# Application of LSTM neural network in the research of cyclical industries: taking the automotive industry as an illustrative case

Ziheng Wang [1, #], Heng Quan [2, *, #], Yu Gao [3, #]

[1] School of Energy and Environment, Southeast University, Nanjing, China, 211102

[2] School of Management & Engineering, Nanjing University, Nanjing, China, 210023

[3] School of Mathematics, Southwestern University of Finance and Economics, Chengdu, Sichuan, 611130

* Corresponding author: 211275004@smail.nju.edu.cn

[#]These authors contributed equally.

**Abstract.** Anticipating future developmental trends within an industry stands as one of the paramount responsibilities for industry researchers. Nevertheless, owing to the intricate and subjective nature inherent in industry research, coupled with researchers usually having to rely on their own subjective judgments when making predictions, the forecasting outcomes frequently prove unsatisfactory. In order to address this issue, this study introduces deep learning technology into the industry research field and employs Long & Short-Term Memory (LSTM) networks to forecast a typical cyclical industry - the automotive industry. In the specific experiment, Adam optimization algorithm is employed to enhance model training speed, followed by utilization of the grid search algorithm for optimal parameter selection. Subsequently, the obtained optimal model is utilized for LSTM multivariate sequence prediction in order to forecast the future development trend of the automotive industry over the next six months. The ultimate outcome demonstrates that the LSTM model accurately anticipated the forthcoming U-shaped rebound trend in the automotive industry, albeit with some limitations in predicting precise numerical values and timing. Nevertheless, it aligns closely with the actual developmental trajectory. This investigation represents a constructive application of deep learning technology in industrial research, and its final findings hold significant implications for future industry research endeavors.

**Keywords:** LSTM, industry research, automotive industry, time series prediction, Adam optimization.

## 1. Introduction

Since the establishment of China's Science and Technology Innovation Board (STAR Market) in 2019 and the implementation of comprehensive registration system reform in 2023, there has been a substantial expansion in the scale of China's A-share market, with the total number of companies exceeding 5,000. For investors, identifying the appropriate industry and company for investment from a vast array of options is paramount and warrants meticulous consideration. A commonly employed investment strategy involves allocating funds to cyclical industries in accordance with the macroeconomic cycle and the industry's own lifecycle, including but not limited to automotive, real estate, machinery, and consumer goods sectors. This approach requires industry researchers to accurately assess the current stage of the industry within its cycle and anticipate future developmental trends in order to produce high-quality, informative industry research reports for investors as a reference.

However, due to the intricate and subjective nature of industry research, industry researchers are typically limited to offering ambiguous future projections based on their own assessment of the industry and economic cycles, rather than accurately forecasting the future developmental trends of the industry and providing investment guidance. Furthermore, as a result of the prevalent tendency for researchers to exhibit excessive optimism, their assessments of the industry's future may demonstrate considerable deviations and inaccuracies. The challenge of making reasonable predictions and conducting thorough research on cyclical industries is a topic worthy of investigation.

Fortunately, the advancement of computational technology and deep learning has presented novel approaches to address this issue. Researchers are continuously integrating machine learning techniques across diverse financial domains with the aim of generating precise forecasts for future trends [1]. This study aims to incorporate LSTM neural network into the industry research of cyclical industries. Using a typical cyclical industry – the automotive industry as a classic example, it will conduct a comprehensive analysis of the economic cycle and current industry conditions, and provide an objective and rational research report along with future predictions.

## 2. Literature review

In the current field of financial data time series forecasting, there are two mainstream research methodologies. The first method is traditional time series analysis models such as Autoregressive Conditional Heteroskedasticity (ARCH), Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH); the alternative approach involves the increasingly prevalent machine learning and deep learning methodologies, such as Support Vector Machine (SVM) and Recurrent Neural Network (RNN), which have gained significant traction in recent years.

