Real-time Impact of Securities Price Volatility Based on Big Data Mining

Fanjin Feng

Nanchang Branch of Jiangxi Radio and TV University

939696850@qq.com

Keywords: Big Data Mining, Stock Price Fluctuation, Internet Information, Stock Research

Abstract: With the development of big data, a large number of fast information on the Internet provides rich data support for predicting the changes of the securities market, so how to use the big data analysis technology to mine the real-time changes of the securities market is a hot issue of current research. This paper aims to study the real-time impact of securities price volatility based on big data mining. This paper uses AISI model to study the index change value of market influence. If it is negative, it means that it has a negative impact on the market; if it is positive, it means that it has a positive impact on the market. In this paper, according to the recent time of Internet information on the stock market change prediction experiment, data collection, analysis and prediction are carried out, and the actual verification is carried out according to the results and the subsequent situation. The experimental data show that the extraction method, change relationship and influence model of information content and each characteristic factor in the Internet are constructed, and the Internet information indicators for the large market, industry and individual stocks are discussed to reflect the support degree of the data. The experimental results show that the AISI prediction indexes of 8 trading days in 10 days are: - 2, 1, - 5, - 1, 3, - 22, 1 and - 3, which can realize the method of forecasting the stock market based on the comprehensive characteristic factor measurement. Therefore, big data mining method has good feasibility and will bring obvious academic and commercial value.

1. Introduction

The impact of information on the securities market is the core issue of financial research and an important basis for the majority of investors to make investment decisions [1]. Decision errors caused by information asymmetry have been the main reason for investment failure for a long time, so how to effectively obtain and analyze information and make market forecasts has always been a hot topic in the industry and academia. With the rapid development of Internet technology, the rich data of Internet big data has transformed the securities market from information shortage to information surplus. The rapid growth of securities information on the Internet has brought valuable opportunities for securities analysis and forecasting, and it is also a severe challenge [2 -3]. On the one hand, various emerging online social media have become the main platforms for information transmission and communication, which contain massive amounts of information and provide rich data support for investment decisions; on the other hand, the challenges of big data brought by the development of the Internet And problems also plague the effective analysis and use of information.

Yoshino et al. introduced the commendation and devaluation of Internet financial information in the SVM prediction model, and revealed the correlation between financial market volatility and Internet news through experiments, and proposed an effective stock market prediction method [4]. However, these methods only studied the relationship between the proportion of propensity and the volatility of the stock price of financial market companies, and did not conduct in-depth research on the content of the information itself. Zhang et al. proposed a rule acquisition algorithm based on the largest grains. According to the hierarchical division space formed by the conditional attribute domain, the algorithm gradually extracts the rules corresponding to the largest grains from top to bottom, which improves the generalization of rough sets. Ability to achieve rule-based knowledge extraction [5].

As people gradually rely on Internet information to judge affairs, the Internet is gradually being accepted and used by investors [6-7]. The Internet has not only changed the way investors participate in investment activities, but has also become the main channel for the dissemination of investment information in the stock market. More and more investors use the Internet to search for financial information and exchange investment experience with others [8-9]. The timely and effective dissemination of information from various channels on the Internet has caused major changes in the information dissemination method of the capital market, effectively solving the problem of information asymmetry in the securities market, and the new market information structure also has an impact on the behavior of investors It has a huge impact, which in turn affects the pricing of stock assets and the allocation of financial resources [10].

2. Research on Real-time Impact of Securities Price Volatility Based on Big Data Mining

2.1 Data Mining Technology

Data mining is a new commercial information processing technology. Its main feature is to convert, extract, analyze and other model processing of large amounts of data in commercial databases, and extract key data that assist business decision-making. Therefore, data mining Technology is widely used in various business fields. With the rapid development of the social economy, the development of China's securities market has resulted in the accumulation of securities transaction data and the formation of a dense and huge data set. Using data mining technology, these complex data sets are used to dig out useful data for investors. information. Through the analysis of this data and the use of mathematical tools to study the securities market and the development trend of securities, and to further predict the securities market, it has played an important role in promoting the research of my country's securities field and has become the focus of attention in the securities industry.

