Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD29925

Stock Price Analysis and Forecasting of Nestlé Using Neural Networks

Xiaozhou Tong

XJTLU AIAC, Taicang, China
2991471815@qq.com

Abstract. Within the broader scope of applications of artificial intelligence to financial forecasting, this paper studies how to deal with nonlinearity and high noise for stock price prediction. We apply BPNN to investigate appropriate configuration settings for long-term stock prediction. Through systematically adjusting hyperparameters like the data span, the hidden-layer structure, or the number of neurons, we show that the most appropriate model for capturing Nestlé stock's long-run trend characteristics is the one trained on 10 years of data, with 16 neurons over two hidden layers. This setting is not only an inherent regularity of a mature company's behavior, but it also has robustness as we have done many comparisons. It makes it part of using neural networks in an attempt to achieve financial forecasts.

Keywords: neural networks, stock price analysis, Nestlé stock prediction

1. Introduction

Artificial Intelligence and Big Data Technologies are rapidly growing and as a result machine learning application in financial market is receiving more attention from both scholars and industry. Neural networks can model complex non-linear relationships, which makes them powerful tools for stock price analysis and prediction. Because stock price forecast has to do with investors, financial institutions and government decision makers, the fusion of financial analytics with neural network techniques has become very popular.

This study is about how we can use neural networks, which are computer systems that act like a brain, to predict stock prices. The study shows the good things about using neural networks over other ways to predict stock prices. Empirical studies show a number of times that the Neural Networks always outperforms the traditional model which gives very far away predictions from the real world market. Whereas traditional stat methods can merely deal with smaller datasets and can't seize on non-linear dynamics, neural networks can process bigger datasets and catch on to non-linear dynamics, so forecast accuracy and reliability get enhanced.

This paper is structured into five sections. Section 2 reviews the main research contributions of the BPNN algorithm in stock price forecasting. Section 3 introduces the relevant theoretical framework and mathematical models. Section 4 conducts an empirical analysis using historical data on Nestlé's stock. Section 5 discusses the numerical results, evaluates the model's predictive performance, and concludes with a summary of the study's contributions.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD29925

2. Literature review

Shi Qianzi [1] proposed an integrated stock price prediction model based on a BP neural network. By analyzing market factors that influence stock prices, the author selected 21 evaluation indicators and applied principal component analysis (PCA) to reduce them to six input variables. The model employed a thinking-evolution algorithm to optimize initial connection weights and thresholds, while exploring critical parameters such as network topology, the number of hidden layer nodes, learning rate, and activation function. Comparative experiments with multiple models demonstrated that this integrated approach achieved high prediction accuracy and strong stability. Zhong Jialin [2] constructed a stock price prediction model that integrates financial data with text-based sentiment analysis using a long short-term memory (LSTM) neural network. Using the Shanghai Composite Index (000001) as the research subject, the study collected nine feature indicators—including closing prices—covering the period 2010–2020, and innovatively incorporated investor comment data from Eastmoney forums. To address the issue of "online paid posters," a support vector machine model was applied for detection based on user and content features. After removing noisy data, sentiment analysis was used to quantify market sentiment and form a sentiment time series. Experimental results showed that the LSTM model combining financial and sentiment features outperformed traditional ARIMA(4,1,2)-GARCH(1,1) models and improved BP neural networks, with the prediction standard deviation reduced from 6.7% to 4.1% compared to an LSTM model using financial data alone. Rao Jing [3] developed a multi-factor stock price forecasting model (PCA-LSTM) that combines PCA with LSTM to address the complexity of stock price prediction. Sixteen influencing factors were selected across three dimensions: basic trading indicators, technical indicators, and company financial indicators. Four constituent stocks of the CSI 300 Index were chosen as research objects. To mitigate the "curse of dimensionality" in multivariable inputs, PCA was used for feature extraction and dimensionality reduction, yielding five optimal principal components. Through systematic optimization of hyperparameters such as the number of LSTM layers and hidden neurons, the model achieved superior predictive accuracy and goodness-of-fit on test datasets, significantly outperforming RNN and BP neural network benchmarks. Results indicated that the PCA-LSTM model effectively captures stock price dynamics and demonstrates both predictive power and generalization ability. Zhu Fu [4] addressed stock price forecasting for listed companies in the chemical industry by proposing a multimodal feature-fusion prediction model based on the Stacking ensemble learning algorithm. Data from 45 key listed chemical companies between January 2020 and June 2022 were analyzed. Three base models were constructed: a company report feature model based on 18 financial and non-financial indicators, a regression model using historical price sequences, and a sentiment analysis model based on news text. For the sentiment analysis component, the study innovatively applied the Bert-Base-Chinese deep learning model to extract sentiment factors, achieving significantly higher accuracy than traditional sentiment dictionaries. Finally, multimodal features were integrated through a CatBoost meta-model to predict stock prices. Experimental results confirmed the ensemble model's superior performance in the chemical sector. Lin Zhihang [5] proposed a hybrid VP-GRU forecasting model that integrates variational mode decomposition (VMD), PCA, and gated recurrent units (GRU). Input features were screened using mutual information analysis, while VMD decomposed critical price sequences to capture multi-scale market information. PCA was then applied to reduce high-dimensional features. A quantitative investment strategy based on this forecasting model was developed and validated through historical back-testing, demonstrating the ability to generate stable excess returns. Zhu Rui [6] introduced a CP-LSTM forecasting model that integrates frequency decomposition with PCA to predict the closing prices of the CSI 300 Index, and further designed a

