Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 3, No. 4, 2024

Deep Learning Approaches for Stock Price Prediction: Methods and Trends

Avery Thompson

Northern Arizona University

9avery.thompson@na.edu

Abstract: The ability to anticipate potential opportunities or crises in the stock market has always been highly valued by investors. This skill became particularly crucial during the Covid-19 global pandemic, as effective risk management became essential for navigating the volatile environment. Beyond traditional business analysis strategies, there is a growing demand for a robust intelligent system capable of accurately predicting stock prices to inform investment strategies. Currently, a significant body of research focuses on predicting stock price trends, predominantly employing deep learning techniques. Despite the success of these studies in achieving favorable outcomes, there has been a scarcity of comprehensive surveys summarizing the deep learning methods utilized in stock price prediction.Therefore, this paper aims to provide a thorough review of the machine learning techniques applied in stock price forecasting, outline the development context of this field, and analyze emerging trends based on previously published research. The reviewed papers were categorized according to the deep learning methods they employed, including Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Recurrent Neural Networks (RNN), and various hybrid deep learning models. Additionally, this paper identifies key datasets, variables, models, and results from each study. The survey presents these findings using widely adopted performance metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Mean Square Error (MSE).

Keywords:machine learning; deep learning; stock price prediction; LSTM; GRU

1.Introduction

The fluctuation of stock prices is often indicative of the global economic environment, signaling periods of prosperity or recession. Traditional stock traders frequently criticize the authenticity and accuracy of modern machine learning techniques for stock price prediction, arguing that the stock market's inherent complexity makes reliable forecasting impossible. However, prior research has demonstrated the practical utility of these machine learning models. Researchers, aiming to predict potential risks that necessitate immediate buying or selling actions, have developed various hybrid models to forecast the trends of candlestick charts for individual stocks. The success of machine learning applications in other time series prediction problems suggests a promising outlook for stock market analysis.

Historically, researchers have extensively utilized support vector machines (SVM) and neural networks for predicting price trends, not just in stock markets but in other domains as well. The performance of a prediction model largely hinges on the features used, leading previous research to focus on historical stock data. To incorporate the influence of public and private information, scholars have worked on leveraging textual data from social media, news, and official company announcements. This information from diverse sources is integrated into machine learning systems.

Deep learning models, in contrast, can integrate all these steps into an end-to-end training model. In recent years, deep learning has not only excelled in computer vision tasks but has also achieved remarkable results in natural language processing. This versatility has encouraged researchers to apply deep learning to various tasks. Early machine learning methods relied on manually constructed features for classification or regression, but their performance was dependent on the quality of these features. Poor feature construction could adversely affect the model's performance. Deep learning addresses this issue effectively by enabling models to automatically learn relevant features through deep neural networks and nonlinear transformations.

The success of deep learning in providing accurate prediction results has garnered global attention, underscoring the potential of these techniques. Properly configured, deep learning can be a powerful tool across numerous academic and industrial fields, significantly advancing human civilization. Although deep learning methods such as Graph Neural Networks (GNN), Gated Recurrent Units (GRU), and Long Short-Term Memory (LSTM) have been successfully applied to forecasting tasks, there is still a need for a comprehensive review summarizing these recent advancements in stock price prediction.

Therefore, this article first provides a detailed introduction to the background of the task, including its subtasks, datasets, and variables. Next, it introduces some of the latest neural network models used in each task, explaining their underlying principles. Additionally, the article reviews recent work, highlighting the problems and methods each task addresses. Finally, based on an analysis of previous research, the article identifies potential challenges in the field. This review aims to serve as a valuable guide for beginners, offering insights into the current state and future directions of stock price prediction using deep learning techniques.

2.Background

2.1.Stock Price Prediction (The overall task)

To effectively trace the empirical evidence of stock market prediction based on machine learning and deep learning models, this survey divides the general task into several subtasks. Five research questions are constructed and listed as follows:

1. What datasets are normally used for forecasting in the stock market?

2. What models are usually adopted to predict the stock market?

3. What different performance parameters are used during the process of the experiment?

4. What metrics are set to serve as a standard to evaluate the effectiveness of these models?

5. What are the most representative works and contributions for stock market forecasting?

These questions will be answered sequentially in the remaining survey.

2.2 .Datasets/Benchmarks

2.2.1.Time Series Data in Finance

The task centers on time series stock data, encompassing various prices at the stock's opening, closing, current closing period, price changes, volume, and daily closing offer price. Predominant research focuses on indices such as Dow Jones, Nasdaq, NYSE, S&P 500, and the Hong Kong Index. These data sets are accessible for different time spans (frequencies) and are sourced from platforms like Google Finance, Bloomberg, and Yahoo Finance.