Traditional time series analysis methods have inherent limitations. The ARCH model is capable of handling heteroscedastic time series data, but it does not account for the autoregressive impact on volatility. The ARIMA model is adept at effectively managing data with autoregression, but the differencing of original data in this model leads to neglect of long-term trends and undermines the economic significance of the data [2]. The GARCH model, which is an extension of the ARCH model and includes a moving average component to address autoregression effects, simultaneously considers both autoregressive effects and heteroscedasticity. Despite its theoretical comprehensiveness and completeness, practical implementation is hindered by challenges in order and parameter estimation, leading to increased complexity.

In recent years, driven by the growing demand for precise prediction and the increasing popularity of machine learning techniques, there has been a widespread application of Support Vector Machines (SVM) and Recurrent Neural Networks (RNN) in the financial sector for the classification and regression of time series data [3]. While SVM and RNN demonstrate superior modeling effects in comparison to traditional time series models, it is noteworthy that as a shallow machine learning algorithm, SVM exhibits suboptimal performance in data prediction, and although RNN performs well in time series analysis, it is also prone to issues related to vanishing gradients and exploding gradients.

Therefore, this paper employs Long & Short Term Memory (LSTM) networks as the experimental model. The LSTM was first proposed by Hochreiter and Schmidhuber in 1997, optimized from the basis of RNN neural networks [4]. It effectively addresses the issues of long-term memory decay and gradient instability encountered in RNN, rendering it a widely adopted technique in the financial domain in recent years. Moghar and Hamiche used LSTM models to predict stock market trends, resulting in favorable investment returns [5]. Mehtab and Dutta also utilized LSTM to simulate and forecast the NIFTY 50 index in the Indian stock market [6]. Due to the outstanding performance of the LSTM neural network and the research objectives of this paper, this study will conduct time series data analysis using LSTM neural network in a typical cyclical industry - the automotive industry. The goal is to accurately assess the position of cyclical industries within the economic cycle and make objective and rational predictions about future returns. It will also explore the application of deep learning technology in the field of industry research.

# 3. Methodology

## 3.1. Model

The Long Short-Term Memory (LSTM), a specialized recurrent neural network structure, is extensively employed in deep learning for tasks necessitating the handling of long-term dependencies, such as time series prediction [7].

The LSTM architecture comprises three crucial gating units, namely the input gate, forget gate and output gate as shown in Figure1. Each of these gates consists of one or more neurons that are connected in a weighted manner. These gated units learn to regulate the flow of information, thereby enabling LSTM networks to handle long-term dependencies more effectively [8].

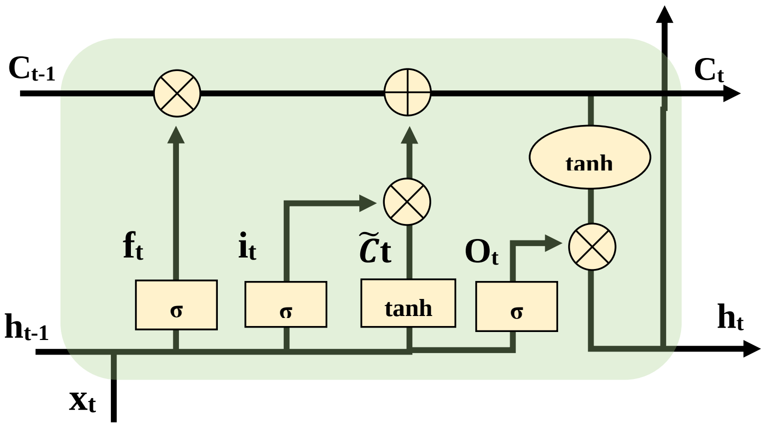


**Figure 1.** The repeating module in LSTM

In Figure1, the expressions of each gate and memory cell in the model are as follows.

