ISSN 2616-5902 Vol. 5, Issue 10: 72-77, DOI: 10.25236/AJBM.2023.051011

Predicting TSLA Stock Price Based on LSTM and GRU Models

Yishan Zhang[1,a,*], Yuge Cui[1,b]

[1]Actuarial Science, Xi'an Jiaotong-Liverpool University, Renai Roud, Suzhou, China
[a]Yishan.Zhang21@student.xjtlu.edu.cn, [b]Yuge.Cui21@student.xjtlu.edu.cn
**Corresponding author*

Abstract: *In recent years, advancements in artificial intelligence technologies have facilitated their successful application in time series forecasting and analytical tasks. With the objective of fostering growth in the stock market, this study employs two well-established deep learning forecasting models, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), to predict TESLA stock prices. Initially, historical data is utilized to project future stock prices, and a comparative analysis of the two models is conducted, concluding that a hybrid LSTM+GRU model demonstrates superior performance compared to a standalone LSTM model when provided with identical data. Furthermore, to enhance prediction accuracy, time series data is smoothed using either moving average or exponential average techniques, resulting in a significant improvement in the performance of the LSTM model as compared to both its previous performance and the LSTM+GRU model.*

Keywords: *Tesla, Green Stock, Stock Market, LSTM, GRU*

1. Introduction

The prioritization of economic growth in nations, coupled with an increasing demand for green development, has resulted in the rapid expansion of the green stock market. As a global leader in the electric vehicle industry, Tesla's stock serves as a representative example of green stocks. The stock market is characterized by its non-linear, non-continuous, and volatile nature, influenced by a multitude of factors such as market sentiment, financial operations, political events, rumors, news, and company transactions [1].

Various studies have employed different approaches to predict stock prices. Alkhatib et al. [2] used six variables (High, Low, Open, Volume, HiLo, and OpSe) to enhance prediction accuracy; Ratchagit & Xu [3] employed a two-delay combination model and compared individual and combined prediction methods; Albahli [4] proposed a stock market prediction method based on the ELM classifier and used ELM and RNN for stock price prediction; Lv et al. [5] developed a LightGBM-optimized LSTM for short-term stock price prediction, demonstrating superior accuracy. Wang et al. [6] utilized candlestick chart data to develop a neural network with enhanced attention for predicting stock market price fluctuations. Additionally, numerous studies have explored the impact of human emotions on green stock market price fluctuations, employing various sentiment analysis techniques [1][7][8].

Despite the complexity and unpredictability of the stock market, predicting its movements remains a critical strategy for achieving profitability and promoting economic growth across nations. As such, this study aims to use a deep learning model, specifically the LSTM model, to predict the development trends of Tesla's stock based on historical data, thereby assisting investors in reducing investment risks. This paper delves into the application of the LSTM model for predicting Tesla's stock prices, discussing the methodology, data preparation, and evaluation of model performance. Furthermore, this study seeks to provide a comprehensive guide for those interested in utilizing the LSTM model for stock price prediction.

This paper first introduces the model principles based on the LSTM, then describes the specific experimental process, compares the prediction results of LSTM and LSTM & GRU, and finally obtains the model with the highest accuracy of predicting stock price with historical data.

ISSN 2616-5902 Vol. 5, Issue 10: 72-77, DOI: 10.25236/AJBM.2023.051011

2. Methodology

In this section we fist introduce LSTM model and then introduce LSTM+GRU model.

2.1. LSTM Model

Hochreiter and Schmidhuber [9] proposed LSTM and then Graves [10] improved and promoted it. In this paper, the LSTM model learns the correlation and rules between different times by training data of historical stock price data. The LSTM model inputs the current stock price data into the model, and then predicts the future stock price.

Cell state, input gate and forget gate are C_t, i_t and f_t, which maintain and control information. At time t, x_t is the input data, h_t represents the current output, C_t is the value from the input gate, tan*h* is hyperbolic tangent function, σ is the sigmoid function, W represents the matrix weight, and b is the bias. The operation formula of LSTM is as follows.

Forget gate:

$$f_t = \sigma(W_f \cdot \ [\ h_{t-1}\ , x_t\ ] + b_f) \tag{1}$$

Input gate:

$$i\ _t = \sigma(W_t\ \cdot [\ h_{t-1}\ , x_t\ ] + b_i) \tag{2}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

Cellular Sate:

$$C_t = f_t C_{t-1} + i \cdot \tilde{C}_t \tag{4}$$

Output Gate:

$$o_t = \sigma\big(W_o \cdot \big[h_{t-1}\ , x_t\ \big] + b_o\big) \tag{5}$$

$$h_t = o_t tanh(C_t) \tag{6}$$

2.2. LSTM+GRU Model

GRU is one of the variants of RNN which is introduced by Cho et al. [11]. Combining LSTM GRU will have more preferable memory ability and fewer parameters. The LSTM+GRU model just adds the reset gate based on LSTM.

