ITM Web of Conferences **70**, 04008 (2025) https://doi.org/10.1051/itmconf/20257004008
DAI 2024

Enhancing Stock Price Forecasting Accuracy Using LSTM and Bi-LSTM Models

Hao Wang

School of Information, Beijing Wuzi University, 101149 Beijing, China

Abstract. Accurately predicting stock price trends is of critical importance in the financial sector, enabling both individuals and enterprises to make informed and profitable decisions. In recent years, researchers have employed a variety of techniques to forecast stock market trends, yet the challenge of improving accuracy remains. This research introduces an innovative approach to predicting stock prices, employing two sophisticated models: Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (Bi-LSTM) networks. Through rigorous analysis, the research demonstrates that, with proper hyperparameter tuning, LSTM models are capable of making highly accurate predictions of future stock trends, a capability that is also exhibited by Bi-LSTM models. The study evaluates the models by measuring the Root Mean Square Error (RMSE) while varying key factors. Publicly available stock market information, such as the highest and lowest prices, and opening and closing prices, is utilized for evaluating model effectiveness. The results indicate that the Bi-LSTM model is superior to the LSTM model in terms of RMSE, making it a more effective methodology for stock price forecasting and aiding in strategic decision-making.

1 Introduction

The advent of the Internet has resulted in a heightened interdependence between stock markets and the digital domain, with the former becoming increasingly susceptible to influences emanating from the latter. Stock markets occupy a pivotal position within a country's financial system, and fluctuations in stock prices can serve as indicators of broader economic trends. Previous researchers have typically regarded the trajectory of stock prices as a significant economic concern [1]. However, in the context of accelerated global economic growth, the size of the stock market has gradually expanded, leading to a more intricate financial landscape. The forecasting of stock markets represents one of the most significant challenges currently facing the financial markets. This is attributable to the fact that time series data is characterized by a multitude of variables and complex properties that characterize these markets [2]. These factors are closely associated with changes in stock markets and lead to high volatility and uncertainty, which in turn make stock forecasting a challenging endeavour.

Corresponding author: wanghao1204@ldy.edu.rs

In the past few years, a range of deep learning algorithms have been utilized in different global stock exchanges. To assess the true impact of integrating financial news sentiment into stock market prediction, a study was carried out by Shahi et al. [3]. The normalized stock market forecasting results of the Gated Recursive Unit (GRU) and Long Short-Term Memory (LSTM) models were evaluated under identical control conditions. Kamalov et al. assessed the performance of various neural network architectures in forecasting stock prices regarding their future value [4]. Shen et al. collected data spanning two years from the Chinese stock market and proposed an extensive, feature-engineered, personalized, deep-learning model that exhibits strong overall accuracy in predicting stock market trends [5]. Sunny et al. demonstrated precise forecasting of future stock movements through the application of a recurrent neural network (RNN) model, which comprised an LSTM model and a Bidirectional Long Short-Term Memory (Bi-LSTM) model [6]. Gülmez et al. demonstrated the efficacy of an artificial rabbit optimization algorithm model (LSTM-ARO) in predicting stock market prices [7]. This was achieved by processing data with multiple input and output time steps. Htun et al. concentrated on the diverse machine learning-based techniques employed in stock market forecasting [8]. Their research demonstrated the benefits of utilizing popular feature selection and extraction techniques based on different deep-learning methodologies for stock price prediction. Niu et al. advanced an innovative hybrid variational mode decomposition (VMD) model that fuses a variational model with an LSTM network decomposition [9]. The model that demonstrated such remarkable predictive capabilities was designated VMD-LSTM.

The main goal of this research is to evaluate the practicality and efficiency of LSTM and Bi-LSTM models in forecasting stock prices. Initially, the study synthesizes and organizes relevant concepts and background information essential for understanding stock price prediction. Following this, a comprehensive analysis of the core technologies underlying deep learning is provided, including an introduction to the fundamental principles that drive these models. The research then compares and analyses the advantages of LSTM and Bi-LSTM models. The findings suggest that deep learning methodologies have the potential to achieve a substantial level of accuracy in predicting stock prices, highlighting their significance in financial forecasting.

The section starts by providing an overview of utilizing deep learning for forecasting stock prices and then proceeds to thoroughly examine fundamental concepts and principles in Chapter 2. Chapter 3 analyses the methodology of this study in detail and systematically discusses the experimental results, while Chapter 4 provides recommendations and summarizes the study. This investigation presents a comprehensive evaluation of the utilization of deep learning models, illustrating their potential to enhance the precision of stock price predictions.