2.2 Artificial Neural Network

Since the securities market is a nonlinear dynamic system, there are many factors that affect the securities market, and the securities index shows complexity and uncertainty. Early securities analysis methods and forecasting tools, such as traditional securities index technical analysis methods, are obviously not accurate enough due to some factors and limitations of calculation accuracy. Using the advantages of neural networks in nonlinear systems to make predictions, through the learning and training process of neural networks, the speed and accuracy of prediction can be improved.

Neural networks have some advantages that other methods can't match, including:

(1) The artificial neural network has a self-learning function, which is particularly important for prediction. The artificial neural network that can predict the future can provide economic forecasts, market forecasts, crisis forecasts and benefit forecasts, with a wide range of applications;

(2) The neural network can approximate any nonlinear continuous function with arbitrary accuracy. For non-independence, the totality of factors can still handle the complex nonlinear relationship;

(3) The feedback network of the artificial neural network has the function of associative storage;

(4) The artificial neural network has the ability to find optimal solutions. Designing a feedback artificial neural network for a specific problem can find the optimal solution at a high speed and improve the high-speed computing capability of the computer.

2.3 The Influence Model of Internet Information and Securities Market Changes

Taking various types of information on the Internet as the data source, a new "Internet Information Index" (ISI) was constructed to supplement and test the volume index. The ISI indicator has the following characteristics:

(1) The correlation between ISI indicators and the securities market, based on which it is judged whether the information content obtained on the Internet is significantly related to changes in the

securities market;

(2) The degree of influence of the coexistence model of ISI index and trading volume index on the changes in the securities market, and based on this, study the trend analysis of the influence of the index data on the changes in the securities market.

The main economic purpose of constructing this indicator is to find an indicator that can enhance the ability to explain changes in the securities market. The economic meaning of this indicator is to add up "securities-related" Internet information; and to consider whether the indicator can directly reflect open source the quality of the information and the ability to deal with "noise".

According to different analysis angles, the ISI indicator is studied from three levels, namely, the comprehensive Internet information indicator AISI, industry Internet information indicator IISI, and individual stock Internet information indicator CISI.

Define the factors that can be obtained in the Internet, including economic cycles, fiscal policies, interest rate changes, exchange rate changes, price changes, inflation, political policies, industry changes, operating conditions, upstream and downstream influences, and construct a vector that can be continuously observed on a daily basis, Among them, Vi represents the i-th influencing factor:

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_9, v_{10}\} \qquad (1)$$

The AISI model of comprehensive Internet information indicators for the market is as follows:

$$AISI = \sum_{i=1}^{10} \sum_{j=1}^{k_i} f_i \times \omega(N_j, \Delta t) \qquad (2)$$

For the current time T, if in a recent continuous observation period [t, T], there are k_i information about the influence factors of vi on the Internet, where the j-th message information $N_j=(j=1,....,k_i)$ The comprehensive influence on the market is f_j, which is defined as the index change value that affects the market. If it is negative, it means that it has a negative impact on the market. If it is a positive number, it means it has a positive impact on the market. This value is obtained through historical machine learning and empirical judgment for various types of information events. The occurrence time of the message j is t_j, and the weight function of the influence change with time is $w(N_j, \Delta t)$, where $\Delta t = T-t_j$, which represents the time interval from the occurrence time, and describes the change over time The impact of the news on the current market situation.

2.4 Investor Attention Measurement Based on Network Big Data

Compared with traditional transaction data or financial statement data, network big data has its unique advantages: (1) Scale. The huge scale of network data provides ideal measurement samples for related research, improves the fault tolerance rate of data processing, and ensures the accuracy of measurement results. (2) Diversity. Network data includes information left by investors voluntarily, such as basic personal information, life records, and online messages, as well as data left unintentionally by investors, such as surfing habits, social relationships, and emotions. These data are the key advantages of online big data to portray investor psychology and behavior. (3) High speed. Big data on the Internet records the behavior of investors all the time and can directly reveal the psychological state of investors. But at the same time, there is a lot of noise in network big data, and researchers need to have certain big data processing capabilities to obtain "clean" data.

