang 524088, Guangdong, China

*Corresponding author: Rongjie Cai

+Co-first author: Chang Liu, Rongjie Cai

Abstract: With the rapid development of financial market in China, systematic risk control has become an important research topic in securities investment. In recent years, financial markets have experienced frequent and significant fluctuations around the world, bringing unprecedented challenges to investors, especially in the prevention and control of systemic risks. Systemic risk, i.e. market risk, is the risk that affects all assets and cannot be completely eliminated by diversifying the investment. In order to deal with this kind of risk, based on the financial mathematical modeling method, this paper constructs a comprehensive risk control model for securities market investment, in order to help investors and regulators to manage risks more scientifically and effectively, and to improve the stability of investment and the rationality of return. In the aspect of systematic risk prediction, this paper introduces machine learning technology to construct a systematic risk model with stronger prediction ability. By selecting multiple risk factors such as market volatility (MVI), liquidity (LSI), linkage (MCI) and volume pressure (VPI), a multi-factor prediction model is established, which uses machine learning algorithms such as XGBoost and stochastic forest to predict the change trend of future systematic risk. During the training process, the model improves the ability to capture market risks through feature engineering and parameter optimization, so that it can more effectively predict the possible market fluctuations in the future. At the same time, this paper makes a feature importance analysis on the model, which helps to identify the factors that have a key impact on the systematic risk and provides further basis for the application and improvement of the model. On the whole, the model proposed in this paper has high practicability and algorithm efficiency while effectively controlling the risks, which provides a scientific basis for the risk control in the securities market.

Keywords: Securities Market, Systematic Risk, Risk Measurement, Risk Control System, Machine Learning.

1. Introduction

1.1. Background of the Issue

With the rapid development of China's financial market, a market-oriented and diversified financial system has gradually formed. The flow and allocation of capital among different industries have significantly enhanced the vitality of the market. However, the openness and liquidity of the market also increase the probability of occurrence of systemic risks. Systemic risk is the risk that affects all assets and cannot be completely eliminated through diversification. In the 2008 global financial crisis, the destructive power of this risk was fully reflected, which brought serious losses to the global capital market. The China market has also experienced risk shocks, such as a sharp stock market shock in 2015 and sharp stock market volatility in early 2024. Therefore, effective control of systemic risk has become an issue that cannot be ignored in securities market investment.

1.2. Questions raised

In view of the demand for risk management in the securities market, this research uses mathematical modeling method to provide new ideas for the quantification and monitoring of

systematic risks.

The main problem of this study is to design a model that can effectively predict the systemic risk. By selecting a series of key indicators (such as volatility, liquidity and market sentiment, etc.) to build a multi-factor risk prediction model, and using machine learning technology (such as XGBoost and random forest) to train the model, the risk early warning information can be provided in a timely manner when the future market fluctuates.

This modeling design aims to construct a risk control system for the securities market through scientific methods, which not only provides a tool for effective market management, but also provides a reasonable income reference for ordinary investors.

2. Problem analysis

Systematic risk is usually affected by a combination of many factors, so it is very important to design a multi-factor prediction model for effectively predicting future systematic risk. In this problem, the following key indicators are selected as input features, and a prediction model is constructed by using machine learning method:

1. Volatility Indicator (MVI): The 20-day rolling volatility

is used to measure the volatility of the market as a benchmark measure of market risk.

2. Liquidity indicator (LSI): Liquidity is measured by turnover rate and volume pressure (such as relative changes in volume). This indicator reflects the trading activity of assets in the market and is an important indicator of market health.

3. Linkage indicator (MCI): It is used to calculate the co-activity of assets in the market, which is quantified by rolling the correlation between stocks. Systemic risk increases when there is a strong synergistic movement of market assets.

4. Volume Stress Indicator (VPI): helps capture extreme market behavior by analyzing the rate of change in volume to reflect the market's trading pressure.

Based on feature engineering, this paper constructs a risk prediction model by using XGBoost, random forest and other machine learning algorithms. The main purpose of this model is to identify potential risk signals in a timely manner in the future market fluctuations. Through the analysis of feature importance, this paper also identifies the most critical factor for systematic risk prediction, which further improves the interpretation of the model.

3. Model Assumptions and Symbols

3.1. Model Basic Assumptions

1. Historical data representation assumption: It is assumed that the market's revenue and risk characteristics can be represented by historical data. This assumption facilitates the use of historical data to build models that predict future market risks.

2. Investor behavior ignores the assumption: assuming that the short-term trading behavior of a single investor has less overall impact on the systematic risk of the market, the model does not consider the micro-behavior changes of individual investors, but extracts the systematic risk characteristics through the overall data of the market, thus simplifying the complexity of the model.

3. Normal distribution assumption: It is assumed that the distribution of market returns and other relevant data conforms to or approximately conforms to the normal distribution. This assumption facilitates the use of standard statistical methods to measure market risk (e.g. when calculating VaR and CVaR), even though there may be skewness and kurtosis in actual market returns.

