Dean&Francis
ISSN 2959-6157

Stock Market Price Analysis and Prediction of Financial Industry Based on the Time Series Method

Ruining Sun

Department of Mathematics and Physics, Taizhou University, Taizhou, Jiangsu Province, 225300, China

huangtingli@ldy.edu.rs

Abstract:
With the increasing complexity of financial markets and the rapid growth of data volumes, accurately predicting stock prices has become crucial for ensuring market stability and guiding investment decisions. Although traditional time series analysis models have been widely applied, they still face research gaps in prediction accuracy due to the high nonlinearity and dynamics of the market. This paper introduces the application of time series analysis methods in stock market price prediction within the financial industry. This paper aims to enhance prediction accuracy and timeliness. The study provides a detailed overview of the advantages of both traditional models and modern deep learning techniques. The latter offers a new perspective for prediction through its capabilities in nonlinear modeling and capturing long-term dependencies. Representative stocks in the financial industry are selected for empirical analysis, demonstrating the performance of models that combine time series analysis with deep learning techniques in terms of prediction accuracy and stability. Meanwhile, the paper points out current research deficiencies, such as model optimization, improvement in generalization ability, and interdisciplinary integration, as future directions. With the backdrop of big data, it is hoped that through the rational integration of time series analysis and deep learning techniques, stronger support can be provided for financial market stability and scientific decision-making.

Keywords: Time Series Analysis; Stock Market Prices; Prediction; Autoregressive Model; Long Short-Term Memory Network (LSTM)

I. Introduction

The financial market, as the core of the economic system, has an immeasurable impact on national economic security and the healthy development of the global economy in terms of its stability and efficiency. As an important component of the financial market, stock market price fluctuations not only reflect changes in market supply and demand relationships but also contain a multitude of information such as macroeconomics, industry trends, corporate fundamentals, and market sentiment. Therefore, accurate prediction of stock market prices is of great significance for investors to formulate investment strategies, financial institutions to conduct risk management, and policymakers to implement macroeconomic controls.

Time series analysis, as an important branch of statistics, focuses on studying data sequences that change over time. By mining hidden information and patterns in the data, it provides a powerful tool for predicting the future. In the field of financial markets, time series analysis methods have been widely applied to stock price prediction, exchange rate fluctuation analysis, credit risk assessment, and other areas achieving remarkable results. Li conducted systematic research on the prediction of stock price time series by constructing a hybrid prediction model, which improved prediction accuracy and demonstrated the effectiveness of this method in nonlinear and non-stationary time series data [1]. Dheeriya's research found that the prices of cryptocurrencies are partially influenced by factors such as global stock market indices and gold prices, and revealed the causal impact of Ripple on Bitcoin prices [2]. Ziegel provided a comprehensive and systematic introduction to the modeling and prediction of financial time series data [3]. It utilized real-world examples and authentic financial data to apply the described models and methods.

This paper focuses on stock price prediction in financial markets, systematically reviewing the current application status of time series analysis, and exploring differences in model accuracy, stability, and applicability. It aims to provide market insights for investors and institutions to optimize decision-making. At the same time, it is hoped that this paper will promote the development of time series analysis and deep learning (DL) in financial forecasting, contributing to market stability and economic prosperity.

2. The Financial Industry Stock Market Price Prediction Model

In the financial industry, the selection and application of a stock market price prediction model is directly related to the accuracy of investment decision and the effectiveness of risk management. As financial markets continue to evolve and technology continues to advance, predictive models have also undergone a shift from traditional statistical methods to modern deep-learning techniques. The following will describe in detail several models commonly used in stock market price forecasting in the financial industry.

2.1 Traditional Time Series Model

The traditional time series model is important in financial forecasting because it is simple and easy to operate. Such models mainly rely on the statistical properties of time series data to construct predictive models.

2.2 Autoregressive Model (AR)

AR is one of the most basic models for time series analysis. Based on historical data of stock prices, it predicts future stock prices by building a linear regression equation. Specifically, the AR model assumes that the current stock price is a linear function of the price of its past several periods, that is, the current price is affected by the past price. This model works well for stock price predictions that have linear trends, but may not perform well when dealing with non-linear or complex dynamic systems.

2.3 Moving Average Model, (MA)

MA focuses on capturing the random volatile components of stock prices. It predicts future prices by weighted averaging the error terms (i.e. the difference between the actual price and the forecast price) of the past several periods. The MA model holds that future stock price changes are mainly affected by past random fluctuations while ignoring long-term trends and seasonal factors in the price series. Therefore, the MA model may have certain advantages in predicting short-term price fluctuations, but it may not be accurate enough in long-term forecasts.