The calculation expression of the input gate is:

$$i_t = sigmoid(W_i[h_{t-1}, x_t] + b_i) \tag{1}$$

The calculation expression of the forget gate is:

$$\tilde{C} = tanh(W_C[h_{t-1}, x_t] + b_C) \tag{2}$$

The expression for calculating the new state is:

$$\begin{gathered} f_t = sigmoid(W_f[h_{t-1}, x_t] + b_f \\ C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \end{gathered} \tag{3}$$

The calculation expression of the input gate is:

$$\begin{gathered} o_t = sigmoid(W_o[h_{t-1}, x_t] + b_o) \\ h_t = o_t * tanh(C_t) \end{gathered} \tag{4}$$

Where Wi, bi is related to the learnable parameters, it is related to the new vector, and hi is related to the previous state.

The Adam algorithm, an adaptive learning rate optimization technique for neural network training, combines the benefits of momentum gradient descent and adaptive learning rate methods. It is widely applicable and exhibits efficient performance.

The algorithm is based on the concept of the Gradient Descent algorithm (SGD), combined with Adagrad and RMSProp algorithms. The computation relies on the first derivative of the objective function, ensuring a relatively low computational load. Adam supports sparse gradients, where the parameter update size remains unchanged regardless of gradient scaling. The step size boundary for parameter updates is limited by hyperparameter settings, eliminating the need for a fixed objective function.

Representing the objective function as f(θ), random optimization typically aims to minimize E(f(θ)). However, since we are working with small batches each time, we actually deal with

$f_1(\theta)\ldots f_T(\theta)$ denoting gradients corresponding to steps t as $g_t=\nabla_\theta f_t(\theta)$. The Adam method estimates both first and second moments of gradient E ($g_t$).

### 3.2. Variables

Dependent variable. The article selects the automotive industry as the object of industry research. As a typical cyclical industry, the automotive industry exhibits robust cyclical characteristics and possesses the advantageous attributes of substantial industry scale, consistent market performance, and transparent public data, rendering it a fitting subject for research. Consequently, this article utilizes the index value of the A-share market's automobile vehicle selected index (wind code 8841392.WI) as the dependent variable for data fitting and future prediction.

Independent variables. In the realm of industrial research, the PEST model and Porter's Five Forces analysis framework are widely recognized as prominent industry research models. As a method for macro factor analysis, PEST highlights the necessity of conducting a comprehensive analysis encompassing the political (P), social (S), technological (T), and economic (E) aspects. Simultaneously, Porter's Five Forces model underscores the significance of considering internal industry competition, as well as the influence of upstream suppliers and downstream buyers. The independent variable for the model will be determined based on the analysis of these two models. All the independent variables are presented in Table 1.

**Table 1.** Independent variables

| Factor | Description |
|---|---|
| One-year Loan Prime Rate (LPR) | The benchmark for the loan market |
| Manufacturing Purchasing Managers' Index (PMI) | A barometer for the economic health of the manufacturing sector |
| Macro-economic prosperity index | An indicator of future economic prospects |
| Industrial value added of enterprises above designated size | A reflection of the prosperity of industrial development |
| Consumer Price Index (CPI) | The predominant measure of inflation |
| Narrow money supply (M1) | The aggregate of cash and demand deposits |
| Automobile sales volume | The total sales revenue per month |
| Threaded steel (Φ16) prices | The predominant raw materials utilized in automobile manufacturing |
| Aluminum alloy (A356) prices | |
| Concentration of the automobile industry | The ratio of the cumulative market value of the leading five companies |

### 3.3. Sample

This study collected monthly data of the aforementioned dependent and independent variables from January 2014 to December 2023, spanning a decade, sourced from Wind Financial Terminal and the official website of the National Bureau of Statistics of China. The dataset comprises 1320 indicator data points spanning 120 months. These data will be used to create input vectors for the LSTM model in order to make predictions on multivariate time series. The primary focus of this study will be on the dependent variable, specifically the selected car index value. Furthermore, this study also gathered variable data from January to April 2024, which will be utilized for comparative analysis with the projected values.