2 Methodology

2.1 Dataset description and preprocessing

The data set was obtained from Yahoo Finance, a publicly accessible source [10] (Some of the data are presented in Table 1). The experiment was conducted using Google stock market data from 19 August 2004 to 4 October 2019. The data were presented in numerical format. The dataset contains a wide range of information about the stock, which can be employed in the prediction of future trends. The total number of days included in the data set is 4,170. Initially, the experiment meticulously pre-processed the data using various methods and subsequently partitioned the processed dataset into two segments: a test dataset and a training dataset. The test set encompasses 12% of the data, while the remaining 88% is allocated to the training set. Then, the training data set and a variety of different tuning parameters of the LSTM are combined with the Bi-LSTM model. Finally, an overall predicted stock price is generated. Subsequently, this prediction dataset was compared with the test dataset to evaluate prediction accuracy.

Table 1. Some examples of the data.

Date	Open	High	Low	Close	Adj Close	Volume
Aug 19, 2004	2.49	2.59	2.39	2.50	2.50	897,427,216
Aug 20, 2004	2.52	2.72	2.50	2.70	2.69	458,857,488
Aug 23, 2004	2.76	2.83	2.72	2.72	2.72	366,857,939
Aug 24, 2004	2.77	2.78	2.58	2.61	2.61	306,396,159
Aug 25, 2004	2.61	2.69	2.59	2.64	2.64	184,645,512
Aug 26, 2004	2.61	2.69	2.61	2.69	2.68	142,572,401

2.2 Methods

The objective of this research is to utilize deep learning models to predict future movements in the stock market. In particular, the efficacy of LSTM and Bi-LSTM models for stock price forecasting will be evaluated. The study utilized Google stock market data obtained from Yahoo Finance and employed the min-max feature ratio method to partition the dataset into a training set comprising 88% of the data and a test set containing 12%. The experiment took place within the Google Collaboratory simulation environment, utilizing the TensorFlow framework to train and predict using LSTM and Bi-LSTM models with various tuning parameters. Finally, model-generated predictions were compared with test dataset results to assess prediction accuracy (Fig. 1 displays the study's flowchart).

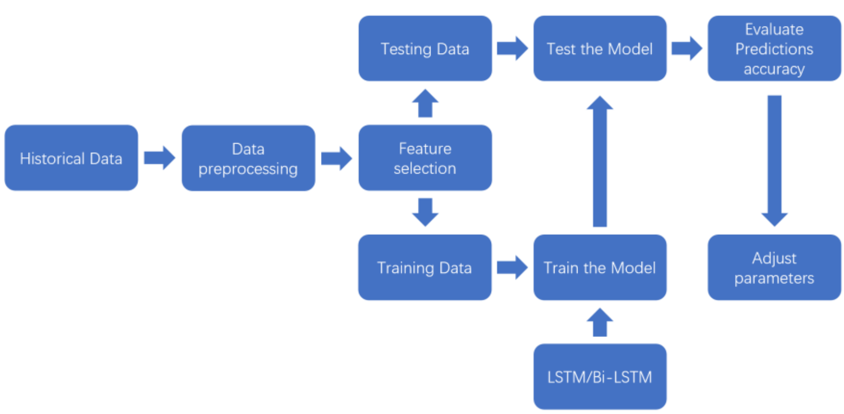

Fig. 1. Research Flowchart (Picture credit: Original).

2.2.1 LSTM

LSTM networks represent a specialized type of RNN first described by Hochreiter and Schmid Huber in 1997. Since their initial description, LSTM have grown in popularity within the field of artificial intelligence [11]. In particular, it has been developed to solve problems associated with time series forecasting. This pertains to the capacity to retain information for extended periods when processing continuous data. For this investigation, the LSTM analysis utilizes time series data, and the objective is to extract the sequential features of the data. In this study, the input to the LSTM is the adjusted closing price and volume data. LSTM networks are capable of capturing long-term dependencies. They are furnished with a memory unit capable of retaining information over an extended period. One limitation of traditional RNN is that they are susceptible to gradient disappearance and explosion when training the model on long sequences. The LSTM network addresses this issue by employing a gating mechanism that enables selective recall or forgetting of information. Furthermore, LSTM enable the model to capture and retain crucial contextual information even when there are significant time intervals between related events in the sequence. It should be noted, however, that LSTM do not present a wholly unproblematic solution. In comparison to more straightforward architectures, such as feedforward neural networks, the computational expense is greater. Furthermore, the training of LSTM networks is more time-consuming than that of simpler models due to their increased computational complexity.