In addition to price data at different frequencies, the volume, which represents the total number of traded stocks within a period, and the adjusted closing price, which accounts for dividends, stock splits, and new stock offerings, are widely used. For instance, in [1], researchers utilized trading volumes at different periods for the S&P 500 and Dow Jones indices. In [2], researchers incorporated high and low prices with LSTM analysis to predict soybean futures prices on the Chicago Board of Trade (CBOT).

Various papers addressing financial market forecasting employ different frequencies of trading data. Articles [3] and [4] utilized daily stock data in their research. In [5], data from both Nikkei 225 and S&P 500 were used for prediction with deep learning models. Researchers in used data from Google Finance with a frequency of 15 minutes, deriving technical indicators and incorporating sentiment embeddings for market prediction. In [6], financial news was employed to perform experiments on data from the Hong Kong Stock Exchange.

2.2.2. Textual Data

The task then incorporates textual data into the analysis. Previous research has shown that combining textual data with historical prices enhances the performance of forecasting models. Textual data comes from various sources, such as internet comments, news sources, company reports, financial reports, blogs, and discussion boards in stock trading applications. Textual information from news, financial blogs, online sentiments, and discussion boards has been utilized for stock market prediction.

News data from journals, newspapers, and websites are among the most common sources of textual data. For example, financial news data were extracted in [7] and [8] from different financial news websites to predict future stock price trends. In [9], a sentiment analytical tool was constructed and applied to a financial news article prediction system. In [10], both news data and technical features were used in a simple regressive model to make predictions.

Nowadays, influential individuals frequently share their opinions about political or financial events on social media, potentially impacting other stock traders' decisions. For instance, [11] demonstrates that President Trump's tweets can provide information for short-term market movements. Social media platforms are extensively used for sentiment analysis due to their capacity to capture personal opinions. Specifically, researchers in [12] used Twitter data to evaluate public sentiment's effect on market predictions. Similarly, [13] utilized tweet data to analyze the relationship between textual data and stock price trends for market prediction.

Company disclosures are more reliable sources as they usually contain the latest officially confirmed information, such as quarterly earnings, board adjustments, and business challenges. For instance, textual data from company announcements were analyzed in [14] to understand their impact on both short-term and long-term stock index predictions. Additionally, based on instant announcements from Thomson Reuters' website, [15] constructed a decision support model that included sentiment analysis focusing on negation scope detection.

2.3.Metrics

After collecting data and incorporating it into the assigned machine learning or deep learning models, it is essential to establish a standard for evaluating the models' performance. In the remainder of Section 2, four of the most widely used performance metrics for model evaluation are briefly introduced to facilitate a better understanding of the tasks. The performance metrics commonly adopted in most research are:

1. Mean Absolute Percentage Error (MAPE): Measures the accuracy of the model by calculating the percentage error between predicted and actual values.
2. Mean Absolute Error (MAE): Represents the average absolute difference between predicted and actual values, providing a straightforward measure of prediction accuracy.
3. Mean Squared Error (MSE): Computes the average squared difference between predicted and actual values, emphasizing larger errors by squaring them.
4. Root Mean Squared Error (RMSE): The square root of the MSE, providing an error metric that is in the same units as the predicted values, which helps in interpreting the model's accuracy.

2.3.1.MAPE

Also known as mean absolute percentage deviation (MAPD), the mean absolute percentage error (MAPE) is a standard metric for evaluating the accuracy of a prediction model. It expresses the performance of the model as a ratio, defined by the formula:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

A_t stands for the actual value, and F_t represents the forecast value. The difference between A_t and F_t is divided by the actual value A_t . The absolute value of this ratio is summed for every forecasted point in time and then divided by the number of fitted points n . MAPE is widely used in stock market forecasting due to its ability to provide a clear and interpretable measure of prediction accuracy.

2.3.2.MAE

Mean Absolute Error (MAE), on the other hand, is a measure of errors between paired observations expressing the same phenomenon.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

Where y_i is the prediction and x_i represents the true value. The mean absolute error (MAE) shares the same scale as the data being measured. MAE is a widely used standard for computing forecast error in stock analysis and is sometimes used in conjunction with the more standard definition of mean absolute deviation.

2.3.3. MSE

The next common measure for the difference/distance between two datasets is the mean squared error (MSE). MSE is computed as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \widehat{Y_i})^2$$

3. Neural Network Models

Various methods and techniques applied to price prediction analysis are presented in the literature. The literature also considers and categorizes some statistical techniques, machine learning, and deep learning methods. Although a subset of machine learning, deep learning techniques have also been specifically introduced due to their recent popularity.