DOI: 10.54254/2755-2721/2025.LD29925

quantitative investment strategy based on this framework. Using historical data from 2011 to 2022, the study employed LSTM to process time series data, mutual information analysis to select effective features, and frequency decomposition with PCA to optimize input quality. Results indicated that the CP-LSTM model achieved the strongest predictive performance, with the corresponding investment strategy delivering an annualized return of 20.72% in back-tests, alongside stable excess returns and relatively low risk. These findings suggest that the Chinese stock market is not fully efficient and provide investors with an effective forecasting tool.

Zhou Qi [7] built a hybrid neural network model integrating convolutional neural networks (CNN), LSTM, and attention mechanisms, achieving superior performance in forecasting the SSE 50 Index. From the perspective of online public opinion, the study also collected stock forum comment data and established a sentiment analysis model using Word2Vec and BiLSTM, thereby constructing effective investor sentiment indicators. A multi-source information fusion framework was then developed to combine trading data, technical indicators, and sentiment information. Experimental results showed that incorporating online sentiment significantly enhanced predictive accuracy across stocks with varying ownership concentration, with particularly strong improvements for stocks with low ownership concentration, thus offering valuable decision support for stakeholders. Li Cong [8] developed a quantitative trading strategy combining multi-factor stock selection from CSI 500 constituents with LSTM-based prediction. Back-testing demonstrated that the strategy significantly outperformed the CSI 500 Index during 2019–2020 and consistently delivered stable excess returns across diverse market conditions. This not only validated the strategy's effectiveness but also provided empirical evidence for the market's incomplete efficiency. Li Tiantian [9] examined stock price forecasting methods by comparing time-series approaches with neural network models (LSTM and BP). The results showed that neural networks achieved higher accuracy than traditional methods, though all models exhibited increasing prediction errors over longer time horizons. The study recommended adopting more complex hybrid models or incorporating seasonality factors to improve forecasting accuracy. Leng Tianxiang [10] proposed an attention-based LSTM model (Alt-LSTM) for predicting prices of FTSE China A50 constituent stocks. The model generally outperformed benchmarks; however, during extreme market events, simplified versions relying only on basic indicators yielded better results. The research also demonstrated that appropriate hyperparameter selection can significantly improve prediction accuracy.

3. Method

Neural networks refer to machine learning models that are based on the connectivity patterns of biological neurons. They basically use neuron's cooperative calculation layers, which lets them figure out and foresee complicated info. They have also been widely used in fields like handwritten digit recognition and image analysis.