The Bi-LSTM model represents a variant of the LSTM model. Integrating forward LSTM and backward LSTM concurrently captures both forward and reverse information in the sequence, facilitating a more comprehensive extraction of textual information. Bidirectional RNNs (BRNNs) can access all preceding information while predicting future data, thereby overcoming the constraints associated with traditional RNN. The architecture of the Bi-LSTM model comprises three key layers: an embedding layer for word representation, a bidirectional LSTM layer for processing input sequences in both forward and reverse directions, and a fully connected layer for generating prediction results. The word embedding layer converts each word into a corresponding vector representation, while the bidirectional LSTM layer processes both forward and reverse input sequences before the fully connected layer generates the final prediction outcome. Bi-LSTM models are extensively employed in natural language processing and time series forecasting tasks. To optimize the Bi-LSTM model, the subtractive averaging algorithm can be employed to dynamically adjust the parameters of the model, thereby enhancing the robustness and accuracy of the model. This approach enables the model to adaptively handle noise and outliers and to adapt to new data rapidly.

This research employs unique data processing methods for individual models to evaluate the feasibility and effectiveness of LSTM and Bi-LSTM models in forecasting stock prices based on actual stock price datasets. Furthermore, it highlights the impact of deep learning algorithms in producing predictions using a wide range of time series data.

3 Result and Discussion

3.1 Discussion of results

This research opted to utilize Root Mean Square Error (RMSE) as the assessment measure for examining the effectiveness of the predictive LSTM and Bi-LSTM models following appropriate data scaling, training, and testing.

In Table 2, as the number of epochs increases continuously in this study, the RMSE of the trained model under LSTM fitting initially exhibits an increase followed by a gradual decrease, deviating from the anticipated results. Conversely, the Bi-LSTM fitted model aligns with expected outcomes, demonstrating a decreasing trend in RMSE with increasing epochs. Fig. 2 also illustrates this pattern. Furthermore, a favorable association has been observed between the quantity of epochs and training time. Fig. 3 and Fig. 4, which depict LSTM models with two and four hidden layers, span 50 epochs, respectively, and demonstrate this association. Fig. 5 and Fig. 6 present Bi-LSTM models with varying numbers of hidden layers, trained for a duration of 50 epochs. The outcomes obtained from these models exhibit comparable findings.

Table 2. An analysis of the RMSE at several epochs in the LSTM and Bi-LSTM models.

NO. of Epochs	LSTM RMSE	Time (min)	Bi-LSTM RMSE	Time (min)
10	0.0011000	3	0.0007167	8
20	0.0007250	6	0.0006459	15
50	0.0004933	15	0.0004219	40
100	0.0004928	30	0.0004127	70
250	0.0031980	75	0.0003568	200

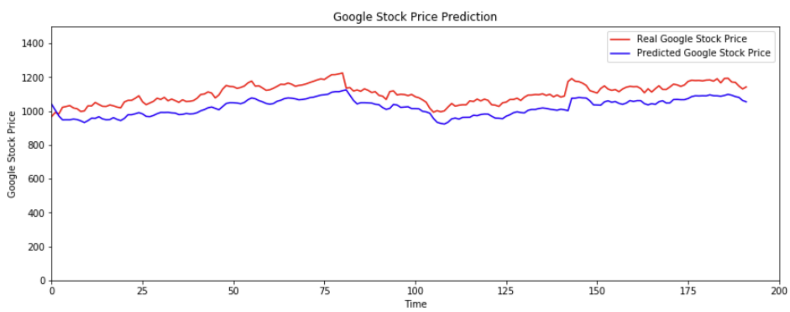

Fig. 2. The output prediction of the LSTM model after 250 epochs using two hidden layers (Picture credit: Original).

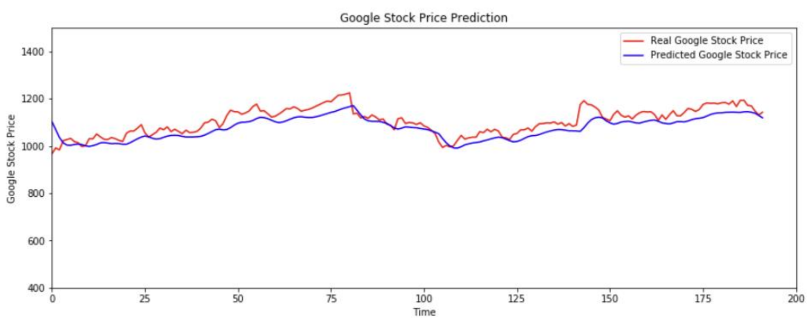

Fig. 3. The output prediction of the LSTM model after 50 epochs using two hidden layers (Picture credit: Original).