As a particular type of machine learning technique, deep learning displays great potential and flexibility by acquiring information to split the task into a nested hierarchy of concepts. It represents the overall task as different concepts, with each concept defined in relation to simpler ones, and more obscure notations computed in terms of more comprehensible ones. The dominant characteristic of deep learning methodology is its ability to represent the semantic composition provided by vectorization and neural processing. Nonlinear network topology in neural networks manages to model complex real-world data by extracting features that capture distinguishing information, creating a more robust model than traditional methods.

The successful application of deep learning models in various fields, such as speech processing and image recognition, has made them an ideal alternative for time series analysis. The adoption of different learning models, such as artificial neural networks (ANN), convolutional neural networks (CNN), deep belief networks (DBN), recurrent neural networks (RNN) [30], and long short-term memory (LSTM), underscores the capability of deep learning algorithms in real-world applications. In [30], researchers applied RNN for learning characteristics and used a reinforcement learning composition for making trading decisions with deep representations.

Numerous experiments have been conducted based on the above machine learning strategies. Below, the paper mainly elaborates on two deep learning architectures commonly used in predicting future stock market trends – LSTM and GRU.

3.1.Long-short term memory (LSTM)

Long Short-Term Memory (LSTM) is an effective model for sequence prediction problems due to its gate structure, which efficiently extracts relevant information while discarding outdated data. The core of LSTM revolves around a memory cell that maintains its state over time and nonlinear gating units that regulate the flow of information into and out of the cell.

In the context of a stock market scenario, LSTM is well-suited for predicting stock prices based on several dependencies:

a) The pattern the stock has been following in recent days (uptrend or downtrend). b) The stock price on the previous day. c) Factors influencing the current day's price, such as market disturbances, changes in company leadership, or fluctuations in company profits. d) Global disturbances, such as wars or pandemics.

These dependencies can be generalized as follows:

a) Previous cell state: The information previously stored in the cell (e.g., stock prices from previous days). b) Previous hidden state: The output of the previous cell. c) Current input state: New information fed into the model.

The LSTM's operation formula begins with the forget gate, which is calculated as follows:

$$f_t = (W_f * [h_{t-1}, x_t] + b_f)$$

Input gate:

$$i_t = (W_i * [h_{t-1}, x_t] + b_i)$$

$$\widetilde{c_t} = \tanh(W_c * [h_{t-1}, x_t]) + b_c$$

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c_t}$$

Output gate:

$$t = (W_o * [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(c_t)$$

While simple models of RNNs can process original text data sequentially to acquire context-based features, issues like vanishing gradients and dependencies on short contexts may still impact forecasting performance. To address these issues, an LSTM-RNN hybrid model has been introduced.

Specifically, LSTM utilizes memory cells to process data with long dependencies [16], leveraging hierarchical structures and multiple hidden layers for feature engineering. For example, in [17], scholars constructed an LSTM-based model that uses parameters such as stock prices and backlog to predict a single stock price. In [18], a hybrid of CNN and LSTM is implemented, where CNN is applied for stock selection, and LSTM is responsible for price prediction. In [19], technical indicators and stock data were used as input for LSTM, and the results were tested on the Bovespa index from the BM&F Ibovespa stock exchange. Moreover, researchers in [20] utilized an LSTM model for single stock market index forecasting, revealing that LSTM outperforms most memory-free models such as random forest and logistic regression classifiers.

In [21], paragraph vectors were incorporated with LSTM in their model, which was later tested on textual data from different companies in the Tokyo Stock Exchange. This study also demonstrated that LSTM can learn effectively even when the input has large dimensions.

4.Conclusion

Stock market prediction is a captivating field of study due to its significant impact on personal financial management. However, financial markets often feature noisy data and unstable fluctuations, prompting the development of modern forecasting models that incorporate textual data for improved prediction accuracy. This paper provides a summary of various machine learning strategies based on previous research. The main contribution of this paper is to introduce current techniques related to adapted methods using various performance metrics. It identifies both financial time series data and textual data, categorizing the techniques used in stock market prediction across different machine learning algorithms. Some selected surveys have adopted hybrid approaches in the financial market, successfully combining the strengths of their original strategies to produce desirable outcomes. When implemented correctly, hybrid models have significant potential and can outperform traditional machine learning methods in several ways.Moreover, LSTM and GRU techniques are commonly used for achieving accurate stock market predictions. The primary challenge in stock market prediction is that existing data mining strategies and dimensionality reduction methods cannot keep pace with advanced machine learning algorithms. Additionally, factors such as domestic governmental decisions and consumer sentiments also influence stock markets. In the future, researchers will strive to address existing problems and improve mechanisms to create a more trustworthy, accurate, and predictable stock market system.

