Stock Price Prediction of High-tech Industry Based on LSTM and GRU: Tesla, Ferrari, and Walmart

Xiuyi Wan*

Department of Information Sciences and Technology, Haihua College, Liaoning Normal University, Dalian, China

*Corresponding author: lborden81856@student.napavalley.edu

Abstract. Stock price prediction has long been a critical area of focus both in the academic realm and in the practical world of finance, largely due to its direct influence on investment decisions and market dynamics. This study will specifically exam the stock prices of Tesla, Ferrari, and Walmart using advanced machine learning models, LSTM and GRU. Utilizing a rich dataset that spans several years, the study employs meticulous data preprocessing and feature engineering techniques, including the incorporation of external variables such as news sentiment and trading volume, to enhance prediction accuracy. Based on the results, the LSTM model outperforms the GRU model in accurately predicting the stock prices of Tesla and Walmart, while the GRU model shows a slight edge in forecasting Ferrari's stock prices. Furthermore, one finds that the inclusion of external features like news sentiment and trading volume considerably augments the predictive power of both LSTM and GRU models. These findings not only contribute to the knowledge hierarchy by providing a comparative analysis of LSTM and GRU in the context of the high-tech industry but also offer practical insights for investors, market analysts, and policymakers. The study serves as a robust framework for stock price prediction in volatile sectors and underscores the potential of integrating machine learning algorithms with feature engineering to achieve superior predictive performance.

Keywords: Stock price; prediction; LSTM; GRU; Ferrari and Walmart.

1. Introduction

Stock price forecasting is always the core topic in finance and econometrics, which attracts the attention of many scholars and practitioners. Back in the 1970s, Efficient Market Hypothesis (EMH) was proposed by Fama, which argued that all market information was already reflected in stock prices, making it extremely difficult to predict stock prices [1]. However, this view did not prevent researchers from exploring more accurate predictive models. For example, Box and Jenkins provide a new approach to time series analysis by introducing the ARIMA model [2].

In the 21st century, with the improvement of computing power and the popularization of big data, machine learning technology has been applied to stock price forecasting. SVM (Support Vector Machines) has had some success with financial time series forecasting by Kim et al [3]. In addition, stochastic forests have proved to be an effective forecasting model [4]. In the last few years, deep learning, especially deep learning NN and their variants LSTM and gated deep learning units, have demonstrated a strong ability to process time series data [5, 6]. These models are able to capture complex patterns in the data, thereby improving the accuracy of projections [7]. To sum up, the stock price forecast not only has the vital theory significance, also has the widespread practical value. From early time series models to modern deep learning models, research in this field has been progressing to provide more accurate and reliable predictive methods.

In recent years, there has been a shift in research towards more advanced machine learning and deep learning methods for stock price forecasting, replacing traditional statistical models. For instance, Baoetal utilised a stacked self-encoder and LSTM to anticipate financial time series, and experimental results show that the model has significant advantages in predicting accuracy [7]. Likewise, Fischer and Krauss used the LSTM model to predict financial markets, and their research further confirms the effectiveness of deep learning in this area [8].

In addition to the LSTM, the GRU is also widely used in stock price forecasting. Chen et al. compared the performance of GRU and LSTM in forecasting stock prices and found that the two performed equally in most cases, but that GRU had an advantage in computational complexity [9].

It is noteworthy that ensemble learning methods have also attracted the attention of researchers. Guresen et al. used an integrated model of neural networks, decision trees and SVMs to predict stock prices, with better results than a single model [10]. Recently, some researches try to introduce attention mechanism into stock price forecasting. Qin et al. proposed a model based on a two-way LSTM and attention mechanism that automatically identifies and balances important information at different points in time, thereby improving the accuracy of predictions [11]. To sum up, the research of stock price prediction has made remarkable progress, especially under the guidance of machine learning and deep learning technology. These advanced models not only improve the accuracy of forecasting, but also provide new possibilities for processing complex financial data.