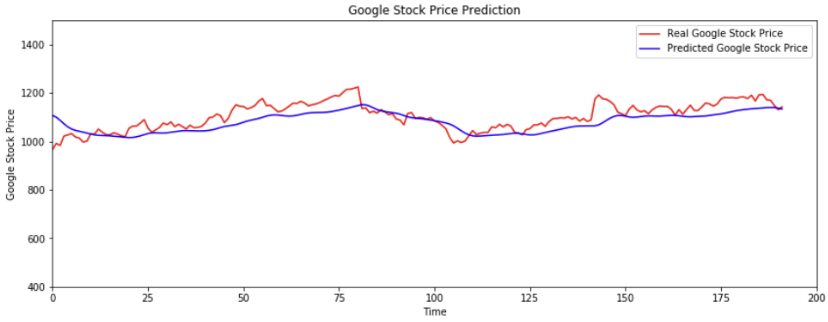

Fig. 4. The output prediction of the Bi-LSTM model after 50 epochs using four hidden layers (Picture credit: Original).

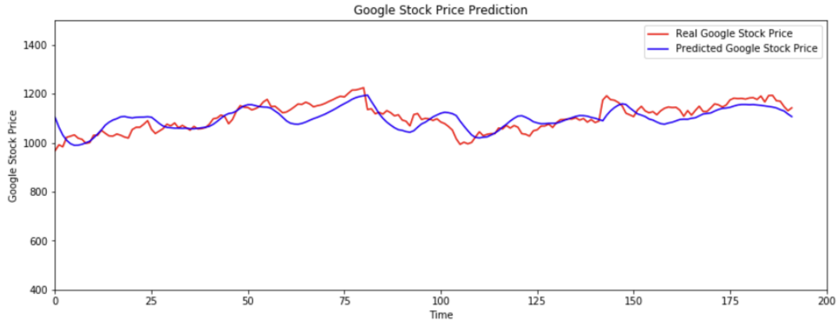

Fig. 5. The output prediction of the Bi-LSTM model after 50 epochs using two hidden layers (Picture credit: Original).

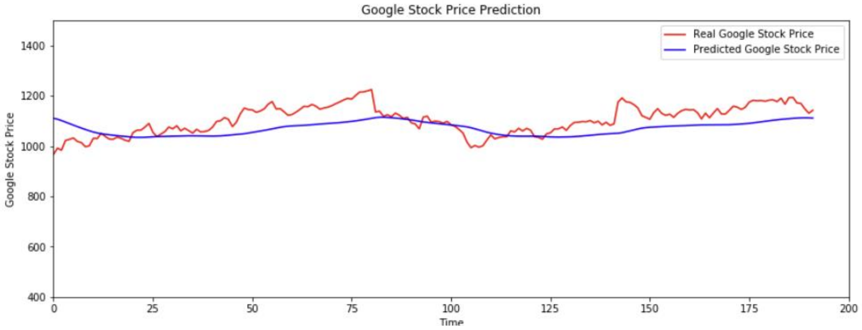

Fig. 6. The output prediction of the Bi-LSTM model after 50 epochs using four hidden layers (Picture credit: Original).

3.2 Discussion

In this study, both LSTM networks & Bi-LSTM networks, two deep learning techniques for stock price prediction, exhibit their distinct theoretical and practical advantages as well as limitations. LSTM networks solve problems such as gradient vanishing that RNN often encounter when learning long-term dependencies. This feature makes LSTM networks particularly adept at capturing temporal dynamics in stock prices that are strongly influenced by historical data. The memory cells in LSTM can extend the retention time of key information, making them suitable for analysing continuous datasets describing trends, cycles, and sudden fluctuations in stock prices. The Bi-LSTM network extends the standard LSTM by bi-directional processing of inputs involving both forward and backward directions. This bi-directional strategy helps to capture non-sequential dependencies by considering both past and future contexts when predicting stock prices. However, due to the presence of multiple unpredictable factors in financial markets, future price movements usually remain highly uncertain. To address these challenges, future research might investigate how the features of Bi-LSTM can be incorporated while preserving the benefits of LSTM with other machine learning techniques (e.g., reinforcement learning or attentional mechanisms), which are effective in dealing with complex spatiotemporal dependencies while enhancing the interpretability of models.