In LSTM+GRU, the input gate decides how much input x_t and previous output h_{t-1} to be passed the next cell and the reset gate is used to determine how much of the past information to forget. The weight W determines the current memory content to ensure that only relevant information needs to be passed to the next iteration.

The main operations in LSTM+GRU are governed by the following formulae.

Input gate:

$$i\ _t = \sigma(W_t\ \cdot [\ h_{t-1}\ , x_t\ ] + b_i) \tag{7}$$

Forget gate:

$$f_t = \sigma(W_f \cdot \ [\ h_{t-1}\ , x_t\ ] + b_f) \tag{8}$$

Cellular sate:

$$\tilde{C}_t = tanh(W_C \cdot [r_t \cdot C_{t-1} + i \cdot h_{t-1}, x_t] + b_r) \tag{9}$$

Reset gate:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{10}$$

Output Gate:

$$o_t = \sigma\big(W_o \cdot \big[h_{t-1}\ , x_t\ \big] + b_o\big) \tag{11}$$

$$h_t = o_t tanh(C_t) \tag{12}$$

ISSN 2616-5902 Vol. 5, Issue 10: 72-77, DOI: 10.25236/AJBM.2023.051011

Among them, it, ft, ot, and rt are the vectors of the input gate, the forget gate, the output gate, and the reset gate. Wi, Wf, Wc, Wo, and Wr are weight matrices, bi, bf, bc, br and bo is offset vector, ht-1 is the hidden state of the previous time step, xt is the input of the current time step, σ is the sigmoid function, tanh is the hyperbolic tangent function.

The gated structure in the neural unit structure of LSTM and LSTM+GRU neural network uses Sigmoid activation function and Tanh activation function.

The Sigmoid function is:

$$f(x) = \frac{1}{1+e^{-x}} \tag{13}$$

And the tanh function is:

$$\mathrm{f(x)} = \frac{\mathrm{e^{x}-e^{-x}}}{\mathrm{e^{x}+e^{-x}}} \tag{14}$$

The difference between the two functions is that the Sigmoid function has a range of (0,1), while the tanh function has a range of (1,1). Both activation functions increase with the increase of variable x to realize the gating structure.

3. Experimental and Analysis

3.1. Dataset

Considering the significant growth of Tesla's stock in recent years, we chose daily TSLA price among green stocks. After the stock selection, the historical stock prices are fetched from Yahoo Finance. The data is saved to a TXT file and contains all five prices, including everyday stock information which are open, high, low, close, volume, and Open Interest. It has 3219 daily records of 1 targeted stock from 2010-6-30 to 2023-4-13. And the data form is shown in the Table 1.

Table 1: Dataset form

Date	Open	High	Low	Close	Volume
2010-10-01	1.379333	1.383333	1.354000	1.373333	8965500
2010-10-11	1.362667	1.380000	1.338000	1.349333	2568000
2010-10-12	1.346667	1.352000	1.335333	1.349333	3660000
2010-10-13	1.376000	1.390000	1.357333	1.369333	4773000
2010-10-14	1.400000	1.402000	1.360000	1.383333	4422000

3.2. Experimental procedure

In this paper the flow chart of stock price prediction is shown in Fig.1.

In the stage of data pre-processing, we use the closing price as the main research object and then divide the entire data set into a train set, a validation set, and a test set. The train set contains the first 70% of the whole data set, the test set contains the last 10% of the whole data set, then the validation set contains the rest of 20% of the whole data set. And in our LSTM model for price prediction, the length of one sequence is 60.

The data need to be normalized before inputting it into the model. Normalization of the characteristics of a numeric type can unify all the characteristics into roughly the same numeric interval. The normalization of data can eliminate dimensionality, avoid the dependence of data on the selection of measurement units and help to improve the performance of the model, facilitate the training of the model, and speed up the convergence of the model. The normalization methods adopted in this paper are:

1) Use MinMaxScaler to scale the data in the same range;

2) Train the scaler with training data and smooth data.

Since we are dealing with a continuous output here, we can use MSE as our loss function. The optimizer's job is to tune the weights and biases of the network. We chose to use the Adam optimizer.