4. Market efficiency assumption: It is assumed that the market is valid, i.e. all public information has been reflected in the price of the securities, and the price changes follow the changes in market information. This assumption is particularly important in risk prediction models as it affects the forecast indicator's response to future systemic risks.

5. Risk indicator independence assumption: It is assumed that each indicator in the risk measurement is relatively independent (such as volatility, liquidity, mood and other indicators), i.e. the interference between each other is small, which can reflect the market risk characteristics from different perspectives. This assumption facilitates the decomposition of the model structure and the analysis and prediction item by item.

3.2. Symbolic Description

Table1: Symbolic Description

symbol	meaning	unit
R_i	Yields on individual stocks	%
R_m	Market rate of return	%
V_t	Market liquidity	
RSI	relative strength index	
β	Systematic risk factors	
VaR	Value at risk	
CVaR	Conditional value at risk	

4. Model Establishment and Solution

4.1. Solution

In order to effectively predict the market systematic risk, this paper constructs a multi-factor risk prediction model, and uses machine learning method to model various risk factors. Systematic risk is influenced by many factors. This model selects volatility, liquidity, market linkage and other risk factors as feature inputs, and learns the relationship between these factors and market risk through machine learning model, thus identifying potential systematic risk signals.

The specific solution is as follows:

1. Feature selection and handling: Select a number of important indicators based on the characteristics of systemic risk, such as volatility, liquidity, market linkage, etc. These characteristics can reflect the risk status of the market from different dimensions.

2. Selection and comparison of machine learning model: XGBoost model is mainly used, which is suitable for multi-factor risk prediction due to its strong non-linear fitting ability and feature selection ability. At the same time, the stochastic forest model and the LightGBM model are compared to ensure the selection of the best model. XGBoost and LightGBM are tree models based on gradient lifting, while random forest improves generalization ability by integrating multiple decision trees.

3. Model training and verification: historical data are divided into training sets and test sets. The performance of the model is evaluated through cross-validation to prevent over-fitting consolidation and improve the robustness of the model to future data.

4. Feature importance analysis: after the training of the model is completed, the key factors that have great impact on the systematic risk are identified through the feature importance analysis to provide further explanatory support. Feature importance analysis helps identify key risk drivers and ensures the model is transparent and interpretable.

4.2. Model Establishment

In order to realize the prediction of systematic risks, the 'SystemicRiskPredictor' class is constructed in this paper, which mainly realizes the functions of feature engineering, model training, prediction and evaluation, etc. The structure and steps of this class are as follows:

1. Feature engineering:

The 'SystemicRiskPredictor' class first performs feature engineering on multi-factor data, including lag features, rolling mean and standard deviation, trend features, etc. The specific steps are as follows:

Hysteresis characteristics: Multiple hysteresis period values are added to each factor to capture the hysteresis effect of market changes.

Rolling statistical features: calculate the rolling mean and

standard deviation of each feature to help the model identify short-term and long-term risk fluctuations.

Cross-features: Build interactive items between features (such as the product of volume and volatility), enrich feature space and improve the expressiveness of the model.

2. Model training and parameter adjustment:

After the feature engineering is completed, 'SystemicRiskPredictor' is modeled using the XGBoost model. The specific process includes:

Data set partitioning: the data is divided into training set and test set in chronological order, the training set is used for model fitting, and the test set is used for evaluation.

Super-parameter optimization: optimize super-parameters (such as tree depth, learning rate, number of trees, etc.) through grid search or random search, and select the best combination of parameters to obtain the best prediction effect.

Model evaluation: the performance of the model is evaluated using indicators such as mean square error (MSE) to ensure its accuracy and stability. (R^2)

3. Feature importance analysis:

After the training, the importance of the input features is sorted by using the feature importance function built in the XGBoost model. Through the analysis of feature importance, the factors that contribute greatly to the systematic risk prediction are identified to provide support for the model interpretation. The specific analysis is as follows:

Volatility: Volatility increases significantly during periods of high risk and is the most important risk predictor in the model.

Liquidity: Low liquidity indicates that the market may be exposed to the risk of increased uncertainty and is a key predictor.

Market linkage: High linkage indicates that the linkage between markets is strengthened, which increases the possibility of systemic risk.

4.3. Model Solution

After the model training and feature importance analysis were completed, the 'SystemicRiskPredictor' made a forecast of the future systematic risks. The following are the main results and analysis of the model solution:

1. Systematic risk forecast results:

The forecast output of the XGBoost model shows the trend of systematic risk in the future market. In order to compare the effect, this paper also uses random forest and LightGBM model to test and draw a comparison chart of forecast results. The results show that the prediction trends of the three models are approximately the same, but XGBoost is more accurate in fitting some peaks.

