

Prediction of Stock Trading Prices Based on Big Data

Yunzhe Liu^{1,a,*}

¹*School of Economics and Management, Qinghai Minzu University, Xining, Qinghai, 810007, China*

^a*liuyunzhe0815@163.com*

^{*}*corresponding author*

Keywords: Big Data, Stock Trading, Trading Prices, Forecasting Research

Abstract: With the continuous development of the economy, people have gradually started to enjoy playing with stocks. Although stock trading comes with certain risks, it often comes with returns. Therefore, how to effectively predict stock prices, avoid risks, and increase returns has become the focus of current research. This article studied the prediction of stock trading prices based on big data, aiming to improve the accuracy of stock trading price model prediction through big data technology. This article tested the accuracy of using big data to predict stock trading prices through experiments, with a maximum of 88% and a minimum of 80%. The accuracy of traditional stock trading price prediction models was highest at 72% and lowest at 60%. From this experimental result, it can be seen that big data can indeed improve the effectiveness of stock trading price prediction models, proving the high degree of compatibility between big data and stock trading price prediction.

1. Introduction

With the development of the times, people's economic level is also constantly improving, and their living standards have made a huge leap. While the market economy is developing rapidly, the financial market is also expanding at an unimaginable pace. After people became wealthy with money, they also began to gradually develop a sense of investment. Subsequently, stocks became the focus of attention in people's eyes. If people want to have a clearer understanding of the trends in stock prices, further research is needed on predicting stock trading prices.

Many scholars have conducted certain research on stock investment. Scholar Ramelli S believed that the stock buying to selling ratio of individual investors has seasonal characteristics. In late December, the individual buying to selling ratio was lower than normal, and in early January, the individual buying to selling ratio was higher than normal [1]. Li R believed that the strategy of buying stocks that have performed well in the past and selling stocks that have performed poorly in the past can generate significant positive returns over a holding period of 3 to 12 months. The profitability of these strategies is not due to their systemic risks, nor is it due to the delayed response of stock prices to common factors [2]. Li W believed that when a company announces an equity increase, its stock price usually drops. A standard explanation is that equity issuance conveys negative information to investors. If there are many transactions in the capital market, the information gap between the company and investors may lead to some companies repurchasing

stocks in an equilibrium state [3]. Although there is a lot of research on stock trading, there is little research on stock trading price models.

Many scholars have begun to use relevant technical indicators to predict the stock market. This article focused on the study of big data stock trading prediction, which used big data technology to make more accurate predictions of stock trading prices, in order to facilitate people's purchase of stocks. This article tested the accuracy of stock price prediction using big data technology through experiments and found that the accuracy has been significantly improved compared to traditional methods. A satisfaction survey was also conducted on 100 users who used big data to predict stock trading prices, and it was found that the satisfaction data was high. This indicates that big data has a good effect in predicting stock trading prices.

2. Prediction of Stock Trading Prices Using Big Data

2.1. Overview of Stocks

Stocks are a type of securities that allow people to hold shares, that is, certificates that allow people to receive dividends. The characteristics of securities include profitability, risk, liquidity, persistence, and participation [4-5].

The fundamental characteristic of stocks is their source of income, which comes from two sources. One is obtained from a joint-stock company, which means that the person holding the shares has an economic interest in the issuing company. Its manifestation is the dividends and bonuses distributed by the company. The second is derived from the circulation of stocks, which means that holders of stocks can earn profits between buying and selling.

Stock risk refers to the difference between the actual profit obtained and the expected profit when purchasing stocks. Securities investment is a high-risk and high return behavior, where high risk not only means losses, but also returns [6-7].

2.2. Basic Theory of Stock Prices

Stock price refers to the trading price of a stock in the stock market. In theory, the value of a stock depends on its own value. The formation of stock prices depends on the value of the market, and the price of the stock market is influenced by many factors, the most intuitive of which is the relationship between supply and demand. From the perspective of supply and demand, if the buyer's intention to buy exceeds the seller's intention to sell, the stock price would rise [8-9]. Correspondingly, prices would decrease. Stock prices are influenced by various factors. Therefore, the volatility of stock prices would be significant [10-11].