ISSN 2616-5902 Vol. 5, Issue 10: 72-77, DOI: 10.25236/AJBM.2023.051011

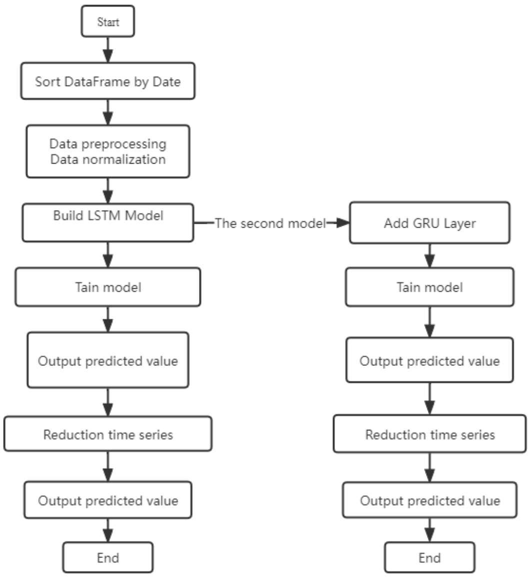
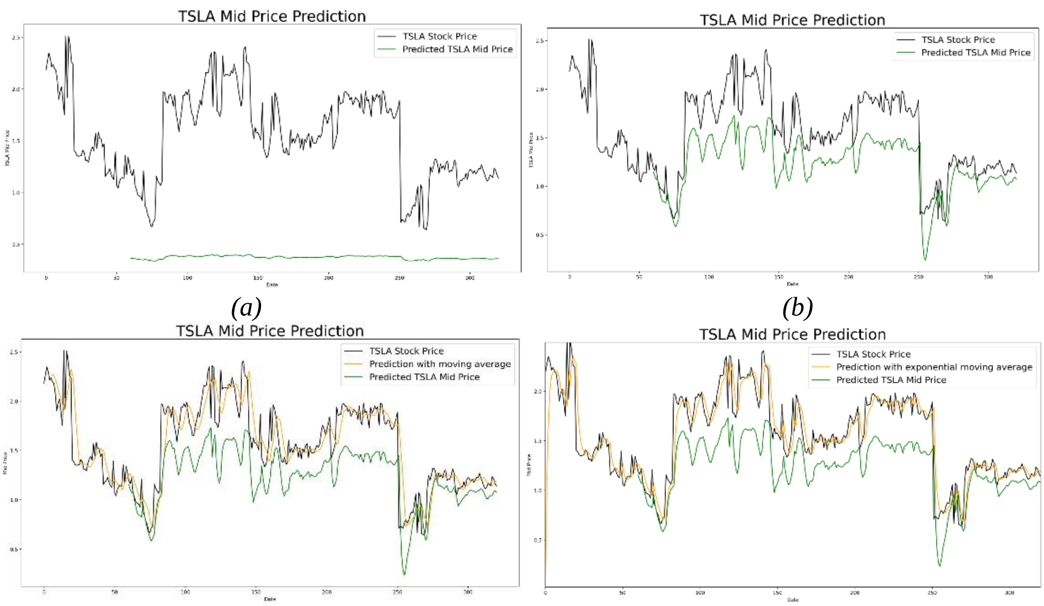


Figure 1: Flow Chart of Experimental.

$$MSE = \frac{1}{n}\sum_{n=1}^{n}(y_k - \widehat{y_k}) \tag{15}$$

In the formula, n is the number of test datasets; y_k is the true value of the k-th sample point; $\widehat{y_k}$ is the model predicted value of the k-th sample point.

3.3. Results



(a) *(b)*

(c) *(d)*

Figure 2: TSLA Mid Price Prediction Based on LSTM.

TSLA stock price prediction using the LSTM model is shown in Fig.2. In Fig.2, subfigure (a) shows the result of TSLA mid-price prediction for 20 epochs in the green line. Subfigure (b) shows the result of TSLA mid-price prediction for 200 epochs in the green line. In subfigure (c), the green line shows the result of TSLA mid-price prediction for 200 epochs and the orange line shows the result of TSLA mid-price prediction with moving average; in subfigure (d), the green line shows the result of TSLA mid-price prediction for 200 epochs and the orange line shows the result of TSLA mid-price prediction with exponential moving average. In subfigure(c) and subfigure (d), the orange lines follow the trend of the black line more closely than the green lines, which shows that LSTM performs much better after smoothing the time series with moving average or exponential average.

TSLA stock price prediction using the LSTM and GRU models is shown in Fig.3 in the green line. The green line has an obvious hysteresis phenomenon to the black line.

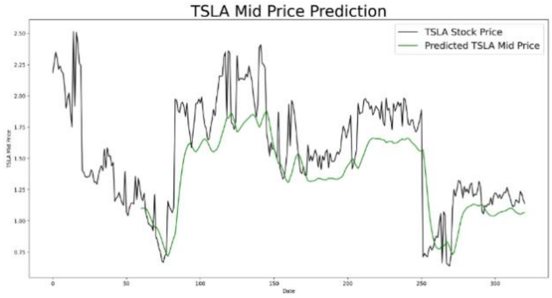


Figure 3: TSLA Mid Price Prediction Based on LSTM&GRU.