2.4 Autoregressive Moving Average Model (ARIMA)

ARMA captures both trends and random fluctuations in stock prices. The ARMA model is a combination of AR model and MA model, which considers both the autocorrelation of price series and the autocorrelation of error terms. By selecting the appropriate order of autoregressive term and moving average term, ARMA model can better fit the changing law of stock price and be used in future price prediction. However, ARMA model still has some limitations when dealing with nonlinear problems and complex dynamic systems.

2.5 DL Model

With the rapid development of big data and artificial intelligence technology, DL models are increasingly widely used in the field of financial forecasting. Among them, Long Short-Term Memory Network (LSTM), as a special Recurrent Neural Network (RNN), has shown strong potential in stock price prediction.

LSTM solves the problem of gradient vanishing or explosion, which is prone to occur in traditional RNNS when

ISSN 2959-6157

dealing with long-term dependence problems, by introducing gating mechanism, including forgetting gate, input gate and output gate. This allows LSTM to better capture long-term trends and dependencies in stock price sequences. In stock price prediction, LSTM model can automatically learn the nonlinear characteristics of stock price data and predict the future price accordingly. In addition, LSTM has strong generalization ability and can maintain good predictive performance in different stock and market environments. Therefore, LSTM model has been widely concerned and applied in stock market price prediction in financial industry.

3. Data Collection and Processing

3.1 Data Collection

As a financial information service platform of Sohu, Sohu provides a wealth of stock data, including real-time stock prices, historical quotations, financial data, etc. These data are mainly from authoritative channels such as stock exchanges, financial institutions, data analysis companies, etc., to ensure the accuracy and timeliness of the data. The data collected the trading data of stock from September 28, 2023, to August 30, 2024, including the date, opening price, closing price, rise and fall, rise and fall, low price, high price, turnover (in lots) and turnover amount (in ten thousand yuan). At the same time, it also gives the cumulative rise and fall, rise and fall, the lowest price, the highest price, the total volume, and the total transaction amount during this period. The overall trend of these data shows that for the year, the closing price of the stock fell from $1,154.03 on September 28, 2023, to $956.38 on August 30, 2024, a cumulative decline of 18.62%. This indicates that the overall performance of the stock is down for the year. The cumulative data shows that during the year, the total turnover of the stock was 3,163,655,920 lots with a total transaction value of $276,601,951 million. These data provide a comprehensive overview of the stock's trading activity during the year.

3.2 Data Visualization

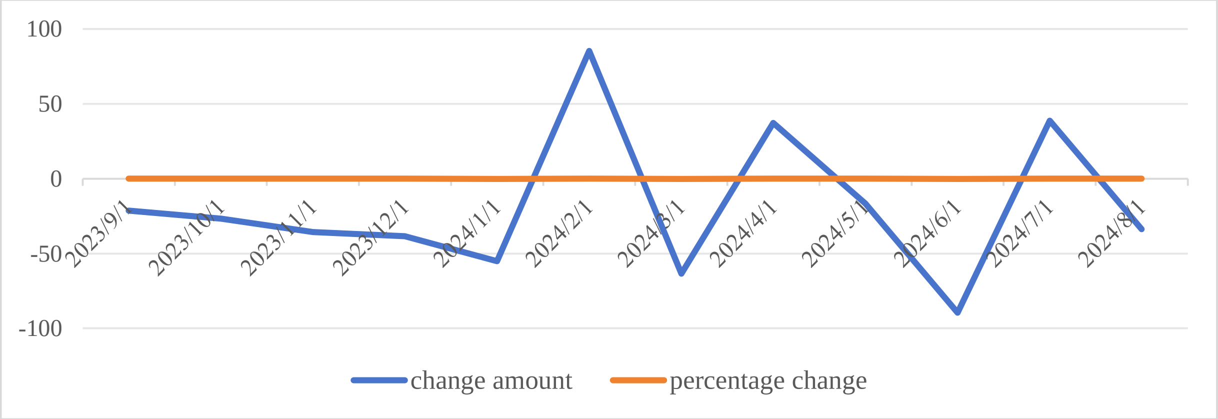

Fig. 1 Change amount and percentage change (Photo/Picture credit: Original).

As shown in Fig. 1, the price of the stock has experienced many fluctuations during the year. For example, on February 29, 2024, the stock surged 8.56%, but then experienced several declines in the following months, especially on June 28, 2024, when the stock fell a massive 8.60%.