3. An Experiment on the Real-Time Impact of Stock Price Fluctuations Based on Big Data Mining

3.1 Setting of Financial Stability

This article believes that financial stability does not only refer to the stability of bank credit or exchange rate fluctuations, but a multi-dimensional indicator system that can comprehensively measure the stability of financial-related variables such as banks, securities, foreign exchange, and macroeconomic conditions. The financial stability indicator system used in this paper is shown in Table 1.

Table 1. My country's Financial Stability Index System

First level indicator	Secondary indicators
World Financial Sentiment Index	Ted Spread
	Non-bank credit as a percentage of total credit
Financial Development Index	Domestic credit
	Securitization rate
	Deposit and loan interest rate difference
	Broad money supply

3.2 Experimental Setup

In this article, you can use Matlab software to minimize error calculations and obtain the desired output value, making it as close to the actual value as possible. Use Newrb to create a radial basis function neural network in Matlab, set the mean square error and radial basis function distribution of the neural network output and target output, and the default function of the neural network performance to calculate the mean square error. When the radial basis function neural network approaches this function, Newrb automatically increases the hidden layer neurons of the radial basis function and iteratively adjusts the network until the average Hall error is satisfied, thereby reducing the value of the performance function. The preset error accuracy is reached. The larger the extended range of the radial basis function, the smoother the fit of the function.

4. Discussion on Real-time Impact of Securities Price Fluctuation Based on Big Data Mining

(1) Through the mining and analysis of information events on the Internet, this paper constructs the extraction method, change relationship and influence model of the information content and various characteristic factors in the Internet, and proposes Internet information indicators for the market, industry, and individual stocks to reflect the data The degree of support. And according to the recent time, data collection, analysis and forecasting were carried out, and the actual verification was carried out according to the results and the subsequent situation. By collecting all relevant information events that occurred during the half-month period from April 1, 2020 to April 10, 2020, there are 13,823 related information events after deduplication, with an average of about 922 per day and daily data The distribution is shown in Table 2.

Table 2. Distribution of the number of related information events that occurred

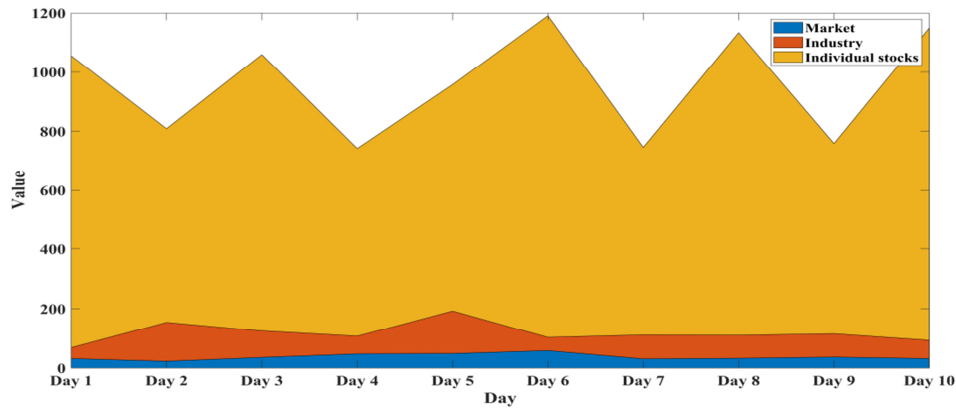| Date	Market	Industry	Individual stocks
2020.4.1	32	37	987
2020.4.2	23	132	654
2020.4.3	36	90	934
2020.4.4	48	60	634
2020.4.5	49	144	765
2020.4.6	59	45	1087
2020.4.7	31	81	634
2020.4.8	33	78	1023
2020.4.9	37	79	643
2020.4.10	32	62	1056

Figure 1. Distribution of the number of related information events that occurred

(2) In order to reduce the calculation difficulty caused by the long-term influence of news, this article assumes that the time period of news is 4 days, that is, 4 days after the news occurs, it will no longer affect the market, and analyze the data based on the comprehensive index of the market. According to the data situation of each information, the AISI index of the 8 trading days in the 10 days is calculated according to the model. Table 3 lists the value change index and error of each trading day predicted by AISI (Indicator of Comprehensive Internet Information of the Market). It can be seen from the results that the trend of the predicted change result is basically in line with the actual situation, and the crossover situation belongs to the abnormal value of the prediction failure. The principles and methods of IISI for the industry and CISI for individual stocks are basically the same.