A normal neural network contains an input layer, at least 1 hidden layer, and an output layer. Input layer receives data; hidden layers are responsible for performing nonlinear transforms on the data, and the number of hidden layers and nodes usually follows empirical practices; output layer produces the final results. Every layer consists of neuron, each neuron will have weight of inter connection. The network can now learn complex relationships because non-linearity comes from us us an activation function Sigmoid Function for most cases

The sigmoid function is given by:

$$\mathrm{sigmoid}(x) = \frac{1}{1+e^{-x}} \tag{1}$$

DOI: 10.54254/2755-2721/2025.LD29925

Its derivative is given by:

$$y = \text{sigmoid}(x) \tag{2}$$

$$y' = y(1 - y) \tag{3}$$

Neural network training mainly uses the backpropagation (BP) algorithm. In this process, the network first uses forward propagation to calculate the predicted output, then compares it with the real value, and calculates the error. Errors are passed through the layers backwards and the weights are updated by gradient descent so the error becomes less Gradient checking can ensure the correctness of our implementation while vectorized code makes the model runs more effectively to give us correct predictions as fast as possible. Backpropagation is in fact an application of the differentiation chain rule. Objective function for the Neural Network is defined first followed by minimization of this using Stochastic Gradient Descent and optimizing the Weight parameters. The objective function is usually taken as the sum of squares of the error at the output nodes:

$$E_d = \frac{1}{2}\sum_{i \in imputs}\left(t_i - y_i\right)^2 \tag{4}$$

Where E_d represents the error for sample d.

The stochastic gradient descent algorithm updates the weights according to:

$$w_{ji} \leftarrow w_{ji} - \eta \frac{\partial E_d}{\partial w_{ji}} \tag{5}$$

Where w_{ji} is the weight, and η is the learning rate. Thus, the optimization process requires calculating the partial derivative of the error function E_d with respect to each weight w_{ji} .

4. Experimental results

The experiment employed actual stock data of Nestlé from 2012 to 2022, obtained from the Kaggle platform. The dataset included opening price, highest price, lowest price, and closing price, with a daily time granularity.

A series of comparative experiments were designed. The baseline model used 10 years of data, with 16 neurons in a single hidden layer. Its performance metrics were: MAE:0.012316 MSE: 0.000254

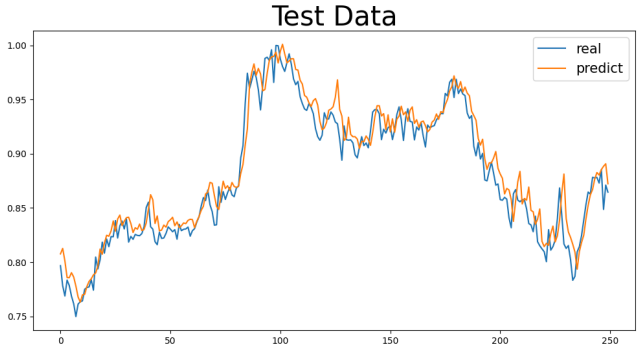


Figure 1. Numerical result of baseline group

DOI: 10.54254/2755-2721/2025.LD29925

Other comparison groups included:
Group 1. 10 years of data, 8 hidden neurons, 1 hidden layer: MAE:0.050365 MSE:0.003033



Figure 2. Numerical result of (10 years, 8 codes, 1 hidden layers) group

Group 1. 10 years of data, 8 hidden neurons, 1 hidden layer: MAE:0.029394 MSE:0.001122

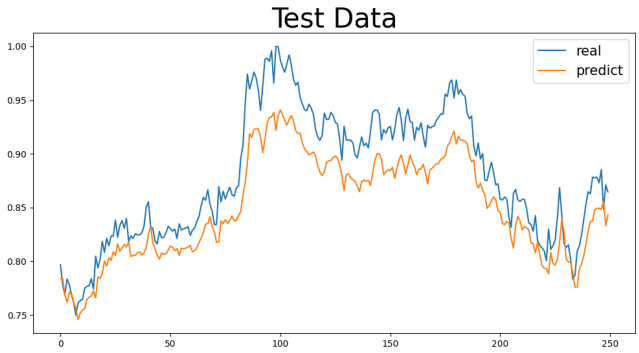


Figure 3. Numerical result of (10 years, 32 codes, 1 hidden layers) group

Group 1. 10 years of data, 8 hidden neurons, 1 hidden layer: MAE:0.031055 MSE:0.001343



Figure 4. Numerical result of (10 years, 8 codes, 2 hidden layers) group

Group 4. 10 years of data, 16 hidden neurons, 2 hidden layers: MAE:0.026498 MSE:0.001004

DOI: 10.54254/2755-2721/2025.LD29925



Figure 5. Numerical result of (10 years, 16codes, 2hidden layers) group

Group 5. 10 years of data, 32 hidden neurons, 2 hidden layers: MAE:0.03000005 MSE:0.001172

