Predicting Tesla's stock based on machine learning

He Yu *

Business School, The University of Nottingham, Nottingham, United Kingdom

* Corresponding Author Email: alyhy29@nottingham.ac.uk

Abstract. This article emphasizes that when using economic theories to predict stocks, a certain logical sequence needs to be followed. Among the many well-known stocks, Tesla's stock is highly volatile. If the stock prediction can be relatively accurate, it will better demonstrate the feasibility of using machine learning to predict stocks. Therefore, the paper takes the prediction of Tesla's stock using machine learning as an example. First, machine learning was used to analyze the data distribution, and then it was used to analyze the data relationships. Based on the results of the first two steps, the model needed for prediction was derived, and then the prediction model (LSTM) was used to accurately predict the stock price fluctuations of Tesla over the past five years. The experimental results indicate that the model has good predictive performance. Finally, it points out the current shortcomings of using machine learning to predict stocks and the future directions for improvement.

Keywords: Predicting Stock, LSTM Model, Machine Learning.

1. Introduction

At present, artificial intelligence plays a significant role in people's society. And machine learning, as an indispensable part of artificial intelligence, is also playing an important role, especially in the aspect of predicting [1]. Data prediction is also indispensable for daily life in society. It is not only used for predicting stock data, but can also be applied to predicting industrial engineering, future weather, natural disasters, and so on. Therefore, developing well the data prediction based on machine learning is not only beneficial to the economy, but also can better ensure people's daily lives. However, the prediction data also follows a set of steps and logic. For instance, when predicting the weather, Scher first assesses the feasibility of using machine learning for weather prediction, then initializes a large-scale atmospheric state, and conducts training on past weather forecasts using a neural network deep learning method, thereby improving the prediction accuracy [2]. Because of these reasons, the main objective of this study is to emphasize the logicality of the process of predicting Tesla's stock prices using machine learning. This article argues that when using machine learning to predict data, emphasis should be placed on the steps and logic. Conducting a comprehensive analysis of the data to eliminate special disturbances before making predictions can enhance the efficiency and accuracy of the predictions. Moreover, this paper is devised to three steps to predict Tesla's stock. First, it will utilize the stock price data of Tesla for the past five years and use machine learning to create charts for analyzing the trends and features. Then, based on the previous steps, the machine learning model (LSTM) is used to obtain the prediction results, and a comparative analysis is conducted. Finally, point out the limitations of this model and its future development directions.

2. Methodology

To ensure the prediction of most scenarios, the Tesla stock data in this article is selected from Kaggle, covering the period from January 1, 2020, to January 1, 2025. This article mainly employs the LSTM model. Figures 1, 2, and 3 were all created by the author of this article in the Python 3.0 environment.

2.1. Data Distribution

First, a tree diagram of Tesla's stock was constructed using machine learning (Figure 1). It can be clearly seen from the first four graphs that the stock price of Tesla has a multi-peak trend. And from

the volume distribution at the bottom, a very obvious long-tail feature can be observed. The LSTM model itself is capable of effectively addressing the issue of data sparsity because it has a memory gating mechanism. Ryu's experimental results precisely confirm this [3]. This kind of data with certain complexity and nonlinearity can provide us with suggestions on selecting models to predict stocks. Therefore, prioritizing the use of machine learning to analyze data distribution can improve observation efficiency and reduce inspection costs.