In terms of phenomenon, the price of stocks, like the price of other commodities, is influenced by the relationship between supply and demand. When supply exceeds demand, the stock price drops, while when supply exceeds demand, the stock price rises. However, unlike other commodities, stocks are just a type of voucher and have value because they can bring profits to their owners. The so-called stock certificate is actually a type of certificate for buying and selling stocks, used to obtain profits [12-13].

2.3. Factors Affecting Stock Price Fluctuations

The stock market price is influenced by various factors and analyzed from three perspectives. One is the company's performance. Generally speaking, the stock price is proportional to the company's performance, and the better the performance, the higher the stock price [14-15]. On the contrary, if a company's operations are not good, its stock price usually also falls. The second is

internal factors within the industry. After long-term analysis of the fluctuations in stock prices in different industries, it can be seen that within a certain period of time, an industry often exhibits the same fluctuations, that is, the same reasons are affecting the stock prices of the industry, which can be referred to as internal reasons within the industry. The third is macroeconomic and policy, which have a significant and widespread impact on stock prices [16-17]. In addition, the fluctuation of stock prices is influenced by other factors. For example, among political and other irresistible reasons, the most influential one is war. Once an unpredictable war occurs, it has a significant impact on the stock market, leading to a decline. The formulation of important economic policies also has a certain impact on investors' investments. At the same time, the psychology of investors also affects the changes in stock prices to a certain extent, and risk averse individuals usually hold a more optimistic attitude. Therefore, rushing to buy stocks drives up stock prices. Those who want to avoid risks hold a pessimistic attitude and a wait-and-see attitude towards the stock market, which leads to a decline in stock prices. This makes the stock prices of listed companies highly uncertain and carries high risks and returns [18-19].

2.4. Stock Price Movement

Through observation of the stock market for a period of time, it can be seen that stock prices have daily small fluctuations, short-term fluctuations, and trend changes, just like the ups and downs of tides. These three different forms of motion coexist and interact with each other [20]. These three movements are referred to as daily changes, secondary changes, and main trend changes in this article.

The daily changes are shown in Figure 1, which shows that the daily changes are random and influenced by good news, bad news, and other technical factors, ranging from a few hours to a few days. Because the time is too short, investors do not care.

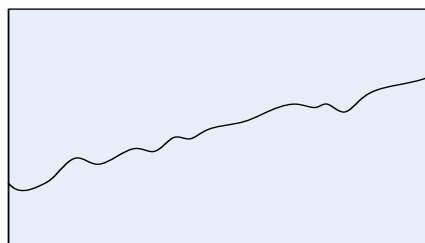


Figure 1: The daily movements of the stock market

Secondary movement, as shown in Figure 2, refers to the phenomenon of a sudden drop in stock prices during an upward trend or a mid-term rebound during a downward trend. Secondary exercise usually occurs between a week and a few months.

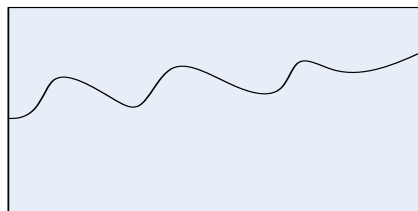


Figure 2: Secondary movements in the stock market

The main trend movement is shown in Figure 3, which refers to the long-term upward trend in stock prices, also known as the bullish market. And there is a long-term downward trend, which is to look bad at the market. Once a major trend appears, it usually lasts for one to several years.

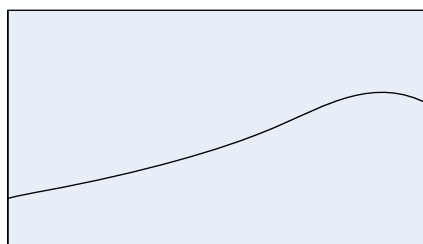


Figure 3: Major movements in the stock market

2.5. Overview of Big Data

Big data refers to a collection of data that cannot be captured, managed, and processed using ordinary software tools for a period of time. From the perspective of the computer industry, big data is a huge collection of data. Big data is no longer in the form of text, but may also be video, image, voice, etc. These data may have no obvious rules, most of them are unstructured data, and can not be expressed in two-dimensional mode; A large amount of irrelevant information, in order to mine valuable data, one must search for gold in the sand, which is a manifestation of low data value density.

Big Data Technology: The core technologies of big data include: big data collection, preprocessing, storage, analysis and model prediction, and presentation of big data results.