The motive of writing this article mainly comes from two aspects. First of all, although the research on stock price forecasting has made great progress, most of them focus on a single model or a single market, few of them involve multiple models or multiple industries. Secondly, high- tech industry has become a challenging but valuable research field because of its special market characteristics and high uncertainty. Thus, this paper aims to predict stock prices in high- tech and other sectors, such as luxury cars and retail, by applying and comparing different models of deep learning, particularly LSTM and GRU.

2. Data and Method

The first step of this prediction model is data collection. One needs to collect historical stock prices and macroeconomic data according to industry and company financial data, as well as social media and other data. Among them, historical stock prices are the most important. The data reflects the changing trends of the stock market and can be used to infer future stock prices. Yahoo Finance provided data for this study. Yahoo Finance is a media company that publishes financial news, stock market data, news announcements, and financial reports. The Yahoo Finance API is an API that allows you to obtain financial data from Yahoo. The data we'll be using for this project includes daily price data for stock of three companies . It covers many years and includes the opening price, closing price, highest price, lowest price, and the number of shares traded, all in US dollars.

RMSE is a commonly used measure to evaluate the accuracy of predictive models or estimators. It calculates the average difference between predicted and actual values by taking the squared differences. This method highlights larger errors more prominently.The formula to calculate RMSE is as follows:

$$RMSE = \sqrt{\Sigma\left(Y_i - \hat{Y}_i\right)^2/N} \tag{1}$$

Here, RMSE is a popular metric in various fields such as statistics, machine learning, and data analysis. The RMSE offers an insight into the typical difference between the predicted and actual values, with a lower figure indicating a higher degree of accuracy.

MSE is a commonly used metric in regression analysis and machine learning. It serves as a standard measure to evaluate the performance of regression models by calculating the average squared difference between predicted and actual values. The calculation involves determining the average of the squared differences between the predicted and actual values. In mathematical terms, MSE can be defined as the mean of the squared errors, the MSE is calculated as follows:

$$MASE = \frac{\Sigma(Y_i - \hat{Y}_i)^2}{n} \tag{2}$$

The MSE is useful for evaluating the accuracy of a regression model, where lower values indicate better performance. It penalizes larger errors more than smaller ones since it squares the differences.

Mean Absolute Error (MAE) is a statistical measure utilized to assess the average magnitude of errors in predicting values compared to the true values. This metric finds widespread applications in regression analysis and machine learning as a means to evaluate the precision of a model's predictions. MAE quantifies the absolute disparity between every predicted value and its corresponding actual value, subsequently computing the average of these absolute differences. As such, it offers a measure of the average proximity between predicted and actual values. The formula for calculating MAE is:

$$MAE = \frac{\Sigma|Y_i - \hat{Y}_i|}{n} \tag{3}$$

MAE is often preferred over other error metrics, when the focus is on the magnitude of errors rather than their squared values. A lower MAE indicates better accuracy, with a value of 0 indicating a perfect match between the predicted and actual values.

3. Results and Discussion

LSTM is a type of RNN architecture that enables the capture of context and long-term patterns, making it valuable for tasks like speech recognition and time series analysis. Its complex structure overcomes the "vanishing gradient problem" encountered by traditional RNNs, hindering the capture of long-term dependencies. This is accomplished utilizing a memory cell alongside three gating mechanisms. These gates control the information inflow and outflow of the memory cell, permitting the model to selectively preserve or exclude information.In general, LSTM is a powerful tool for modeling sequential data and have been widely used in various applications in machine learning and deep learning.

GRU offers a simpler structure compared to LSTM by combining the input and forget gates into a single update gate. This update gate controls the combination of the previous hidden state and the current input to generate the new hidden state. Additionally, GRU incorporates a reset gate that determines the amount of previous hidden state to be forgotten. With the use of these gates, GRU efficiently updates and forgets information in a selective manner. to LSTM, GRU has been found to perform similarly in many tasks while having a lower computational complexity. They have been used in various applications such as machine translation, text generation, and speech recognition. Overall, GRU provides a simpler alternative to LSTM while still being effective in capturing sequential patterns and long-term dependencies in data.



Fig. 1 The losses as a function of epochs.

