# Optimization of Stock Prediction Methods based on Long Short-Term Memory Model

Zeyan Liu *

SUZHOU FOREIGN LANGUAGE SCHOOL, Suzhou, Jiangsu, China

* Corresponding Author Email: jennifer20240601@gmail.com

**Abstract.** These days, with the fast-paced development of the economy, a growing number of individuals have started making stock market investments, aiming to earn more money with wise decisions. However, a high level of risk has become inevitable because of the high level of stock price volatility, which is impacted by several variables. As a result, forecasting stocks has emerged as a critical component of financial research. One of the fundamental stock prediction models, Long Short-Term Memory (LSTM), is introduced in this article. With the consideration of the increasing level of requirement of the stock prediction, this dissertation also discusses the different ways of optimization of the LSTM models from hybridization, genetic algorithm, attention mechanism to the variant LSTM. Although the optimized models have been experimented with to show better performance in stock prediction, there are still several limitations waiting to be improved in the future. This dissertation has shed light on the analysis and comparison of stock prediction models based on the LSTM, with paramount significance to the future development of stock prediction.

**Keywords:** Stock Prediction; Long Short-Term Memory; Optimization; Hybridization.

## 1. Introduction

Attracted by the high possibility of earning more from the stock investment, more and more residents begin to invest in the stock market. Therefore, the stock market has been regarded as an important part of the financial field for a long period of time, acting as a place where stocks are transferred, traded, and circulated by people with the goal of making more profit [1,2]. However, the stock market is always labelled as high uncertainty, high volatility, and non-stationary with nonlinear stock prices [3,4,5]. As a result, investors are always faced with high risks due to the high level of fluctuations of stock price brought by several factors affecting the stock market, both internally and externally, including the global economic environment, industrial outlook, government policies, investors' moods and so on [3]. Due to the rising investment risk level, it is of paramount significance to make a reasonable and wise prediction of the future stock price in order to reduce the investment risks of investors. Only with more reliable stock predictions, are investors allowed to generate better investment strategies and achieve a higher level of returns from their investments [2,6].

For the predicting technique, in the past, simple linear mathematical models were widely used by scholars for stock prediction [2]. However, due to the increasing amount of noise, high level of uncertainty of the stock information and the requirement of longer stock prediction periods, the traditional linear methods were gradually becoming out-of-date, clearly regarded as less useful with many limitations nowadays [2].

These days, the use of artificial intelligence (AI) has caught much attention in various fields [7]. Undoubtedly, AI technology is also becoming popular in the financial field, stock prediction included. An essential component of artificial intelligence technology is machine learning. And deep learning is logically a machine learning concept [8]. This more sophisticated branch works especially well for prediction projects involving stochastic processes [6]. One benefit of deep learning is its ability to tackle challenging tasks like picture identification and natural language processing [6]. Many different kinds of artificial neural networks exist, including Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), and others. Among these neural networks, Long Short-Term Memory (LSTM), a prominent deep learning research technique, has made great strides in stock price prediction and has emerged as one of the cutting-edge techniques due to its capability to comprehend and analyze data with long-term dependencies, overcoming the problem of gradient exploding, gradient vanishing and insufficient long-term memory of its original model-Recurrent Neural Networks models [2,4,8,9,10].

Although the Long Short-Term Memory models have already been a better choice for stock prediction compared to the RNN models. As the expected performance of the stock prediction rises, the LSTM model seems to be out-of-date with limitations, which include the lack of stability, accuracy, robustness and so on. As a result, the optimization of the LSTM-based stock prediction models has become a hot direction for researchers.

In this article, the basics of Long Short-Term Memory models are introduced. With the rising hope of better stock prediction results, the optimization research of LSTM models, including hybridization of different models, cooperation with attention mechanisms and improved LSTM variants with the help of an algorithm, would also be analysed. Some limitations and improvements of these studies are also presented in this article.

## 2. Long Short-Term Memory

In 1997, Long Short-Term Memory was proposed by Hochreiter and Schmidhuber as a kind of artificial neural network [3,4]. This neural network, which can process sequential input like time series, audio, and so forth, was primarily created using Recurrent Neural Networks model and addresses the long-standing issues of gradient vanishing and gradient exploding [1,2,8,10]. The LSTM model is especially capable of dealing with long-term dependent data [1]. The breakthrough of Long Short-Term Memory basically lies in the gate mechanism presented in each of its units [4]. For the gate mechanism, there are generally three main gates, which are the forget gate, the output gate, and the input gate [4]. The functions of various gates in the gate mechanism vary. The input gate is responsible for selecting the information which will be altered [4,10]. Besides, the output gate decides the data that will be transmitted to the following unit, and the forget gate determines the data that should be removed [4,10].