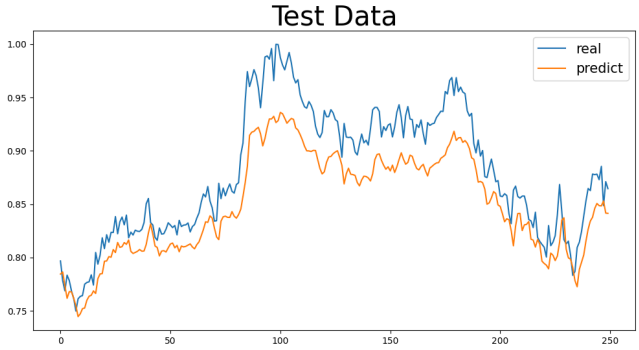


Figure 6. Numerical result of (10 years, 32codes, 2 hidden layers) group

Group 6. 5 years of data, 8 hidden neurons, 1 hidden layer: MAE 0.017821 MSE 0.000492

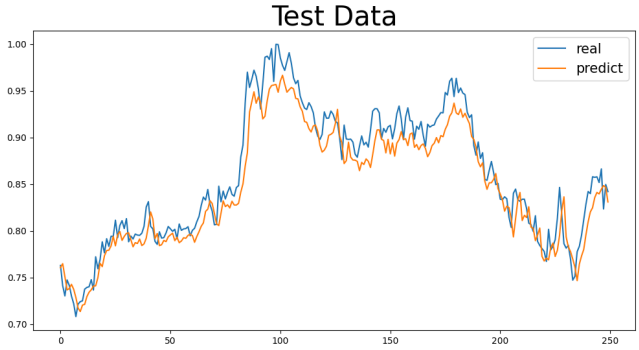


Figure 7. Numerical result of (5 years, 8codes, 1 hidden layers) group

Group 7. 5 years of data, 16 hidden neurons, 1 hidden layer: MAE:0.034532 MSE:0.001639

DOI: 10.54254/2755-2721/2025.LD29925



Figure 8. Numerical result of (5years, 16 codes, 1 hidden layers) group

Group 8. 5 years of data, 32 hidden neurons, 1 hidden layer: MAE:0.047174 MSE:0.003192

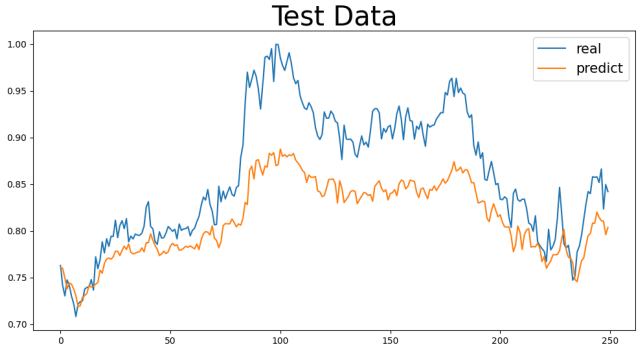


Figure 9. Numerical result of (5years, 32 codes, 1 hidden layers) group

Group 9. 5 years of data, 8 hidden neurons, 2 hidden layers: MAE= 0.076790,MSE= 0.007462



Figure 10. Numerical result of (5years, 8 codes, 2 hidden layers) group

Group 10. 5 years of data, 16 hidden neurons, 2 hidden layers: MAE 0.040479 MSE:0.002373

DOI: 10.54254/2755-2721/2025.LD29925



Figure 11. Numerical result of (5 years, 16 codes, 2 hidden layers) group

Group 11. 5 years of data, 32 hidden neurons, 2 hidden layers: MAE: 0.040875 MSE:0.002493

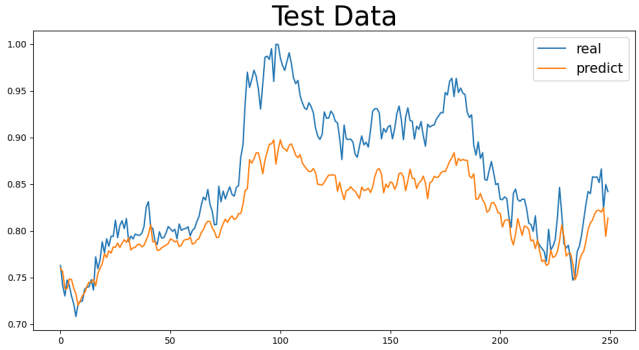


Figure 12. Numerical result of (5 years, 32 codes, 2 hidden layers) group

Group 12. 3 years of data, 8 hidden neurons, 1 hidden layer: MAE:0.052048 MSE:0.003908