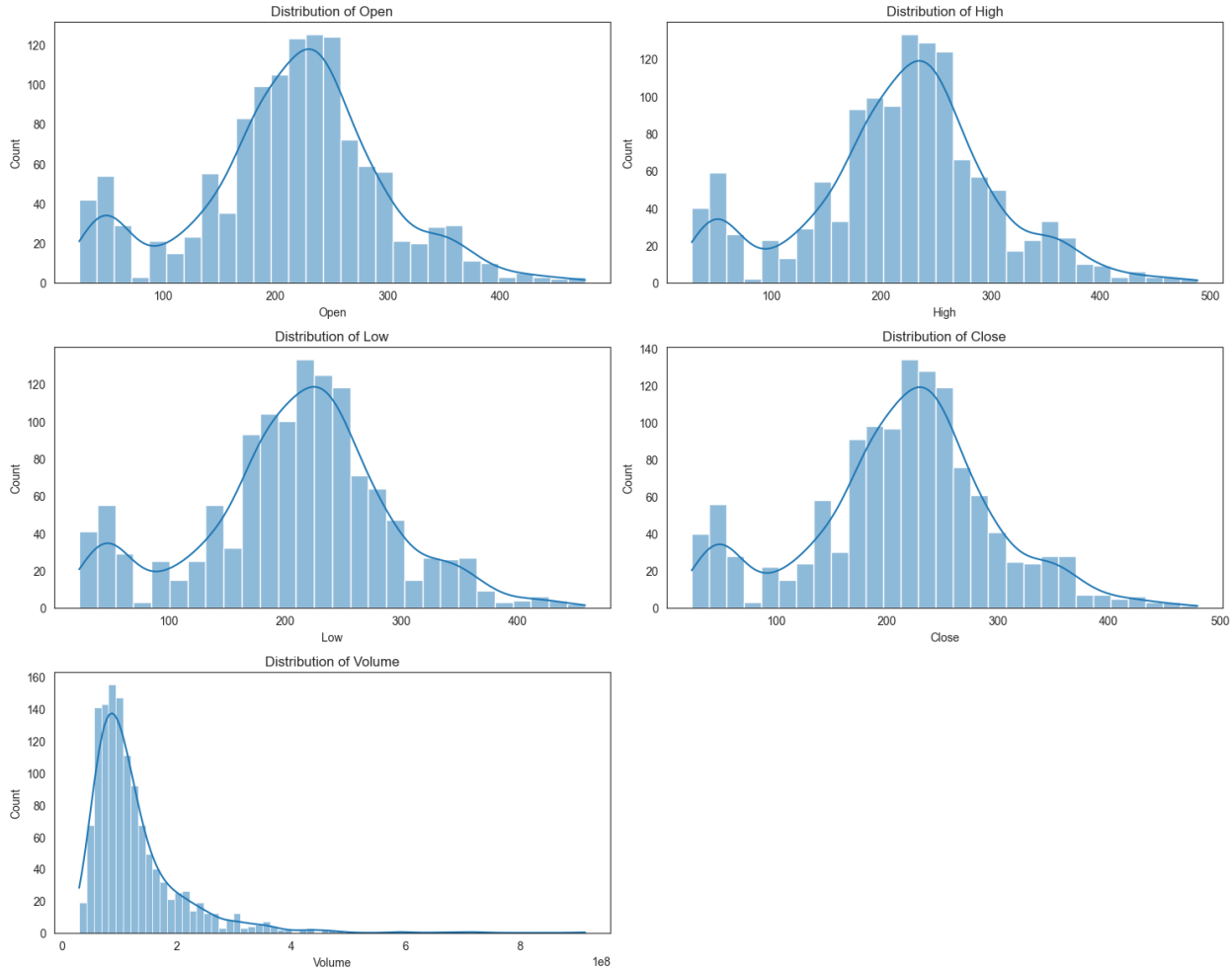

Figure 1. Tree diagram of Tesla's stock

2.2. Data relationship

First, a tree diagram of Tesla's stock was constructed using machine learning (Figure 1). It can be clearly seen from the first four graphs that the stock price of Tesla has a multi-peak trend. And from the volume distribution at the bottom, a very obvious long-tail feature can be observed. This kind of data with certain complexity and nonlinearity can provide us with suggestions on selecting models to predict stocks. Therefore, there is an absolute relationship between trading volume and stock prices. Rui collected data from nine national markets and also proved that there is a positive correlation between the absolute value of trading volume and stock price changes [4]. This also supports that we can use this to improve the accuracy of predicting stock prices.