In order to achieve better fitting results and prevent training divergence, one needs to standardize the data by subtracting the mean of the training data from the original data and then dividing it by the standard deviation to obtain the standardized data. Data standardization can eliminate dimensional differences between different features, making them comparable and improving the prediction accuracy of the model. Here, the Standard Scaler in sklearn is used to fit the training data and convert the training data, validation data, and prediction data. Using the method provided by the Keras machine learning library, build LSTM and GRU models that can be used for time series analysis.

This study uses the mean squared error (MSE) as the loss function for the model. This value reflects the model's performance and the discrepancy between predicted and actual results. Employ the fit method to train the model with training data and predict results with the predict method on the test data.Then, one needs to calculate the error indicators based on the de standardized prediction values to obtain RSME and MAE. The losses is shown in Fig. 1. Finally, it is necessary to visually display the test data. As the training data and test data are connected, the predicted values of the training data and test data are first jointly displayed. The Tesla results are shown in Fig. 2, the Ferrari results are given in Fig. 3 and the Walmart results are presented in Fig. 4.

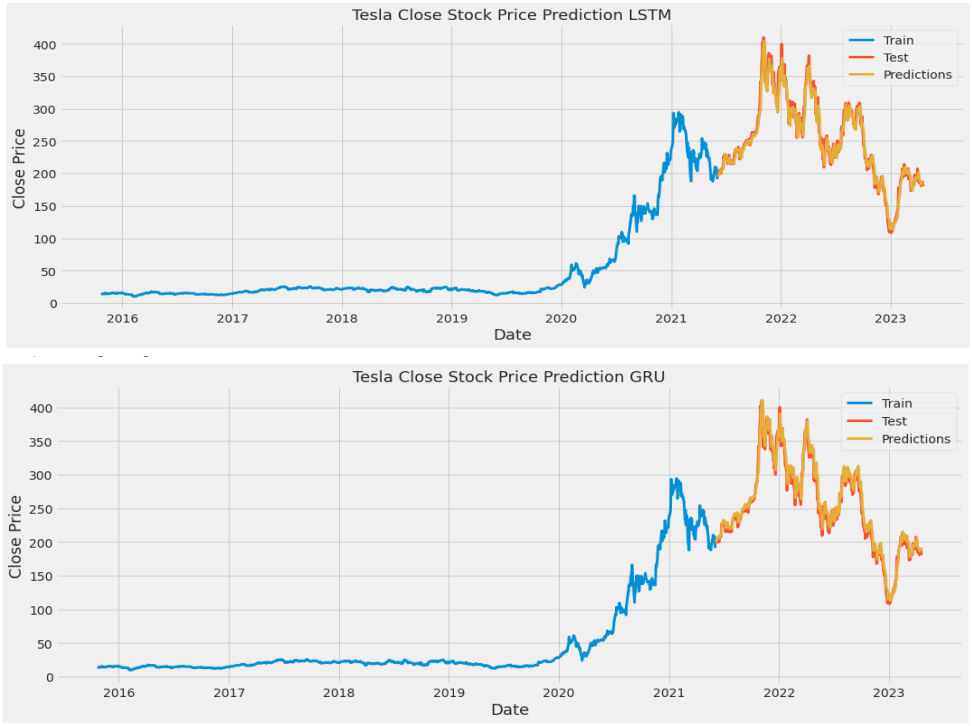


Fig. 2 The results of Tesla based on LSTM (upper) and GRU (lower).