Big data collection: In the collection of web data, the most common technique is web crawler technology (for example, input method development, which can be used to obtain a large amount of vocabulary data due to the need to continuously update the vocabulary and adjust the frequency of vocabulary usage); Data on the platform is collected, such as Taobao.

Big data preprocessing: Data preprocessing refers to the necessary cleaning, integration, transformation, discretization, and reduction of raw data before data mining, in order to meet the minimum specifications and standards required for knowledge acquisition research in mining algorithms. **Large capacity storage:** database, file storage, etc. **Big Data Analysis and Mining:** the use of technologies such as neural networks, predictive models, pattern recognition, time series analysis, machine learning, and natural language analysis.

Presentation of big data calculation results: Present the analyzed and calculated results in the most direct and easily understandable way, such as charts. **The application of big data:** Big data has important applications in various fields of society, as well as in the financial field. In the financial industry, compared to the banking and insurance industry, the development speed of the securities market is much slower. Currently, the application of the securities market mainly focuses on three aspects: first, predicting stock prices; Another is customer relationship management; The other is the “Investment Prosperity Index” Next, some feasibility analysis and research on the stock price prediction direction of the securities industry in the big data environment are conducted.

2.6. Steps for Stock Forecast Analysis

By comparing the construction models of Gaussian kernel functions, polynomial kernel functions, etc., the optimal kernel function can be obtained. On this basis, genetic algorithm, grid optimization, and particle swarm optimization methods are used to obtain the optimal core parameters and predict them. The specific content includes: 1) By collecting historical data of predicted bank stocks, they can be divided into two groups: one is the training set, and the other is the testing set. 2) In order to accelerate the convergence of the algorithm, standardized preprocessing was performed on the original stock data. 3) Using genetic algorithms, particle

swarm optimization, grid parameter optimization, and other techniques to find the best parameters. The optimal parameters obtained from various optimization algorithms are utilized to train stock market data and determine the optimal parameters. Through experimental comparison, the optimal support vector machine prediction model determined by each kernel function is obtained. 4) The established optimal support vector machine prediction model is applied to the stock price data of Bank of China, and corresponding prediction results are obtained.

2.7. Big Data in Stock Price Forecasting

Big data is used in stock price prediction and analysis through regression analysis, which is a statistical tool used to study the correlation between variables and has high reliability and wide applicability. Regression analysis is usually used to study the specific dependencies of a variable on other variables, and can be based on observational data to determine the causal relationship between the independent and dependent variables, thus constructing a regression model. Principal component analysis is a multivariate statistical method that converts multiple related variables of the tested object into a few unrelated variables, and these unrelated composite variables also contain most of the information of the original variable. By selecting the first few important principal components according to certain standards, the original problem is greatly simplified, and the size of the impact of these comprehensive indicators, namely the principal components, on the studied object can be determined.

Big data also uses grey correlation analysis in stock price prediction and analysis. Grey correlation degree refers to the similarity of development trends between factors, mainly based on the approximation degree of sequence curve geometry. Grey correlation degree is used to measure the contribution of factors to the behavior of system agents. Generally speaking, the closer the curve formed by the comparison sequence and the curve formed by the reference sequence are geometrically, the higher their correlation degree, that is, the higher their grey correlation degree.

2.8. Prediction Model for Stock Prices

The stock prices at different times, especially those at continuous times, are closely related to each other, which means there is a correlation between prices. Therefore, an autoregressive model can be used to describe it. In addition, since stock prices are influenced by various external factors, it is best to construct an autoregressive model with control variables that represent the impact of external factors to simulate changes in stock prices. Without losing universality, the representation method of the stock price model is shown in formula 1:

$$y = at + b \quad (1)$$

In formula 1, y is the stock price at time t , and a and b are the values of the control variables. If the least squares parameter is used to estimate stock prices, and the forgetting factor is set to i , $0 \leq i \leq 1$, then formula 2 can be obtained by recursion using the least squares method:

$$y = \frac{i}{t+b} \quad (2)$$

The value of the forgetting factor determines the speed of stock price evaluation. To maximize the evaluation speed, the value of i is needed to reach its maximum. The calculation method of i is shown in formula 3:

$$i = \sqrt[2]{ab} \quad (3)$$

3. Simulation Experiment of Big Data Application in Stock Price Prediction Model

This article uses stock data as training data to estimate stock prices and tests the accuracy of the constructed stock model using a function. In stock forecasting, the main target is the closing price. This article mainly focuses on short-term stock prediction, predicting stock trends over the next 10 days, and then comparing the accuracy of stock price prediction models under big data and traditional stock prediction models through research to verify the effectiveness of big data in stock price prediction models. The accuracy comparison is shown in Figure 4.