## 3. Optimization of LSTM for Stock Prediction

With the increasing level of stock market fluctuations and higher requirements of stock predictions, the single Long Short-Term Memory model seems to lack robustness, accuracy and stability. As a result, researchers become more focused on the optimization of the LSTM-based stock prediction models. In 2020, Wenjie Lu, Jiazheng Li, Jingyang Wang and Lele Qin had proposed a hybridised stock prediction model based on the LSTM model, which is a CNN-BiLSTN-AM method. This hybridized model takes advantage of each of the single models composed in the hybridization. Convolutional Neural Networks model, with a high level of capability in picture and natural language processing, is responsible for extracting the input data characteristics [2]. While the interdependence of the stock time series data is adequately discovered by Bidirectional Long Short-Term Memory (BiLSTM) which is a model with two LSTM layers acting in both forward and backwards directions [2]. Moreover, the Attention Mechanism (AM), which draws inspiration from the concept of human visual attention, helps to capture how historical characteristic states of time series data affect stock price [2]. This hybridized model was experimented with based on the Shanghai Composite Index (000,001) stock, which has been proven to have a higher level of accuracy, robustness and better trend adaptation compared to a single LSTM model. However, this research only did next-day closing price predictions and was experimented with based on solely one stock, so it's hard to prove the generalization ability of this model over longer periods of time. Aiming to provide investors with better stock prediction results, a genetic algorithm optimized hybridized model based on the LSTM model was proposed by Heon Baek in 2023, which is a genetic algorithm optimization-based CNN-LSTM stock prediction model [3]. Genetic algorithm (GA) in the improved model could help to discover the optimal solution to the problem of large search spaces, contributing to the model training stage [3]. This GA assisted model has also been proven to have higher accuracy with better performance in stock forecasting. However, this optimized model is also only short-term tested, which means the model lacks the evidence to prove the middle or long-term prediction generalisation ability. Moreover, solely the Korea Stock Index data was chosen for the model evaluation, showing a low level of data diversity. It has also been noted by other researchers that the combination of optimization algorithms may result in increased computing complexity and longer training times [11]. More training time may further lead to a higher risk level for investors as stock price fluctuations often happen in a short time. In the same year, Burak Gülmez proposed another optimized model, which is a deep LSTM network assisted by an artificial rabbits optimization algorithm (ARO). Artificial rabbits optimization algorithm is a novel algorithm, mimicking rabbits' foraging and hiding, which is used for global exploration and local exploitation respectively [1]. The ARO is helpful in offering optimal LSTM hyperparameters [1]. This model seems to be more accurate with generalization ability [1]. On the other hand, rabbit mimicking has no real relationship with the financial field, so it may lead to some bio-inspired limitations due to the disciplinary relevance. In 2024, Shuai Sang and Lu Li focused on enhancing LSTM in order to produce a variant LSTM model that could be used in the BiLSTM model [4]. In their improvement research, one more forget gate was designed internally for the long-term memory in order to selectively forget long-term information, leading to the reduction of noisy data no longer useful [4]. Their experiments were put forward based on stock index data and futures data [4]. The experimental results gave support to the optimized model. The model was proven to possess higher accuracy, robustness, generalisation capability and faster running speed [4]. However, similarly, non-technical factors were not considered and the model was only experimented on short term. A new version of the LSTM model with an attention mechanism was proposed by Shuai Sang and Lu Li in the same year [8]. Their investigation revealed that traditional LSTM had inadequate noise-handling abilities and was comparatively sensitive [8]. To tackle these limitations, for the variational LSTM, the forget and input gates are designed to be coupled, enabling the model to consider both new and past information simultaneously [8]. This helps to reduce the processing pressure for the input gate and reduce the parameters at the same time [8]. Additionally, the enhanced model includes a simpler forget gate for the long-term cell state that could selectively forget long-term memory to increase the noise handling capacity [8]. This improved model had been tested to have a higher level of robustness, accuracy, generalisation capability and faster convergence speed. However, this model was also only short-term experimented with limited generalizability.