Fortunately, due to the exceptional performance of LSTM networks in managing non-stationary sequential data, there is no need to employ unit root tests, ADF tests, and other methods to assess the stationarity of the data. Prior to entering the model, all data have been preprocessed using the Z-score Normalization method. The standardized data value = (original data value – mean value) / standard deviation. Figure 2 below illustrates the time series of the dependent variable "automobile vehicle selected index," with pre-2024 data (in blue) utilized for model training and post-2024 data (in red) employed for comparison against predicted outcomes.

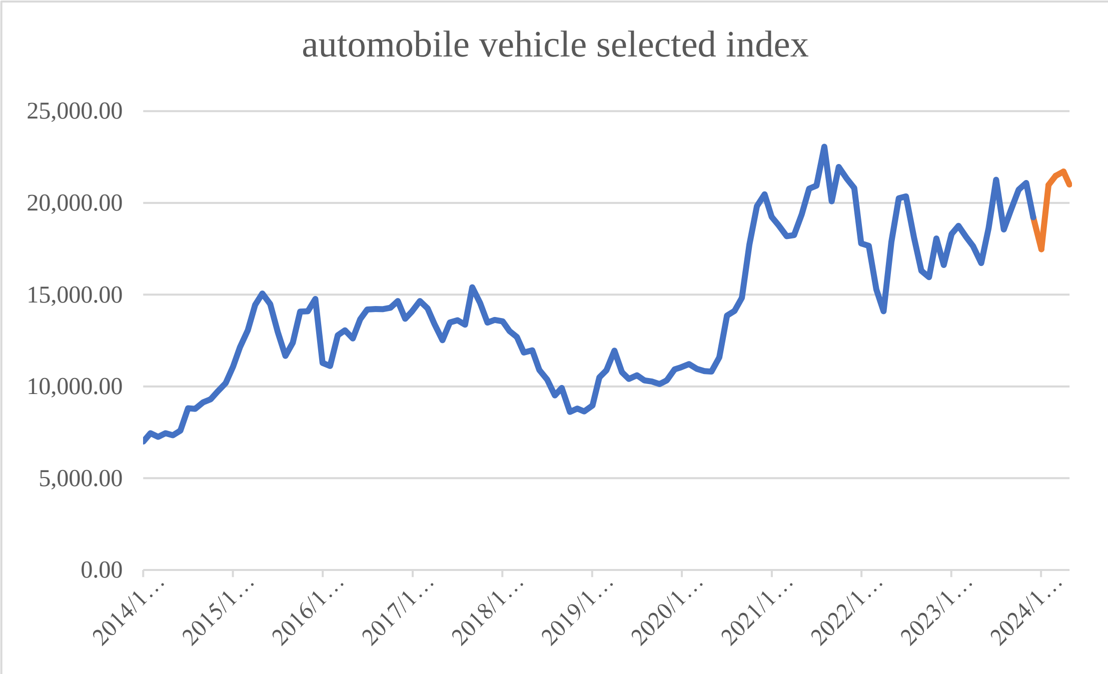


**Figure 2.** Automobile vehicle selected index

## 4. Analysis and discussion

### 4.1. Model training

Due to the intricate nature of LSTM in managing multivariate time series and the extensive array of variables in this experimental design, this study will employ the Adam optimization algorithm for model gradient descent. The Adam algorithm is capable of dynamically adjusting the learning rate without manual intervention or calibration, thereby facilitating rapid convergence of the model and enhancing training efficiency [9]. The experimental findings suggest that all training models exhibit convergence after 500 iterations. This study specifies the number of epochs as 500 rounds, the number of epochs for reducing the learning rate as 250, the Factor for decreasing the learning rate as 0.2, the Decay rate of gradient moving average (Beta1) as 0.9, and the Decay rate of squared gradient moving average (Beta2) as 0.999. These parameters represent the settings utilized for implementing the Adam optimization method in this research.