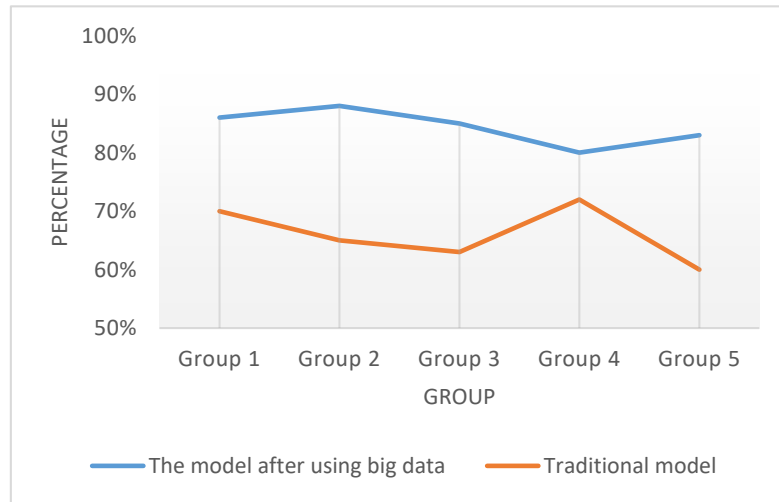


Figure 4: Comparison of prediction accuracy using big data model and traditional model

From the experimental results in Figure 4, it can be seen that the accuracy of the stock price prediction model using big data is the highest at 88% and the lowest at 80%, while the accuracy of traditional stock price prediction models is the highest at 72% and the lowest at 60%. Through this experimental data, it can be seen that the accuracy of stock price prediction has improved significantly after using big data, which proves that the use of big data in stock price prediction models is very successful.

This article also tested the satisfaction of 50 users who used the big data stock price prediction model with the big data stock price prediction model. The satisfaction of these users is shown in Figure 5.

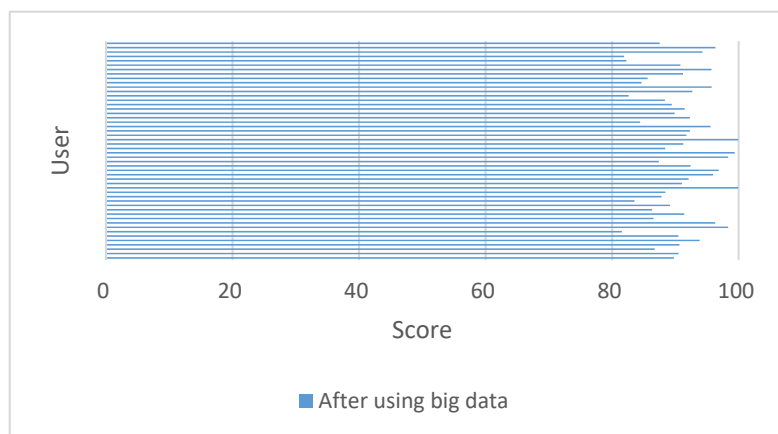


Figure 5: User satisfaction with big data stock price prediction model

From the survey in Figure 5, it can be seen that the satisfaction of 50 users is above 80 points, with the highest reaching a full score, indicating that users are very satisfied with the big data stock price prediction model. This also indirectly indicates that the use of big data in stock price prediction models is very effective.

4. Conclusions

Nowadays, stocks are a relatively popular thing, loved by both men and women of all ages. Although playing with stocks often comes with some risks, it also brings significant returns to investors. Since stocks are so common, it is necessary to develop stock price prediction models to help people better play with stocks. People can only buy the stocks they want with a clearer understanding of their stock prices, so that even if they are losing money, they are willing to do so. Therefore, this article studied the prediction of stock trading prices based on big data, aiming to improve the accuracy of stock trading price model prediction through the information technology of big data. More accurate data can enable people to buy stocks that best meet their needs. This article tested the prediction accuracy of using big data to predict stock trading prices and the accuracy of traditional stock trading price prediction through experiments. From the experimental data, it can be found that the accuracy of the stock trading price prediction model has been improved after using big data, indicating that big data has achieved good results in the stock trading price prediction model. Due to space limitations, the experiments conducted in this article are still insufficient and would be improved in the future. Finally, it is wished that the stock trading price model could become better and better.

