

# A time-series stock price prediction and asset portfolio optimization model based on long- and short-term memory networks

Ziqi Wang<sup>1,✉</sup>

<sup>1</sup> BI Norwegian Business School, Oslo, 0445, Norway

## ABSTRACT

Aiming to address shortcomings in existing time series prediction models, this paper proposes an LSTM model enhanced by fused multi-scale convolutional attention (MCA-LSTM). We design the experimental parameters, construct a stock price dataset, and model the improved LSTM using individual stock closing prices, with prediction accuracy evaluated via RMSE, MAPE, and MAD. To assess the arbitrage and generalization performance of the MCA-LSTM portfolio model, we compare the application of the MCA-LSTM-BL model. Furthermore, within the framework of a mean semi-absolute deviation (MSAD) portfolio optimization model, we develop a new portfolio optimization approach based on return forecasting (MCA-LSTM+MSAD). The asset values and return predictions of various portfolio models are analyzed under transaction cost considerations, and the proposed MCA-LSTM+MSAD model achieves an excess return of 56.98%, consistently maintaining the highest portfolio value throughout the trading period. Overall, our findings indicate that the MCA-LSTM+MSAD model is a promising tool for portfolio optimization and warrants further development for real investment applications.

*Keywords:* LSTM model, attention mechanism, stock price prediction, msad, portfolio

## 1. Introduction

Portfolio optimization refers to choosing the optimal portfolio in the investment process in order to maximize the investment return or reduce the investment risk. For individual and institutional investors, portfolio optimization is an important means to achieve financial goals [7, 3, 2]. Stock market is a high-risk and high-return field, every day the stock market is constantly fluctuating, for

✉ Corresponding author.

*E-mail address:* [norgeziqiwang@163.com](mailto:norgeziqiwang@163.com) (Z. Wang).

Received 21 June 2024; Accepted 27 September 2024; Published Online 17 March 2025.

DOI: [10.61091/jcmcc124-17](https://doi.org/10.61091/jcmcc124-17)

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

investors, how to accurately predict the stock price is a very important issue. In the stock market, the trend prediction of stock price time series has been a key research direction in academia [10, 20, 18, 4], with the goal of grabbing relevant information from historical stock index prices and predicting future stock price time series data. With the development of machine learning and artificial intelligence, stock price prediction models have gradually received widespread attention [6, 9, 15].

Stock price prediction models can be applied in stock trading to guide investment decisions by predicting future price trends. At the same time, the stock price prediction model can also be applied in financial risk management to avoid financial risks by predicting the fluctuations of the stock market [16, 12, 13, 5]. In addition, the stock price prediction model can also be applied in the analysis of industry and enterprise development, through the prediction of industry trends, to guide the development strategy of enterprises [19, 17, 22]. At the same time, the stock price prediction model can also be applied in economic forecasting, through the prediction of the stock market, to predict the trend of economic development [8, 14].

In this paper, we propose to improve the basic architecture of LSTM model, on the basis of LSTM model, its encoder extracts feature information through multi-scale convolution and fuses the attention mechanism to generate a new stock price prediction model. Preprocess the stock dataset to screen the vacancy information and clarify the meaning of each code. The prediction accuracy of the improved model is evaluated using three indicators, RMSE, MAPE and MAD, as evaluation indicators. Analyze the quantitative investment model based on LSTM, and form the portfolio optimization model based on improved LSTM by combining portfolio related algorithms. Simulated trading under the consideration of risk hedging and simulated trading under the consideration of transaction cost are carried out respectively to verify the feasibility of the improved LSTM-based stock portfolio optimization model proposed in this paper.

## 2. Improved LSTM-based stock price prediction

### 2.1. LSTM model

Long Short-Term Memory Networks (LSTM) belong to a special class of Recurrent Neural Networks (RNN) and are intended to solve the problem of long term dependency that exists in recurrent neural networks [21, 1].

The LSTM model has four elements per cell, which are cell state  $C_t$ , input gate  $i_t$ , output gate  $O_t$  and forgetting gate  $f_t$ . where the relevant formulas are given below:

$$f_t = \text{Sigmoid}(W_f o[h_{t-1}, x_t] + b_f), \quad (1)$$

$$i_t = \text{Sigmoid}(W_i o[h_{t-1}, x_t] + b_i), \quad (2)$$

$$C_t = f_t C_{t-1} + i_t \tanh(W_c o[h_{t-1}, x_t] + b_c), \quad (3)$$

$$O_t = \text{Sigmoid}(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (4)$$

$$h_t = O_t \tanh(C_t). \quad (5)$$

In the above equation  $x$  is the input vector.  $h$  is the output vector,  $C$  is the unit state,  $t$  subscripts are the moments. Sigmoid, tanh are the activation functions,  $w$  is the weight matrix,  $b$  is the deviation matrix. The specific structure of the LSTM neuron is shown in Figure 1.