This study partitioned the initial 100 data points into a training set and utilized the remaining 20 data points as a test set. In order to mitigate the potential for significant overfitting in the model, this study adopts a forget rate of 0.2 for the forget gate based on insights from two scholarly articles employing LSTM neural networks for stock prediction [5] and [6]. To determine the optimal initial learning rate and number of hidden units for time series prediction, the grid search algorithm is employed to systematically explore all possible combinations derived from the Cartesian product of these two sets. The algorithm will employ the root mean square error (RMSE) as the evaluation metric, and the set of indicators with the smallest RMSE in the test set will be utilized as the initial parameters for the predictive component of the model. Table 2 presents the results of the grid search. The horizontal axis of the table represents the number of hidden units, and the vertical axis represents the learning rate.

**Table 2.** Grid search

| | | | | | | |
|---|---|---|---|---|---|---|
| 0.0001 | 5237 | 5297 | 5714 | 4866 | 5735 | 6193 |
| 0.0005 | 5115 | 6296 | 5223 | 5849 | 6117 | 6761 |
| 0.001 | 6123 | 6681 | 5170 | 4977 | 4784 | 4530 |
| 0.002 | 5701 | 4372 | 3346 | 3053 | 2985 | 4345 |
| 0.005 | 6113 | 3043 | 2486 | 3856 | 4273 | 5950 |
| Parameter | 50 | 100 | 200 | 500 | 800 | 1000 |

Based on the search results, it is evident that setting the number of hidden units to 200 and the learning rate to 0.005 yields the model with the lowest RMSE on the test set, indicating its superior fit. It is important to note that while the absolute value of RMSE is relatively large at 2486, considering that most of the test set data falls within the range of 15000 to 20000, this RMSE value can be deemed acceptable. At this point, all model parameters have been set in this paper and the configured model will be used for time series prediction.

### 4.2. Future prediction

Based on the model trained in Section 4.1, this paper utilized the model to forecast the trajectory of the car selection index for the upcoming six months (i.e., January 2024 to June 2024). The final RMSE of the model when fitting the original data is 277.2, and specific future prediction results are depicted in Figure 3.

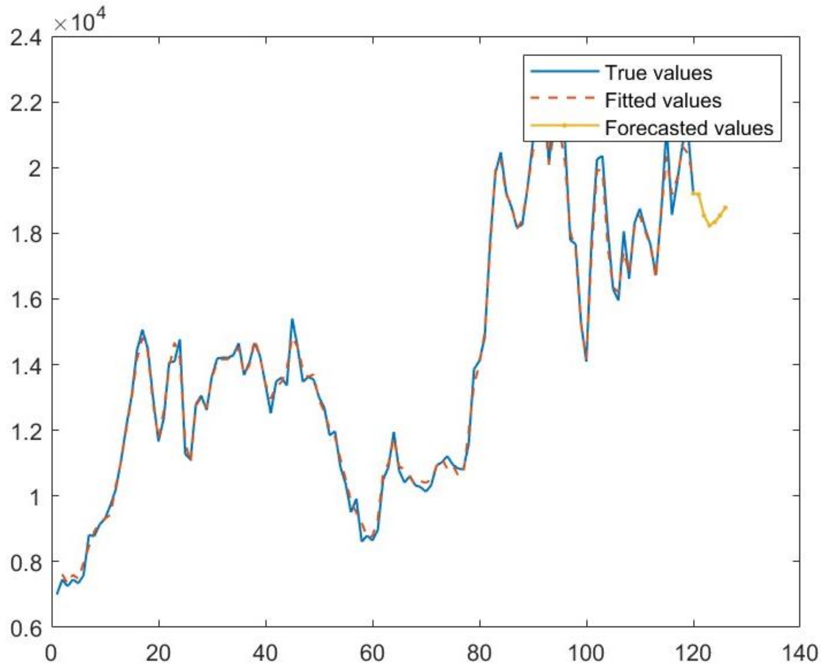


**Figure 3.** Prediction outcome

The predictive findings indicate that the Automobile Select Index experienced a transient trough following a brief additional decrease, followed by a rebound and ascent, demonstrating an overall U-shaped trajectory.