References

- [1] Ramelli S, Wagner A F. Feverish stock price reactions to COVID-19. *The Review of Corporate Finance Studies*, 2020, 9(3): 622-655.
- [2] Li R, Li N, Li J. Short selling, margin buying and stock return in China market. *Accounting & Finance*, 2018, 58(2): 477-501.
- [3] Li W, Chien F, Kamran H W. The nexus between COVID-19 fear and stock market volatility. *Economic research-Ekonomiska istrazivanja*, 2022, 35(1): 1765-1785.
- [4] Moen E R, Nenov P T, Sniekers F. Buying first or selling first in housing markets. *Journal of the European Economic Association*, 2021, 19(1): 38-81.
- [5] Engelberg J E, Reed A V, Ringgenberg M C. Short-selling risk. *The Journal of Finance*, 2018, 73(2): 755-786.
- [6] Wibowo A I L, Putra A D, Dewi M S. Differences in intrinsic value with stock market prices using the price earning ratio (per) approach as an investment decision making indicator (case study of manufacturing companies in Indonesia period 2016-2017). *Aptisi Transactions on Technopreneurship (ATT)*, 2019, 1(1): 82-92.
- [7] Areiqat A Y, Abu-Rumman A, Al-Alani Y S. Impact of behavioral finance on stock investment decisions applied study on a sample of investors at Amman stock exchange. *Academy of Accounting and Financial Studies Journal*, 2019, 23(2): 1-17.
- [8] Ananthi M, Vijayakumar K. Retracted article: stock market analysis using candlestick regression and market trend prediction (CKRM). *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12(5): 4819-4826.
- [9] Hendershott T, Kozhan R, Raman V. Short selling and price discovery in corporate bonds. *Journal of Financial and Quantitative Analysis*, 2020, 55(1): 77-115.
- [10] Barber B M, Huang X, Odean T. Attention-Induced Trading and Returns: Evidence from Robinhood Users. *The Journal of Finance*, 2022, 77(6): 3141-3190.
- [11] Meier C, De Mello L. Investor overconfidence in experimental asset markets across market states. *Journal of Behavioral Finance*, 2020, 21(4): 369-384.
- [12] Atmaz A, Basak S. Option prices and costly short-selling. *Journal of Financial Economics*, 2019, 134(1): 1-28.
- [13] Adiyani D, Mawardi W. The influence of overconfidence, herding behavior, and risk tolerance on stock investment decisions: The empirical study of millennial investors in Semarang City. *Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship*, 2020, 10(1): 89-101.
- [14] Matsubara T, Akita R, Uehara K. Stock price prediction by deep neural generative model of news articles. *IEICE TRANSACTIONS on Information and Systems*, 2018, 101(4): 901-908.

- [15] Purnanandam A, Seyhun H N. Do short sellers trade on private information or false information?. *Journal of Financial and Quantitative Analysis*, 2018, 53(3): 997-1023.
- [16] Kim C, Wang K, Zhang L. Readability of 10-K reports and stock price crash risk. *Contemporary accounting research*, 2019, 36(2): 1184-1216.
- [17] Leippold M, Wang Q, Zhou W. Machine learning in the Chinese stock market. *Journal of Financial Economics*, 2022, 145(2): 64-82.
- [18] Hwang C Y, Zhang S, Zhu Y. Float, Speculation and Stock Price: Evidence from The Split Share Structure Reform In China. *Singapore Economic Review*, 2018, 63(3):701-729.
- [19] Palladino L. Do corporate insiders use stock buybacks for personal gain? *International Review of Applied Economics*, 2020, 34(2): 152-174.
- [20] Duxbury D, Gärling T, Gamble A. How emotions influence behavior in financial markets: a conceptual analysis and emotion-based account of buy-sell preferences. *The European Journal of Finance*, 2020, 26(14): 1417-1438.