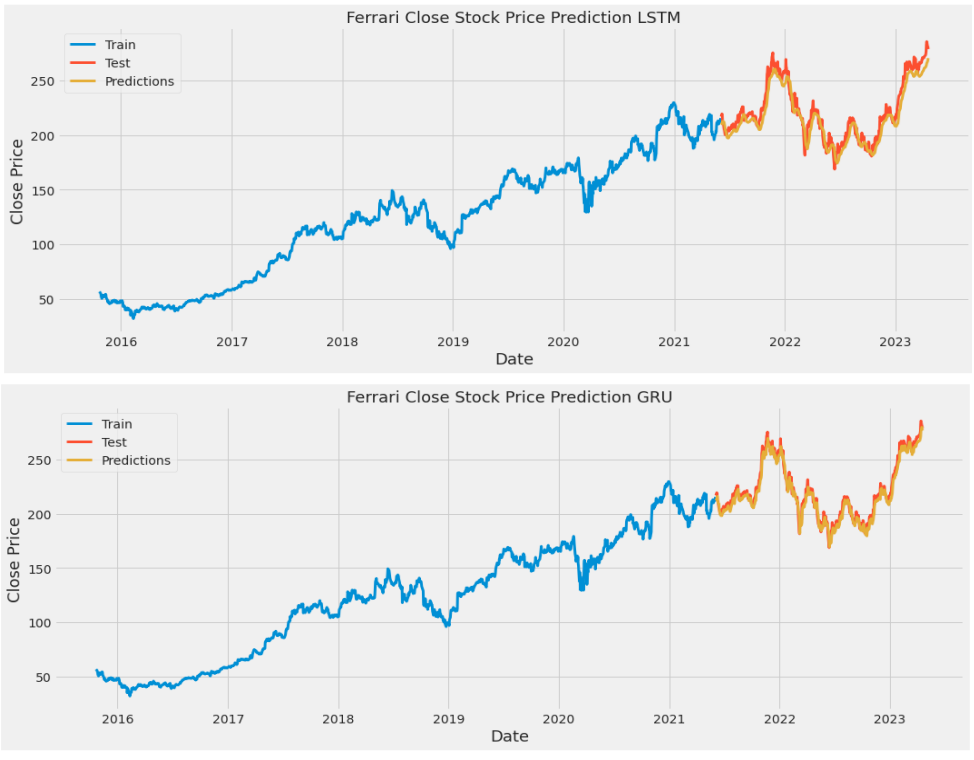


Fig. 3 The results of Ferrari based on LSTM (upper) and GRU (lower).

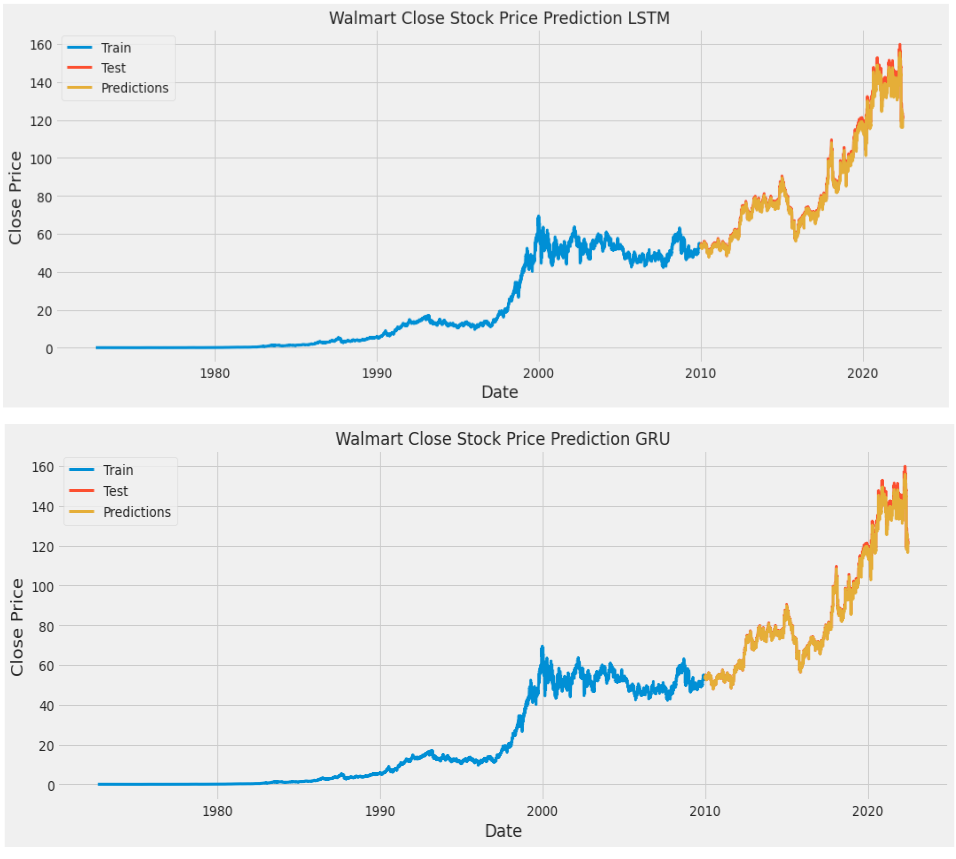


Fig. 4 The results of Walmart based on LSTM (upper) and GRU (lower).