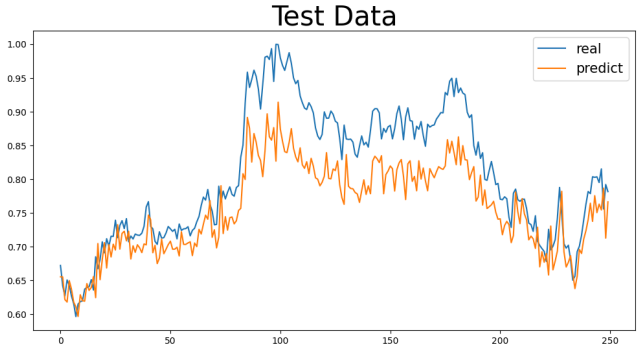


Figure 13. Numerical result of (3 years, 8 codes, 1 hidden layers) group

Group 13. 3 years of data, 16 hidden neurons, 1 hidden layer: MAE:0.056788 MSE:0.004757

DOI: 10.54254/2755-2721/2025.LD29925



Figure 14. Numerical result of (3 years, 16 codes, 1 hidden layers) group

Group 14. 3 years of data, 32 hidden neurons, 1 hidden layer: MAE: 0.046424 MSE: 0.003076

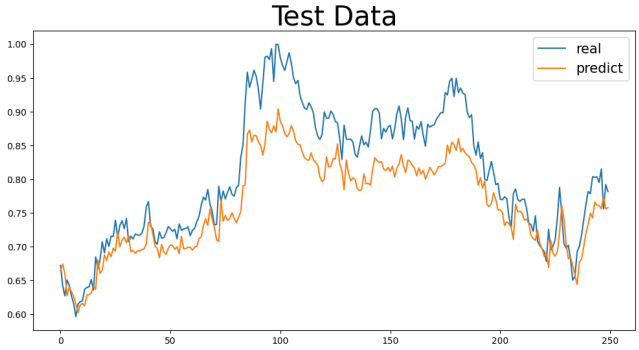


Figure 15. Numerical result of (3 years, 32 codes, 1 hidden layers) group

Group 15. 3 years of data, 8 hidden neurons, 2 hidden layers: MAE:0.046424 MSE: 0.003076

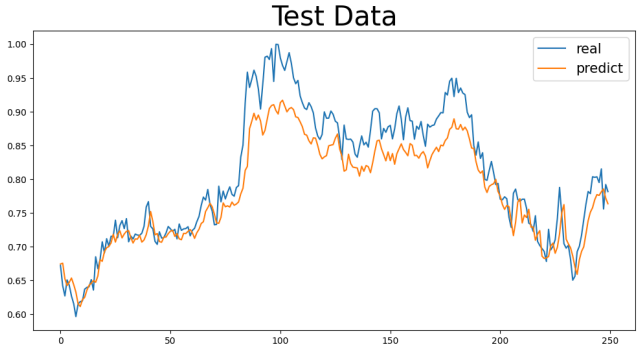


Figure 16. Numerical result of (3 years, 8 codes, 2 hidden layers) group

Group 16. 3 years of data, 16 hidden neurons, 2 hidden layers: MAE: 0.0375384 MSE: 0.002320

DOI: 10.54254/2755-2721/2025.LD29925



Figure 17. Numerical result of (3 years, 16 codes, 2 hidden layers) group

Group 17. 3 years of data, 32 hidden neurons, 2 hidden layers: MAE: 0.043409 MSE: 0.002857