3.4. Discussion

The accuracy score of models is shown in Table 2.

Table 2: Accuracy score of Models.

Model	Accuracy score
LSTM on 20 epochs	0.531
LSTM on 200 epochs	
LSTM&GRU	0.550

The accuracy score is a score obtained by considering factors such as prediction accuracy and the hysteresis phenomenon. With the epochs increasing from 20 to 200, the accuracy of the prediction of the results based on the LSTM model has improved. The accuracy score of the LSTM+GRU model is lower than LSTM on 200 epochs. The prediction results of the LSTM+GRU model have a more serious hysteresis phenomenon, but higher prediction accuracy compared to the results of LSTM. The reason for the lower accuracy score of the LSTM+GRU model compared to that of the LSTM model lies in the significant hysteresis phenomenon exhibited by the former.

4. Conclusions

In this study, leveraging the characteristics of stock time series and incorporating deep learning theory, the Long Short-Term Memory (LSTM) and LSTM+Gated Recurrent Unit (LSTM+GRU) models are employed to predict stock prices and subsequently compared. The findings demonstrate that the accuracy score of the LSTM+GRU model is higher than that of the LSTM model without data smoothing, indicating superior performance when using the same data. When time series data is smoothed utilizing moving average or exponential average methods, the LSTM model's performance significantly improves, surpassing both its previous performance and that of the LSTM+GRU model. However, it is important to note that these models can only predict the future price trends of stocks and not their precise values.

Future research will focus on extracting relevant features from online user comments on stocks, news reports, company statements, and major events, as well as linguistic information obtained from stock websites, forums, and post bars. This information will be incorporated into the model to generate relatively more accurate stock price predictions.

References

[1] Bhadamkar A. and Bhattacharya, S. (2022) 'Tesla Inc. Stock Prediction using Sentiment Analysis', Australasian Accounting Business & Finance Journal, 16(5), pp. 52–66. doi:10.14453/aabfj.v16i5.05 (Accessed: 22 April 2023).
[2] Alkhatib K. et al. (2022) 'A New Stock Price Forecasting Method Using Active Deep Learning Approach', Journal of Open Innovation, 8(2), p. 96. doi:10.3390/joitmc8020096 (Accessed: 22 April 2023).
[3] Ratchagit M. and Xu H. (2022) 'A Two-Delay Combination Model for Stock Price Prediction',

ISSN 2616-5902 Vol. 5, Issue 10: 72-77, DOI: 10.25236/AJBM.2023.051011

Mathematics (2227-7390), 10(19), p. 3447. doi:10.3390/math10193447 (Accessed: 22 April 2023).
[4] Albahli S. et al. (2022) 'A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data', Electronics (2079-9292), 11(20), p. 3414–N.PAG. doi: 10.3390/ electronics 11203414 (Accessed: 22 April 2023).
[5] Lv J. et al. (2021) 'An Economic Forecasting Method Based on the LightGBM-Optimized LSTM and Time-Series Model', Computational Intelligence & Neuroscience, pp. 1–10. doi:10.1155/2021/8128879 (Accessed: 22 April 2023).
[6] Wang J. et al. (2022) 'Predicting Stock Market Volatility from Candlestick Charts: A Multiple Attention Mechanism Graph Neural Network Approach', Mathematical Problems in Engineering, pp. 1–16. doi:10.1155/2022/4743643 (Accessed: 22 April 2023).
[7] Ho T.-T., Huang Y. (2021) 'Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation', Sensors (14248220), 21(23), p. 7957. doi:10.3390/s21237957 (Accessed: 22 April 2023).
[8] Kamal S. et al. (2022) 'Trading Stocks Based on Financial News Using Attention Mechanism', Mathematics (2227-7390), 10(12), p. 2001–N.PAG. doi:10.3390/math10122001 (Accessed: 22 April 2023).
[9] Hochreiter S. and Schmidhuber J. (1997) 'Long short-term memory', Neural Computation, 9(8), p. 1735. doi:10.1162/neco.1997.9.8.1735 (Accessed: 22 April 2023).
[10] Graves A. (2012) Supervised Sequence Labelling with Recurrent Neural Networks. [electronic resource]. 1st ed. 2012. Springer Berlin Heidelberg (Studies in Computational Intelligence: 385). Available at: https://search-ebscohost-com.ez. xjtlu.edu.cn/l ogin.aspx? direct= true&db= cat01010a& AN= xjtlu.0001536424&site=eds-live&scope=site (Accessed: 22 April 2023).
[11] Cho K. et al. (2014) 'Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation'. Available at: https:// search-ebscohost-com. ez.xjtlu.edu.cn/login.aspx? direct=true&db=edsarx&AN=edsarx. 1406.1078&site=eds-live&scope=site (Accessed: 22 April 2023).