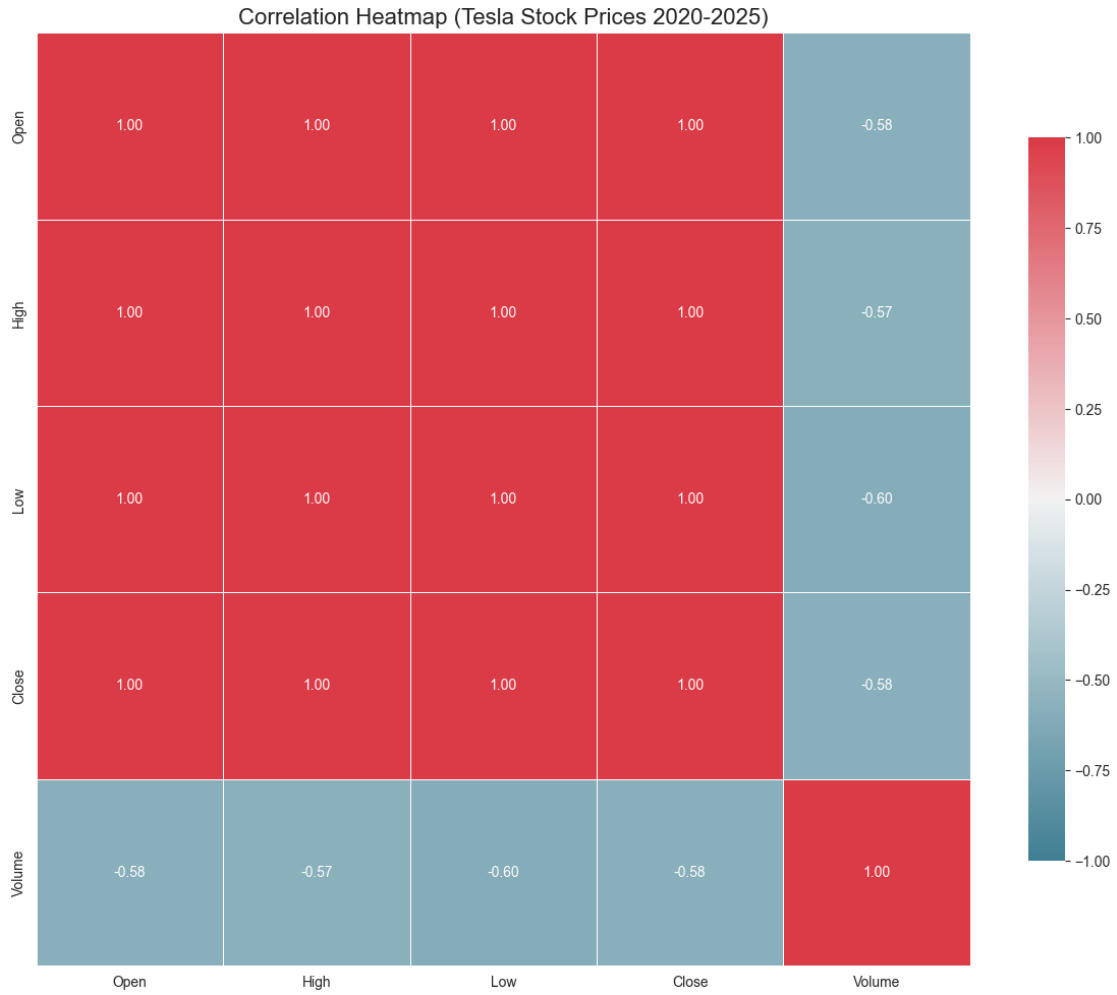

Figure 2. Correlation Heatmap

2.3. The basic principle of LSTM

From the above text, it can be concluded that the most suitable model for predicting Tesla's stock should be LSTM (the specific reasons will be elaborated in the Results section). Therefore, this section briefly introduces the basic principles of LSTM. Long Short-Term Memory (LSTM) is a special type of recurrent neural network, mainly addressing the vanishing or exploding gradient problem in long sequences of data. What makes LSTM unique is its Cell State and Gate Mechanism. The Cell State ensures that LSTM can perform long-term learning with few linear operations. Therefore, compared to other models, LSTM is better at handling long-term data. Another feature of LSTM is the Gate Mechanism, which mainly consists of three gates: the Forget Gate, the Input Gate, and the Output Gate. The Forget Gate determines how much information from the previous cell state should be forgotten, the Input Gate decides how much information should be added to the cell state, and the Output Gate determines what the short-term memory is at the current time step and passes it to the next time step. These three gates and the Cell State allow LSTM to precisely control the inflow, storage, and output of information. These characteristics and advantages make the LSTM model more suitable for stock prediction.

3. Results

After analyzing the data distribution and data relationships, only the LSTM model can accurately predict the stock price of Tesla. Therefore, this paper selects the LSTM model to analyze the stock price of Tesla. Fischer has proposed that LSTM is one of the most accurate and best models for predicting stocks [5]. Zhang has also conducted relevant experiments to prove that although LSTM

performs poorly compared to other models in the short term (1 to 10 days), it is considered one of the most accurate prediction techniques in the long term and over a longer range [6]. The final predicted results (Figure 3) accurately predict the fluctuations in Tesla's stock price. The success of the prediction is inseparable from the previous preparations.