## 4. Limitations and Expectations

All in all, the optimizations of the LSTM-based stock prediction models discussed above have shown different kinds of improvements from distinct aspects. The optimized models mentioned in this dissertation have already been assessed from the standards of accuracy, robustness, generalization capability and running speeds. They have been proven to process better performance than the single LSTM model. On the other hand, most of the optimized stock prediction models still have several limitations. Some of them are only experimented with limited experimental data in the short term, lacking a high generalization ability for middle or long-term predictions, hardly satisfying the lengthening of expected prediction periods. In addition, most of the stock prediction models are all faced with the problem of black-box risks, which means it's challenging to explain the logic behind to the investors. This may lead to a lack of confidence among the investors in using the predicted results brought by these models. Moreover, these models are mainly designed based on technical data instead of some economic indicators, investment sentiment, news and so on. This limitation might lead to a lower level of accuracy in the stock price predictions in the future because there are many other elements contributing to the fluctuations of the price in practical. And financial news that occurs suddenly may influence the stock price in a short period, which means it’s of great importance for future researchers to consider more about these non-technical factors when designing the models.

As a result, in order to improve the model's nature and offer investors better aid in making wiser investment decisions, it may be a good idea to combine hybridized models, attention mechanisms, variant LSTM and optimized genetic algorithms together. Also, if more economic factors, such as the investor sentiment or some industrial news, could be considered in the predictions, the stock forecast may be more accurate, enabling investors to make more reliable investment strategies. Moreover, it's essential to enhance the explainability of the prediction because a more logical explanation of the predicted results would be more reliable for the investors. And it can help increase investors' confidence level.

## 5. Conclusion

To conclude, in the area of stock price prediction, Long Short-Term Memory models have been regarded as a significant outbreak in the past. The basic information of the single Long Short-Term Memory is introduced above. Aiming to assist investors in making better investing decisions and generating wiser strategies, several studies based on the optimization of the Long Short-Term Memory model through variant LSTM, algorithms, hybridization of models and attention mechanisms have been discussed and analyzed. Although most of the optimized models have been proven to have a better performance compared with the single LSTM model in stock prediction, they still have several limitations. Firstly, the problems of black-box risks and the lack of consideration of economic factors are still waiting to be improved, which could be regarded as a future expectation. It is also of great significance for future researchers to pay more attention to the data diversity and long-term reliability when designing the stock prediction models and making the evaluations in experiments.

## References

[1] Gülmez, B. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. Expert Systems with Applications, 2023, 227, 120346.

[2] Lu, W., Li, J., Wang, J., & Qin, L. A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications, 2021, 33(10), 4741-4753.

[3] Baek, H. A CNN-LSTM stock prediction model based on genetic algorithm optimization. Asia-Pacific Financial Markets, 2024, 31(2), 205-220.

[4] Sang, S., & Li, L. A stock prediction method based on heterogeneous bidirectional LSTM. Applied Sciences, 2024, 14(20), 9158.

[5] Bathla, G., Rani, R., & Aggarwal, H. Stocks of year 2020: prediction of high variations in stock prices using LSTM. Multimedia Tools and Applications, 2023, 82(7), 9727-9743.

[6] Ye, P., Zhang, H., & Zhou, X. CNN-CBAM-LSTM: Enhancing Stock Return Prediction Through Long and Short Information Mining in Stock Prediction. Mathematics (2227-7390), 2024, 12(23).

[7] Chung, H., & Shin, K. S. Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. Neural Computing and Applications, 2020, 32(12), 7897-7914.

[8] Sang, S., & Li, L. A novel variant of LSTM stock prediction method incorporating attention mechanism. Mathematics, 2024, 12(7), 945.

[9] Bacco, L., Petrosino, L., Arganese, D., Vollero, L., Papi, M., & Merone, M. Investigating stock prediction using LSTM networks and sentiment analysis of tweets under high uncertainty: A case study of North American and European banks. IEEE Access, 2024.

[10] Landi, F., Baraldi, L., Cornia, M., & Cucchiara, R. Working memory connections for LSTM. Neural Networks, 2021, 144, 334-341.

[11] Zeng, X., Liang, C., Yang, Q., Wang, F., & Cai, J. Enhancing stock index prediction: A hybrid LSTM-PSO model for improved forecasting accuracy. PLoS One, 2025, 20(1), e0310296.