Compared to the observed trend of the stock index in Figure 1 from January to May 2024, it is evident that LSTM accurately anticipated the future trend pattern of the index, namely a U-shaped decline followed by a rebound, however, its performance is suboptimal in predicting specific numerical values and identifying critical time points in this experiment.

### 4.3. Result analysis

The findings of the study illustrate the correctness and applicability of LSTM in forecasting future trends within the industry; however, it exhibits subpar performance in predicting specific numerical values and time points, which can be attributed to various factors.

Factor 1: The peak-heavy tail feature of data. It is widely recognized that financial data display pronounced peak-heavy tail characteristics, deviating from a standard normal distribution [10]. In this study, the Z-score normalization method was employed to initialize the data, despite its suitability being limited to normally distributed data, potentially impacting the accuracy of predictions.

Factor 2: The high volatility feature of market. Different from other major stock markets in the world, the majority of participants in the A-share market are retail investors rather than professional investors. Their irrational behavior often leads to abnormal fluctuations in stock prices, greatly increasing the difficulty of accurate prediction.

Factor 3: The overfitting problem. As the complexity of the neural network increases, overfitting becomes a challenging issue to circumvent. Despite selecting the best-performing model in the test set, this study still grapples with mitigating the impact of overfitting on prediction results.

Despite its poor performance in accurate prediction for time series forecasting, the model is capable of accurately identifying the future development trend of the index. This has significant implications for industry researchers specializing in value investment, as they can integrate the predictive results of the LSTM model with their own research judgments, thereby reinforcing each other and producing more robust industry research reports for investors. In summary, the application of LSTM model prediction can be of significant importance for industry research.

## 5. Conclusion

To address the subjective nature of industry research judgments and to enhance the accuracy of predicting future industry trends, this study incorporates the deep learning network LSTM into the field of industry research. The automotive industry, known for its cyclical characteristics, is chosen as the focal point of investigation. The findings demonstrate that

Due to the complexity and theoretical nature of time series forecasting, and the negative impact of disadvantageous factors, like traditional forecasting methods, LSTM also cannot accurately predict future values. However, LSTM model technology has already been able to achieve correct predictions of future trends in index value, which is of significant importance and practical value for industry research which focuses on the value investment and long-term trend analysis. Additionally, compared with traditional time series prediction method, LSTM model has emancipated itself from the constraints of rigid assumptions, thereby exhibiting enhanced predictive performance and accuracy, as well as heightened practical applicability in real-world scenarios.

In conclusion, this study is a successful exploration and attempt of the application of neural network technology in the industry research field. In the future, it is possible to further optimize the model to improve prediction accuracy, and also to use other deep learning models to extend future predictions to other non-cyclical industries.

## References

[1] Zhong, Y., Luo, L., Wang, X., & Yang, J. (2020). Multi-factor Stock Selection Model Based on Machine Learning. Engineering Letters, 29 (1).

[2] Khan, S., & Alghulaiakh, H. (2020). ARIMA model for accurate time series stocks forecasting. International Journal of Advanced Computer Science and Applications, 11 (7).

[3] Verma, S., Sahu, S. P., & Sahu, T. P. (2023). Stock Market Forecasting with Different Input Indicators using Machine Learning and Deep Learning Techniques: A Review. Engineering Letters, 31 (1).

[4] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735 - 1780.

[5] Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. Procedia computer science, 170, 1168 - 1173.

[6] Mehtab, S., Sen, J., & Dutta, A. (2021). Stock price prediction using machine learning and LSTM-based deep learning models. In Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14 – 17, 2020, Revised Selected Papers 2 (pp. 88-106). Springer Singapore.

[7] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9 (8), 1735 - 1780.

[8] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems, 28 (10), 2222 - 2232.

[9] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980.

[10] Mi Xianhua, Ma Chaoqun, & Zhao Xinwei. (2022). Research on the impact of skewness of stock market returns on risk compensation in China. Chinese Journal of Management Science, 30 (2), 48 - 57.