References

[1] Ankit Thakkar and Kinjal Chaudhari. 2020. Predicting stock trend using an integrated term frequency- inverse document frequency-based feature weight matrix with neural networks. Appl. Soft Comput. 96 (2020), 106684

[2] Chenhao Wang and Qiang Gao. 2018. High and Low Prices Prediction of Soybean Futures with LSTM Neural Network. 2018 IEEE 9th International Conference on Software Engineering and Service Science(ICSESS) (2018), 140–143.

[3] Wang, J., Hong, S., Dong, Y., Li, Z., & Hu, J. (2024). Predicting Stock Market Trends Using LSTM Networks: Overcoming RNN Limitations for Improved Financial Forecasting. Journal of Computer Science and Software Applications, 4(3), 1-7.

[4] T. Matsubara, R. Akita, and K. Uehara. 2018. Stock Price Prediction by Deep Neural Generative Model of News Articles. IEICE Trans. Inf. Syst. 101-D (2018), 901–908.

[5] X. Li, J. Cao, and Z. Pan. 2018. Market impact analysis via deep learned architectures. Neural Computingand Applications (2018), 1–12.

[6] L. Shi, Z. Teng, L. Wang, Y. Zhang, and Alexander Binder. 2019. DeepClue: Visual Interpretation of Text-Based Deep Stock Prediction. IEEE Transactions on Knowledge and Data Engineering 31 (2019), 1094–1108.

[7] Y. Peng and H. Jiang. 2016. Leverage Financial News to Predict Stock Price Movements Using Word Embeddings and Deep Neural Networks. ArXiv abs/1506.07220 (2016).

[8] R. P. Schumaker, Y. Zhang, C. Huang, and H. Chen. 2012. Evaluating sentiment in financial news articles. Decis. Support Syst. 53 (2012), 458–464.

[9] J. L. Wu, C. Su, L. Yu, and P. C. Chang. 2012. Stock Price Predication using Combinational Features from Sentimental Analysis of Stock News and Technical Analysis of Trading Information. In EconomicsDevelopment and Research.

[10] Kun Yuan, G. Liu, J. Wu, and Hui Xiong. 2020. Dancing with Trump in the Stock Market. ACM Transactions on Intelligent Systems and Technology (TIST) 11 (2020), 1 – 22.

[11] J. Bollen, H. Mao, and X. Zeng. 2011. Twitter mood predicts the stock market. ArXiv abs/1010.3003 (2011).

[12] B. Li, K. Chan, C. X. Ou, and R. Sun. 2017. Discovering public sentiment in social media for predicting stock movement of publicly listed companies. Inf. Syst. 69 (2017), 81–92.

[13] S. Feuerriegel and J. Gordon. 2018. Long-term stock index forecasting based on text mining of regulatory disclosures. Decis. Support Syst. 112 (2018), 88–97.

[14] N. Pröllochs, S. Feuerriegel, and D. Neumann. 2016. Negation scope detection in sentiment analysis: Decision support for news-driven trading. Decis. Support Syst. 88 (2016), 67–75.

[15] Sun, M., Feng, Z., Li, Z., Gu, W., & Gu, X. (2024). Enhancing Financial Risk Management through LSTM and Extreme Value Theory: A High-Frequency Trading Volume Approach. Journal of Computer Technology and Software, 3(3).

[16] Yang, H., Zi, Y., Qin, H., Zheng, H., & Hu, Y. (2024). Advancing Emotional Analysis with Large Language Models. Journal of Computer Science and Software Applications, 4(3), 8-15.

[17] S. Liu, C. Zhang, and J. Ma. 2017. CNN-LSTM Neural Network Model for Quantitative Strategy Analysisin Stock Markets. In ICONIP.

[18] D. Nelson, A. M. Pereira, and R. A. de Oliveira. 2017. Stock market's price movement prediction with LSTM neural networks. 2017 International Joint Conference on Neural Networks (IJCNN) (2017), 1419–1426.

[19] T. Fischer and C. Krauß. 2018. Deep learning with long short -term memory networks for financial market predictions. Eur. J. Oper. Res. 270 (2018), 654–669.

[20] R. Akita, A. Yoshihara, T. Matsubara, and K. Uehara. 2016. Deep learning for stock prediction using numerical and textual information. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS) (2016), 1–6.