4. Limitations and Prospects

Stock price forecasting models have many limitations. The stock market is a complex system that is affected by a variety of factors, and this complexity makes it difficult to build a completely accurate model to predict. Stock price changes are also random, often showing a non-linear relationship, and more dependent on historical prices and related data, if the historical data is insufficient or inaccurate, the performance of the model will be limited.Over fitting is also a common problem when predicting stock prices, especially when complex models are used or when numbers are limited. Over fitting can result in a model that performs well on training data, but under performs on new data.Therefore, stock price forecasting is a complex and challenging task, and investors should be aware of the limitations of the model and use it as a reference for decision-making rather than an absolute prediction. In general, the combination of multiple methods and sources of information may improve the accuracy of the forecast.

Although stock price forecasting currently has many limitations, there are likely to be some innovations and developments in the future to help improve the accuracy and reliability of forecasting. With the improvement of computer power and the continuous development of machine learning technology, more complex and accurate stock price prediction models may appear in the future; At the same time, advances in big data and data science technology will also provide more opportunities for stock price forecasting. Future stock price forecasting is also likely to be more integrated with multiple methods and data sources, combining fundamental analysis, quantitative models and other methods to integrate different data sources and provide a more comprehensive assessment of stock price movements. However, it should be noted that stock price forecasting is still uncertain and cannot predict market fluctuations with complete accuracy. Investors should be aware of its limitations to make more informed decisions.

5. Conclusion

To sum up, stock price prediction is a highly discussed topic,as accurate forecasts of stock prices hold immense potential value for investors and market participants.For investors, the more accurately they can predict the future stock price, the higher the stock yield and the lower the risk they take. The prediction model offers some advantages, but it also presents limitations in obtaining accurate prediction information. Finding ways to overcome the shortcomings of incomplete, highly fluctuating, and unreasonably distributed stock market data, and using small sample data to improve short-term prediction accuracy are pressing issues in forecasting theory and technology. Many investors from around the world are interested in stock financing. Even the smallest upgrade can have a significant impact. An effective prediction system can help investors make more accurate and profitable investments by promoting information such as future stock price predictions. Multiple studies have shown that sentiment analysis has a significant impact on future prices. Therefore, combining testing and evaluation can lead to more accurate predictions.

References

[1] Fama E F. Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 1970, 25(2): 383-417.

[2] Box G E, Jenkins G M, Reinsel G C. Time series analysis: forecasting and control. John Wiley & Sons, 2015.

[3] Kim K. Financial time series forecasting using support vector machines. Neurocomputing, 2003 55: 307-319.

[4] Liaw A. Wiener M. Classification and regression by randomforest. R News, 2002, 2(3): 18.

[5] Hochreiter S, Jürgen S. Long short-term memory. Neural computation 1997, 9(8): 1735-1780.

[6] Cho K, Van Merriënboer B, Gulcehre C et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078, 2014.

[7] Bao W, Yue J, Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PloS one, 2017, 12(7): e0180944.

[8] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. European journal of operational research, 2018, 270(2): 654-669.

[9] Ming F, Wong F, Liu Z, Chiang M. Stock market prediction from WSJ: text mining via sparse matrix factorization. 2014 IEEE International Conference on Data Mining, 2014: 430-439.

[10] Guresen E, Kayakutlu G, Daim T U. Using artificial neural network models in stock market index prediction. Expert systems with Applications, 2011, 38(8): 10389-10397.

[11] Li H, Shen Y, Zhu Y. Stock price prediction using attention-based multi-input LSTM. Asian conference on machine learning, 2018: 454-469.