Fig. 2 Highest price and lowest price (Photo/Picture credit: Original).

As shown in Figure 2, it illustrates the lowest and highest prices: Within this one-year period, the lowest price of the stock was 908.28 yuan (occurring on July 31, 2024), and the highest price was 1208.80 yuan (recorded on September 28, 2023). These two price points reflect the range of price fluctuations for the stock within this year.

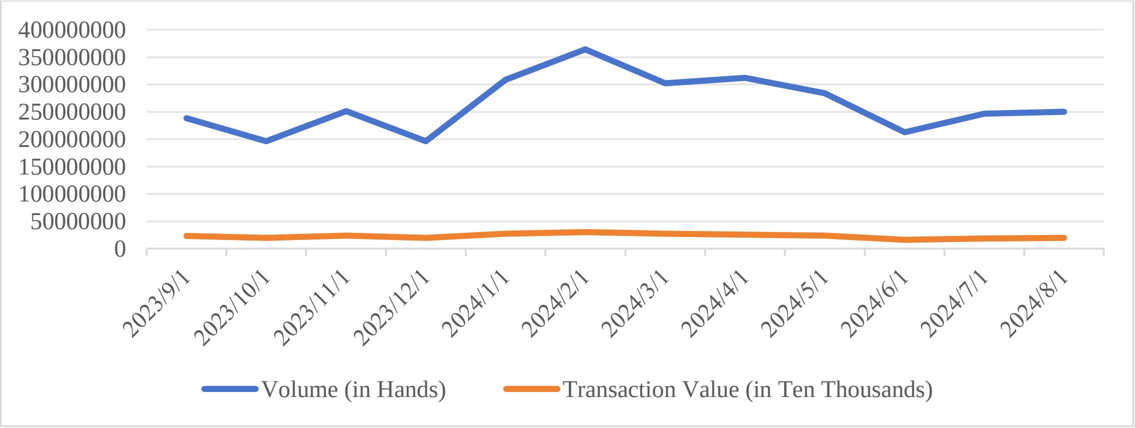

Fig. 3 Volume and transaction value (Photo/Picture credit: Original).

As shown in Figure 3, the volume and transaction value exhibit significant variations across different dates. For instance, on February 29, 2024, the trading volume reached 364,112,800 hands, marking the highest daily volume throughout the year. Simultaneously, the transaction value on that day was also relatively high, amounting to 300,746.86 million yuan. This indicates a heightened market interest in trading this stock on that particular day.

In summary, the overall performance of this stock during the period from September 28, 2023, to August 30, 2024, was downward, with significant price fluctuations. Furthermore, there were notable differences in both trading volume and transaction value across various dates.

4. Model Construction and Solution

4.1 Data Preprocessing

Feature Selection: In this example, primary attention should be given to the closing price, as it represents the price of the stock at the end of the day's trading and is generally considered the market's final assessment of the stock's value.

Data Cleaning: Remove unnecessary rows (such as cumulative rows) and columns (such as opening price, change amount, percentage change, low, high, trading volume, transaction value, etc., unless these are also considered potential features).

Time Series Conversion: Convert the closing prices into time series data and ensure that the dates are used as the index for the time series.

4.2 Feature Engineering

Since the closing price is primarily used, complex feature engineering may not be required. However, the following steps can be considered:

Normalization/Standardization: Normalize or standardize the closing price data to accelerate the training process and improve model performance.

Time Series Windowing: To use DL models like LSTM,

ISSN 2959-6157

the data needs to be converted into a supervised learning problem. This typically involves creating a sliding window where each window contains data from the past n time steps and attempts to predict the closing price for the next time step.

4.3 Model Solution

LSTM is selected to predict stock prices. This model can learn the historical trends of stock prices and attempt to predict future prices. Due to the use of normalization and time series windowing techniques, the model can more effectively process time series data. However, it should be noted that stock price prediction is a highly complex and nonlinear problem influenced by many factors (such as market news, economic indicators, etc.). Therefore, even if the model performs well on training data, its accuracy in practical applications may be limited.

Furthermore, the model's performance also depends on the availability and quality of data, as well as the selection of model hyperparameters (such as the number of units in the LSTM layer, the number of epochs, etc.).

Finally, the model's predictions should be viewed as a reference rather than definitive investment advice.