The role of the unit state is to update the information of the previous moment and integrate the information of the current moment, so as to form a memory for long-term information. The input gate determines whether the new information needs to be memorized or not, it first goes through the

tanh layer to represent the information, and at the same time, it goes through the sigmoid layer to determine which information is important, and finally, it is stored with the output layer of the tanh after computation to the unit state. The result of the output gate is the current unit state weighted by the tanh. The forgetting gate determines the information that is discarded or retained in the unit state through the Sigmoid function.

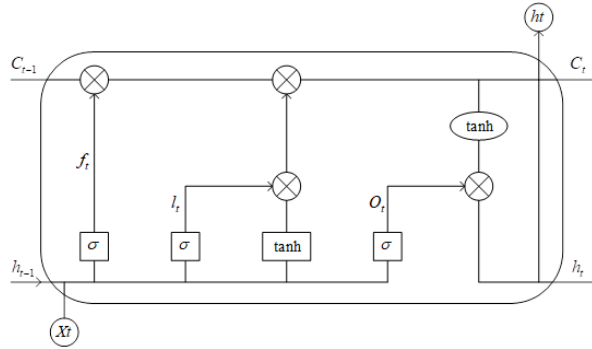


Fig. 1. LSTM neuron structure

2.2. LSTM model prediction

2.2.1. Description of the problem. In the existing time series data prediction models, because of the limitation of the size of the memory cells, the long-term dependency between time steps cannot be effectively preserved while updating the changes of memory information in the short term. In contrast, the stock price prediction studied in this paper is based on a multivariate prediction univariate model, in which all feature variables pass through the hidden layer nodes of equal weight within each time cell, resulting in an inability to distinguish in what way the information about the target sequence to be predicted is dispersed among the nodes of the many hidden layers. In theory, the output of this network structure can be used as the prediction result of  $c$  of any one feature, and cannot accurately fit the real result of the target features.

In order to solve the above deficiencies, this paper splits the stock time series of multivariate features, and takes the daily closing price as the target sequence for prediction, while other variable features are categorized as exogenous sequences. The encoder extracts the features of the exogenous sequence in the time dimension through multi-scale convolution, obtains the exogenous feature information of different time spans, and generates the context vectors of the corresponding moments to be merged with the target sequence through the fusion-attention mechanism of the decoder, and outputs the prediction results through the decoder. This approach not only captures the long time dependency of the target sequence, but also considers the influence of the exogenous sequence on the target sequence at each moment. Given that this paper is forecasting the short-term price of stocks, the effect of the rise and fall of stocks on the average value of the price in the short term can be ignored without considering the short-term price mutability.

2.2.2. Model architecture. Given an exogenous sequence with a time step of  $T$  and a number of variables of  $n$ , the model aims to learn a nonlinear mapping to the current values of the target sequence  $\hat{y}_{T+1}$  through the past values of the target sequence  $(y_1, y_2, \dots, y_T)$ . And the current and past values of the exogenous sequence  $(x_1, x_2, \dots, x_T)$  as shown in Eq. (6). Where  $F(\cdot)$  is the learned target mapping function. When decoding the target sequence, the residual term of the target value at the  $T + 1$  moment is preset to be the mean value of the previous  $T$  moments, and the input at

each moment is to be subtracted from the mean value of the target sequence, and the final prediction result is the last output of the decoder plus the preset residual term. That is:

$$\hat{y}_{T+1} = F(y_1, y_2, \dots, y_T, x_1, x_2, \dots, x_T). \quad (6)$$

In order to integrate the effects of exogenous sequences on the target sequence at different moments, and at the same time to prevent the degradation phenomenon caused by the model being too deep, this paper proposes a new stock price prediction model MCA-LSTM.

(1) *Multi-scale feature extraction*: Although the exogenous sequences are not the prediction task of the model, there is no need to consider how these exogenous features interact with each other. However, each exogenous sequence variable drives the change of the target sequence, and this influence changes with the different values of the target variables at different moments. Therefore, 1D convolution can be used to extract features in the time dimension from the exogenous sequence, and multiple convolution kernels of different sizes can be set up to obtain more complex multi-timescale feature information. And these different scales of information are fused as the output of the encoder, as shown in Eqs. (7) and (8):

$$f_{i,t}^d = \text{Conv} \left( \sum_{k=1}^d (W_{i,k-1}, x_{t+k-1}) + b_i^d \right), \quad (7)$$

$$X_f = \text{Concat}(f^{d1}, f^{d2}, f^{d3}, \dots)^\top, \quad (8)$$

where  $f_{i,t}^d$  denotes the feature value of the  $i$ rd convolutional kernel with step  $d$  of the exogenous sequence, and  $X_f$  denotes the feature after fusing multi-scale information.