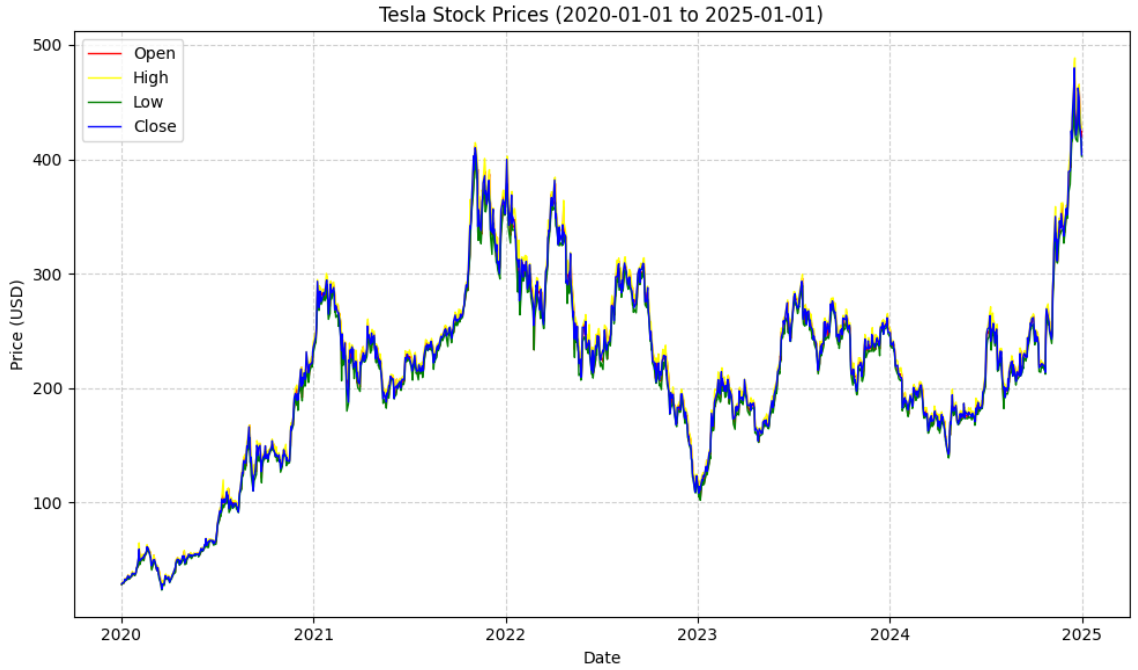

Figure 3. Tesla Stock Prediction (LSTM)

However, even after analyzing the data distribution and relationships and using the model for stock prediction, there are still some shortcomings. For instance, the prediction heavily relies on historical data, and it is unable to simultaneously predict policy changes and market fluctuations. Moreover, sudden wars can also cause stock volatility. This requires us to incorporate multiple factors for stock prediction in future machine learning. Sinkevich proposed that when predicting stocks, news should be considered [7]. Koenecke also suggested that a continuously learning model should be constructed to establish a dynamic model that adapts to market dynamics [8]. These may seem difficult to achieve in the future, but in fact, machine learning has already achieved certain results in this area. Data-intensive machine learning methods can make more well-reasoned decisions in industries such as healthcare, manufacturing, education, and law enforcement [9]. Currently, some relevant personnel have conducted ablation studies on the LSTM-DNN model which provides considerable assistance for accurate stock prediction in the future [10]. With these ideas in mind, we should better analyze the data and find the data relationships and utilize more sophisticated prediction models in the future to make more accurate stock predictions.

4. Conclusion

In conclusion, considering the current development of machine learning, the steps of analyzing data distribution and data correlation before making predictions are indispensable. They can help predictors improve prediction efficiency and reduce testing costs. Of course, the step of selecting an appropriate data analysis model is also essential. For predicting stocks, LSTM is indeed one of the best prediction models at present. In future development, to achieve more accurate prediction results, the continuous development of machine learning and the ability of models to take into account all factors are also indispensable.

References

[1] Chen W, Hussain W, Castrucci F, & Zhang X. Deep learning for financial time series prediction: A state-of-the-art review of standalone and hybrid models. CMES-Computer Modeling in Engineering and Sciences. 2023.

[2] Scher S, & Messori G. Predicting weather forecast uncertainty with machine learning. Quarterly Journal of the Royal Meteorological Society, 144 (717), 2830 - 2841. 2018.

[3] Ryu G A, Chuluunsaikhan T, Nasridinov A, Rah H, & Yoo K H. Sce-lstm: Sparse critical event-driven lstm model with selective memorization for agricultural time-series prediction. Agriculture, 13 (11), 2044. 2023.

[4] Chen G M, Firth M, & Rui O M. The dynamic relation between stock returns, trading volume, and volatility. Financial Review, 36 (3), 153 - 174. 2001.

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

[6] Zhang R. LSTM-based stock prediction modeling and analysis. In: 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022). Atlantis Press, pp. 2537 - 2542. 2022.

[7] Sinkevich Y, McGinnity T M, Coleman S and Beatrice A. Predicting stock price movements based on different categories of news articles. In: 2015 IEEE Symposium Series on Computational Intelligence. Piscataway, NJ: IEEE, pp. 703 - 710. 2015.

[8] Koenecke A and Gajewar A. Curriculum learning in deep neural networks for financial forecasting. In: Workshop on Mining Data for Financial Applications. Cham: Springer International Publishing, pp. 16 - 31. 2019.

[9] Jordan M I, & Mitchell T M. Machine learning: Trends, perspectives, and prospects. Science, 349 (6245), 255 - 260. 2015.

[10] Alam K, Bhuiyan M H, Haque I U, Monir M F and Ahmed T. Enhancing stock market prediction: A robust LSTM-DNN model analysis on 26 real-life datasets, IEEE Access, 12, pp. 122757 - 122768. 2024.