2. Model comparison:

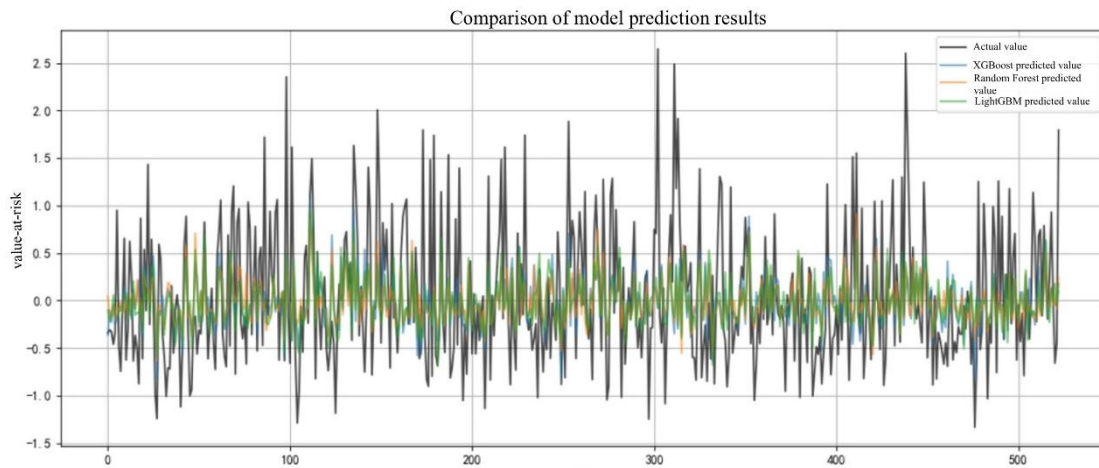


Figure 1. Comparison diagram of model prediction results

The above figure shows the comparison between the predicted results and the actual values of the three models: XGBoost, Random Forest and LightGBM. It can be seen that XGBoost and LightGBM perform well in fitting high-frequency fluctuations, while random forests perform stably in a relatively stable interval. The specific comparison is as follows:

XGBoost: More accurate in capturing risk peaks, suitable for market environments with rapidly changing risks.

Random forest: performs well under stable market conditions and is suitable for low fluctuation environment.

LightGBM: performance is close to XGBoost, but performance is slightly worse with less feature interaction.

3. Feature importance analysis:

By ranking the importance of features, the main factors such as volatility, liquidity and market linkage are identified. The high weight of these factors indicates their significant contribution to the systematic risk prediction.

4. Model performance assessment:

The performance of the model on the test set is verified by

mean square error (MSE) and other evaluation indicators. Cross-validation results show that the XGBoost model can more accurately identify the trend of systematic risk and provide effective risk prediction. (R^2)

To sum up, the 'SystemicRiskPredictor' class has successfully achieved the prediction of systematic risks. The forecast results of the model show the changing trend of the market in the future high-risk period and provide scientific basis for risk management.

5. Model evaluation and promotion

5.1. Benefits of the Model

1. The model combines the reality and reasonably describes the market risk by using multiple risk indicators, which is suitable for the risk management of the securities market.

2. Use machine learning model to improve prediction ability and effectively capture systemic risks.

3. Robust risk control is achieved through dynamic position management.

5.2. model deficiencies

1. Systematic risk changes that may arise from market emergencies are not considered.

2. The linear assumptions of some indicators in the model may not be fully consistent with the actual market conditions.

5.3. Extension of Model

In other securities markets, model parameters can be replaced to achieve risk control; At the same time, the application effect of the model is improved by further optimizing the selection of risk factors.

References

- [1] Zhang Yizhi. Exploring the Problem of Predictive Information Disclosure in Securities Market [D]. East China University of Political Science and Law, 2023.doi: 10.27150/d.cnki.ghdzc.2023.000850.
- [2] Luo Xin. Research on price and volatility forecast of China stock market based on in-depth learning [D]. Zhongnan University of Finance and Economics and Law, 2020.doi: 10.27660/d.cnki.gzcuz.2020.002987.
- [3] Ma Jiahao. China's Stock Market Forecast Model Based on Artificial Intelligence LSTM and Baidu Investor Attention [D]. Southwest University of Finance and Economics, 2019.doi: 10.27412/d.cnki.gxcu.2019.001641.
- [4] Ouyang Tianhao, Lu Xiaoyong. A New Method to Measure the Stability of Securities Markets under the Background of Financial Security-Research on Market Forecast and Arbitrage Value Measurement Based on Big Data Support Vector Machine [J]. Financial Theory and Practice, 2019, 40 (01): 77-83. DOI: 10.16339/j. Cnki. HDXBCJB.2019.01.011.
- [5] From Shanshan. Research on Trend Forecast of Securities Market and Measurement of Market Relative Efficiency Based on Bayesian Method [D]. China University of Metrology, 2018.