This study investigates the utilization of deep learning methodologies, including LSTM networks and Bi-LSTM networks, for the purpose of stock price prediction. Both approaches offer distinct theoretical strengths and practical benefits but also have inherent limitations.

LSTM networks are particularly effective at addressing the gradient vanishing problem commonly encountered in traditional RNN when modelling long-term dependencies. This characteristic enables LSTM to capture temporal dynamics in stock prices, which demonstrates considerable fluctuation in the case of extensive and complex historical information. The memory cells in LSTMs extend the retention period for crucial information, making them well-suited for analysing continuous datasets that reflect trends, cycles, and sudden fluctuations in stock prices. Bi-LSTM networks enhance traditional LSTM by integrating bidirectional input processing. This bidirectional approach improves the model's ability to incorporate both past and future contexts into stock price predictions, thereby enhancing its capacity to capture non-sequential dependencies. However, the inherent unpredictability of financial markets means that future price movements remain highly uncertain. In order to tackle these difficulties, it is essential to conduct additional research into the combination of LSTM and Bi-LSTM models with supplementary machine learning methods, including reinforcement learning or attention mechanisms. These methodologies are renowned for their adeptness in handling intricate spatiotemporal dependencies and enhancing model interpretability, potentially resulting in superior performance when predicting stock prices.

4 Conclusion

This research emphasizes the substantial influence of deep learning algorithms on predicting stock prices, especially by creating forecasting models based on time series data. The findings demonstrate that deep learning algorithms, notably LSTM and Bi-LSTM models, offer superior accuracy compared to traditional regression models. The research underscores the importance of parameter tuning for achieving accurate forecasting results. Effective calibration is essential as it has a direct impact on the predictive precision of the models. Research findings demonstrate that, under identical parameters, the Bi-LSTM model consistently exhibits a lower RMSE in contrast to the LSTM model. This indicates that the Bi-LSTM model is more effective for stock price prediction, making it advantageous for individual investors and businesses engaged in forecasting the stock market. An extensive experimental evaluation of the Bi-LSTM model was conducted, and the results corroborate its superior performance over the LSTM model, evidenced by a reduced RMSE. These findings have the potential to enhance investment strategies and improve financial outcomes for investors. Future studies will prioritize the continued enhancement of the robustness and applicability of the Bi-LSTM model, with ongoing investigations aimed at refining its performance and exploring additional improvements in stock price forecasting.

References

1. R. Vanaga, & B. Sloka, Financial and capital market commission financing: aspects and challenges. Journal of Logistics, Informatics and Service Science, 7(1), 17-30 (2020)
2. A.W. Li, & G.S. Bastos, Stock market forecasting using deep learning and technical analysis: a systematic review. IEEE access, 8, 185232-185242 (2020)
3. T.B. Shahi, A. Shrestha, A. Neupane, & W. Guo, Stock price forecasting with deep learning: A comparative study. Mathematics, 8(9), 1441 (2020)
4. F. Kamalov, L. Smail, & I. Gurrib, Stock price forecast with deep learning. International Conference on Decision Aid Sciences and Application, 1098-1102 (2020)
5. J. Shen, & M.O. Shafiq, Short-term stock market price trend prediction using a comprehensive deep learning system. Journal of big Data, 7, 1-33 (2020)
6. M.A.I. Sunny, M.M.S. Maswood, & A.G. Alharbi, Deep learning-based stock price prediction using LSTM and bi-directional LSTM model. In 2020 2nd novel intelligent and leading emerging sciences conference, 87-92 (2020)
7. B. Gülmez, Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. Expert Systems with Applications, 227, 120346 (2023)
8. H.H. Htun, M. Biehl, & N. Petkov, Survey of feature selection and extraction techniques for stock market prediction. Financial Innovation, 9(1), 26 (2023)
9. H. Niu, K. Xu, & W. Wang, A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network. Applied Intelligence, 50, 4296-4309 (2020)
10. Yahoo, "Yahoo finance", 2020, Retrieved on 2024, Retrieved from https: //finance.yahoo.com/quote/GOOG/history?p=GOOG
11. S. Hochreiter, & J. Schmidhuber, Long short-term memory. Neural computation, 9(8), 1735-1780 (1997)