Table 3. The value change index and error of each trading day predicted by AISI

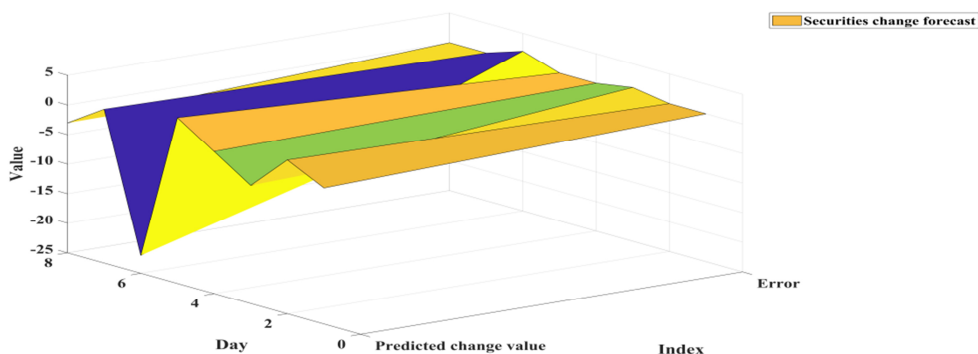| Date	Predicted change value	Error
2020.4.3	-2	0
2020.4.4	1	0
2020.4.5	-5	1
2020.4.6	-1	0
2020.4.7	3	0
2020.4.8	-22	2
2020.4.9	1	0
2020.4.10	-3	0

Figure 2. The value change index and error of each trading day predicted by AISI

5. Conclusion

The securities market is a complex and huge open system. Subjectivity and objectivity are

contradictory and dialectical sides. This article analyzes the information events on the Internet and models 10 factors that affect the changes in the securities market. This method is based on the degree of influence of the information on the Internet to quantitatively analyze the fundamentals of the securities market. It has objectivity and independence, which reflects the nature of price changes in the securities market and is fundamentally different from traditional qualitative analysis. The experimental results show that the AISI forecast indexes for the 8 trading days within 10 days are: -2, 1, -5, -1, 3, -22, 1, -3, which can realize the comprehensive feature factor measurement based on data. Methods of forecasting the securities market.

References

[1] Ding Z, He L, Feng C, et al. The impact of coal price fluctuations on China's economic output[J]. Applied Economics, 2016, 48(24):2225-2237.

[2] Song M, Wang J. Coal price fluctuations in China: Economic effects and policy implications[J]. Journal of Renewable & Sustainable Energy, 2016, 8(6):422-431.

[3] Yoshino N, Taghizadeh-Hesary F. Economic Impacts of Oil Price Fluctuations in Developed and Developing Economies[J]. Ieej Energy Journal, 2016, 9(3):58-75.

[4] Ding Z, Liu Z, Zhang Y, et al. The contagion effect of international crude oil price fluctuations on Chinese stock market investor sentiment[J]. Applied Energy, 2017, 187(FEB.1):27-36.

[5] Boileau M, Normandin M. The price of imported capital and consumption fluctuations in emerging economies[J]. Journal of International Economics, 2017, 108(sep.):67-81.

[6] Ji M, Li H. Exploring Price Fluctuations in a Double Auction Market[J]. Computational Economics, 2016, 48(2):189-209.

[7] Chuku, Chuku, Simpasa, Anthony, Oduor, Jacob. Macroeconomic Consequences of Commodity Price Fluctuations in African Economies[J]. African Development Review, 2018, 30(4):329-345.

[8] Shao Q, Schaffartzik A, Mayer A, et al. The high 'price' of dematerialization: A dynamic panel data analysis of material use and economic recession[J]. Journal of Cleaner Production, 2017, 167(nov.20):120-132.

[9] Zhou D T, Yu H Y, Li Z G. Effects of fluctuations in international oil prices on China's price level based on VAR model[J]. Journal of Discrete Mathematical ences and Cryptography, 2017, 20(1):125-135.

[10] Pomeranets A, Weaver D G. Securities Transaction Taxes and Market Quality[J]. Journal of Financial and Quantitative Analysis, 2018, 53(1):455-484.