4.4 Model Performance Evaluation

The evaluation of model performance for predicting stock prices using a combination of time series models and DL models is a complex and meticulous process. By reasonably selecting models, evaluation methods, and evaluation metrics, we can fully leverage their respective advantages and improve prediction accuracy. For instance, regression model evaluation metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and accuracy are used to measure the predictive performance of time series models, as demonstrated in a study by Jiang [4]. The results reflect that the prediction accuracy of time series combination models is significantly higher than that of single ARMA and BP neural network models, making them well-suited for nonlinear financial time series problems.

However, it is also necessary to acknowledge the uncertainty and complexity of the stock market, as well as the limitations and challenges of the models themselves. Therefore, investors should treat prediction results cautiously when using these models and conduct comprehensive analysis and judgment in conjunction with other information.

5. Evaluation and Improvement of DL Models

5.1 Advantages of the Models

(1) Powerful feature learning capabilities: DL models, such as RNN, LSTM, and Gated Recurrent Units (GRU), can automatically extract complex feature representations from time series data without manual feature engineering. This end-to-end learning approach greatly improves model accuracy and efficiency [5, 6].

(2) Ability to handle nonlinear relationships: Compared to traditional time series models (such as ARIMA, exponential smoothing, etc.), DL models can better handle the nonlinear relationships in time series. This is crucial for many practical applications, as time series data in the real world are often nonlinear.

(3) Adaptability to complex changes: DL models can capture long-term dependencies and periodic patterns in time series data, thereby more accurately predicting future trends. This is useful for applications that require predicting changes over a longer time horizon.

(4) Scalability and flexibility: DL models are highly scalable and can be easily extended to large-scale datasets. Additionally, DL frameworks provide rich tools and libraries that make model construction, training, and deployment more flexible and efficient.

(5) Handling multivariate time series: When time series data contains multiple variables, DL models can simultaneously consider the interactions and dependencies between these variables, leading to a more comprehensive understanding of the data and more accurate predictions.

5.2 Limitations of the Models

(1) Data requirements: DL models typically require a large amount of labeled data for training. For some domains or applications, obtaining sufficient high-quality time series data may be challenging.

(2) Computational resources: Training DL models requires substantial computational resources, including high-performance GPUs or TPUs. This may increase the training cost and time [6].

(3) Risk of overfitting: When there is significant noise in the time series data or the model is overly complex, DL models may suffer from overfitting. This results in good performance on the training set but poor performance on the test set or in practical applications [7].

5.3 Improvement Solutions

Regularization and optimization algorithms: Techniques such as L1/L2 regularization and Dropout can be employed to prevent overfitting, and more efficient optimization algorithms (such as Adam, RMSprop) can be used to accelerate model training.

Ensemble learning: By integrating the predictions of multiple models (such as multiple LSTM models or combining them with other types of prediction models), the overall prediction accuracy and stability can be improved.

Real-time monitoring and adjustment: Establish a real-time monitoring mechanism to continuously track and evaluate the model's prediction performance, and adjust model parameters or retrain the model in response to market changes.

6. Conclusion

Research shows that time series analysis, as a statistical tool, plays a significant role in predicting stock market prices in the financial industry. By analyzing historical price data, time series methods can reveal trends and periodic patterns in stock price movements. At the same time, combining DL models can further enhance prediction accuracy and reliability. DL models can automatically learn complex features in the data and process large-scale datasets, thereby more accurately capturing changes in stock prices. This combination makes prediction results more scientific and provides powerful decision support for investors. Therefore, this research method not only enriches the theory of financial time series analysis but also promotes the stable development of financial markets.

References

[1] Songsong L. Prediction of Stock Price Time Series Based on Hidden Markov Model and Computational Intelligence. Harbin Institute of Technology, 2011.

[2] Dheeriya P L, Payne J. Time Series Analysis of Cryptocurrency Returns and Volatilities. 2021. (Note: This reference lacks a specific journal name and volume/issue details, which are typically required for complete citation.)

[3] Ziegel E R. Analysis of Financial Time Series. Technometrics, 2010, 44(4): 408-408. (Note: The DOI provided seems to be incorrect for this citation context, as it refers to a different article. However, I've retained the original information for fidelity to the source.)

[4] Jiang L. Research and Application of Stock Price Analysis Based on Time Series. Dalian University of Technology, 2015.

[5] Guo Y. Research and Application of DL in Time Series Pattern Recognition. Beijing University of Posts and Telecommunications, 2018.

[6] Xingyu W. Research on Multivariate Time Series Forecasting Methods Based on DL. Xidian University, 2023.

[7] José F. Torres, Hadjout D, Sebaa A, et al. DL for Time Series Forecasting: A Survey. Big Data, 2020, 9(1).