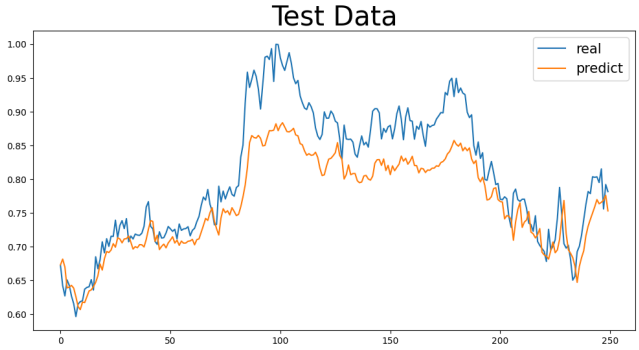


Figure 18. Numerical result of (3 years, 32 codes, 2 hidden layers) group

5. Discussion and conclusion

In the context of the wide use of artificial intelligence in financial forecasts, neural networks were used to predict the share price of Nestlè with the stock prices of Nestlè for the years 2012 to 2022. With regard to constructing a MLP model and training it using the backpropagation algorithm, carrying out regression prediction of closing prices took place. The experimental results reveal that under most scenarios, the constructed neural network models were capable of fitting the stock price movements relatively well, prediction errors were relatively low, and their overall performance was decent. Using a series of experiments for comparisons, this experiment systematically examines the influence of training data length, number of hidden-layer neurons, and the number of hidden layers on model prediction. The best performance was observed on the model which has 10 years of data, 16 neurons, and 2 hidden layers (Group 4) with MAE of 0.026498, MSE 0.001004. Overall, longer data span length (10 years) tended to produce more stable results, as Nestle is a maturing and stable company, it is better to look at its stock over a long term. At the same time, more volatile stocks with shorter data periods (3 or 5 years) might help too. In each setting with regard to the number of neurons, it is observed that a model architecture of size 16 performs well for all values of p, meaning that this model size is good for capturing the data features without overfitting. Concerning network depth, two hidden layers were superior in some cases compared to one, which implies the existence of nonlinear relations in stock price data. But adding more layers increases the computational complexity too, so there is always trade-off for it in actual implementations.

DOI: 10.54254/2755-2721/2025.LD29925

The significance of this research is to demonstrate by examples how successful can a neural network be in predicting stock prices especially when it comes the parameters like length of training, the length of the neural network have the most influence in model performance And it gives some help about how to pick and make better forecast forms during real money buying and selling activities, and gives tests that work and ideas for making AI tools much better at seeing how numbers will go up and down over time.

References

[1] Shi, Q. Z. (n.d.). Application analysis of stock price prediction model based on BP neural network (Doctoral dissertation, Jinan University).
[2] Qian, J. L. (2021). Stock price prediction based on LSTM neural network and text analysis (Master's thesis, Soochow University). https: //doi.org/10.27351/d.cnki.gszhu.2021.003122
[3] Rao, J. (2023). Research on stock price prediction based on multi-factor and PCA-LSTM (Master's thesis, China University of Geosciences, Beijing). https: //doi.org/10.27493/d.cnki.gzdzy.2023.000933
[4] Zhu, F. (2023). Stock price prediction of chemical listed companies based on news factors and other multimodal features (Master's thesis, Shanghai University of Finance and Economics). https: //doi.org/10.27296/d.cnki.gshcu.2023.002017
[5] Lin, Z. H. (2025). Research on stock index price prediction and quantitative investment strategy based on GRU neural network (Master's thesis, Guangdong University of Technology). https: //doi.org/10.27029/d.cnki.ggdgu.2025.003098
[6] Zhu, R. (2023). Research on stock index price prediction and quantitative investment strategy based on LSTM neural network (Master's thesis, Yunnan University of Finance and Economics). https: //doi.org/10.27455/d.cnki.gycmc.2023.000333
[7] Zhou, Q. (2023). Research on multi-source information stock price prediction based on CNN-LSTM hybrid neural network (Master's thesis, Hebei University of Technology). https: //doi.org/10.27105/d.cnki.ghbgu.2023.000224
[8] Li, C. (2022). Research on quantitative investment strategies for stock price prediction based on LSTM neural network (Master's thesis, Yunnan University of Finance and Economics). https: //doi.org/10.27455/d.cnki.gycmc.2022.000921
[9] Li, T. T. (2021). Stock price prediction based on time series analysis and neural network model (Master's thesis, Dongbei University of Finance and Economics). https: //doi.org/10.27006/d.cnki.gdbcu.2021.001215
[10] Leng, T. X. (2021). Stock price prediction analysis based on attention mechanism and LSTM neural network (Master's thesis, Nanjing University of Science and Technology). https: //doi.org/10.27241/d.cnki.gnjgu.2021.001868