(2) *Fusion Attention Mechanism*: Considering that each exogenous feature has its own semantic information, it will affect the target sequence in different dimensions at different moments. So these exogenous influences of different dimensions can be weighted and fused through the attention mechanism at each moment of the decoder. When calculating the weights, it needs to rely on the hidden state information of the decoder at the previous moment, and the fused features are called context vectors, denoted by  $c_i$ , as shown in Eqs. (9) to (10):

$$e_t^i = v_e^\top \tanh(W_e[h_{t-1}; s_{t-1}] + U_e X_f^i), \quad (9)$$

$$\alpha_t^i = \frac{\exp(e_t^i)}{\sum_{k=1}^n \exp(e_t^k)}, \quad (10)$$

$$c_t = \sum_{i=1}^n (\alpha_t^i X_f^i), \quad (11)$$

where  $e_t^i$  denotes the score of the  $i$ th exogenous feature to the decoder at moment  $t$ ,  $\alpha_t^i$  is the weight coefficient of the  $i$ th exogenous feature at moment  $t$ , and  $c_t$  is the context vector computed with the sequence of exogenous features at moment  $t$ .

The  $\tilde{y}_t$  obtained by combining the context vector  $c_t$  and the input  $y_t$  of the decoder can be used to update the hidden state of the decoder at the moment  $t$  as shown in equation (12). Namely:

$$\tilde{y}_t = \tilde{W}^\top [y_t; c_t] + \tilde{b}. \quad (12)$$

### 2.3. Screening and construction of stock dataset

In this paper, we mainly collect the data of constituent stocks in CSI 300 index to construct the stock dataset, and the data are obtained from Yahoo Finance platform [11].

In the process of dataset construction, the individual stock samples that are representative of the market are selected from the 300 stock samples. The data are downloaded from the official channels of Yahoo Finance and saved in CSV format in My SQL database for unified management and maintenance.

The trading data of 20 CSI 300 constituent stocks from January 16, 2012 to December 31, 2022 were obtained through the public data interface provided by Yahoo Finance as the most important data collection object. Through the data acquisition results, the mismatched data volume is eliminated. Due to the lack of the original data, it is necessary to do appropriate preprocessing of the acquired stock price data before it can be stored in the My SQL database and used as the data base for the empirical study.

After data preprocessing, taking the individual stock data of China Merchants Bank (stock code 600036) as an example, the results of data set processing are shown in Table 1. In the table, code is the stock code, date is the date, and YYYYMMDD is the standard format, such as 20120116. Open is the opening price and CLOSE is the closing price. High is the highest price, low is the lowest price, and volume is the turnover. On January 16, 2012, the highest price of China Merchants Bank (stock code 600036) was 18.27.

**Table 1.** Pre-processing data structure

Code	Date	Open	Close	High	Low	Volume
600036	20120116	17.63	17.33	18.27	17.02	756923.21
600036	20120117	18.43	18.15	18.45	17.89	905361.57
600036	20120118	19.04	18.46	19.15	18.55	792213.44
600036	20120119	18.66	17.74	18.71	17.74	102469.06
600036	20120120	16.75	16.85	16.89	16.25	836045.01
600036	20120121	Weekend rest	Weekend rest	Weekend rest	Weekend rest	Weekend rest
600036	20120122	Weekend rest	Weekend rest	Weekend rest	Weekend rest	Weekend rest
600036	20120123	16.75	16.25	16.81	16.21	146926.37
600036	20120124	18.69	17.64	18.74	17.62	113625.08

After preprocessing, the stock price data ensures that the length of time of each constituent stock is consistent and the time order of the data is strictly guaranteed. The time series data in this project is in the frequency of days, which contains more than 620,000 stock price data, and the subsequent experimental research will be carried out on the basis of this dataset.

#### 2.4. Analysis of model prediction results

2.4.1. Experimental setup. The tools and parameter settings used for modeling are shown in Table 2.

2.4.2. Experimental results and evaluation analysis. In this section, the modeling and prediction of the improved LSTM model is carried out using the closing prices of 12 stocks as the data base. And the model prediction accuracy is evaluated with three indicators, namely, RMSE, MAPE and MAD, as evaluation indicators. Taking four of the stock samples as an example, the prediction effect is shown in Figure 2, and Figures 2a, 2b, 2c, and 2d show the stock price prediction comparisons of Guizhou Moutai, Ping An of China, CITIC Securities, and China Merchants Bank, respectively.

**Table 2.** The work and parameter configuration of the experiment

Configuration item	Configuration
Programming language	Python
Third-party libraries used	pandas, numpy, matplotlib.pyplot, tensorflow, scipy.spatial.distance
Neural cryptography	12
Neural network input data dimension	8
Neural network output layer number	2
Batch_size	3
Time_step	9
Learn rate	0.0003

The black line in the figure is the actual value and the red line is the predicted value. From the figure, it can be seen that the improved LSTM model has a smaller prediction error for the four stock samples, and the predicted value of the stock price of CITIC Securities and China Merchants Bank is better. The stock price predictions for Guizhou Maotai and Ping An of China have some errors at the highest price. The improved LSTM model predicts that the highest price of Guizhou Moutai occurs on the trading day of 275-300 days. And the stock price during that trading day does not have the highest price although it has an upward trend.

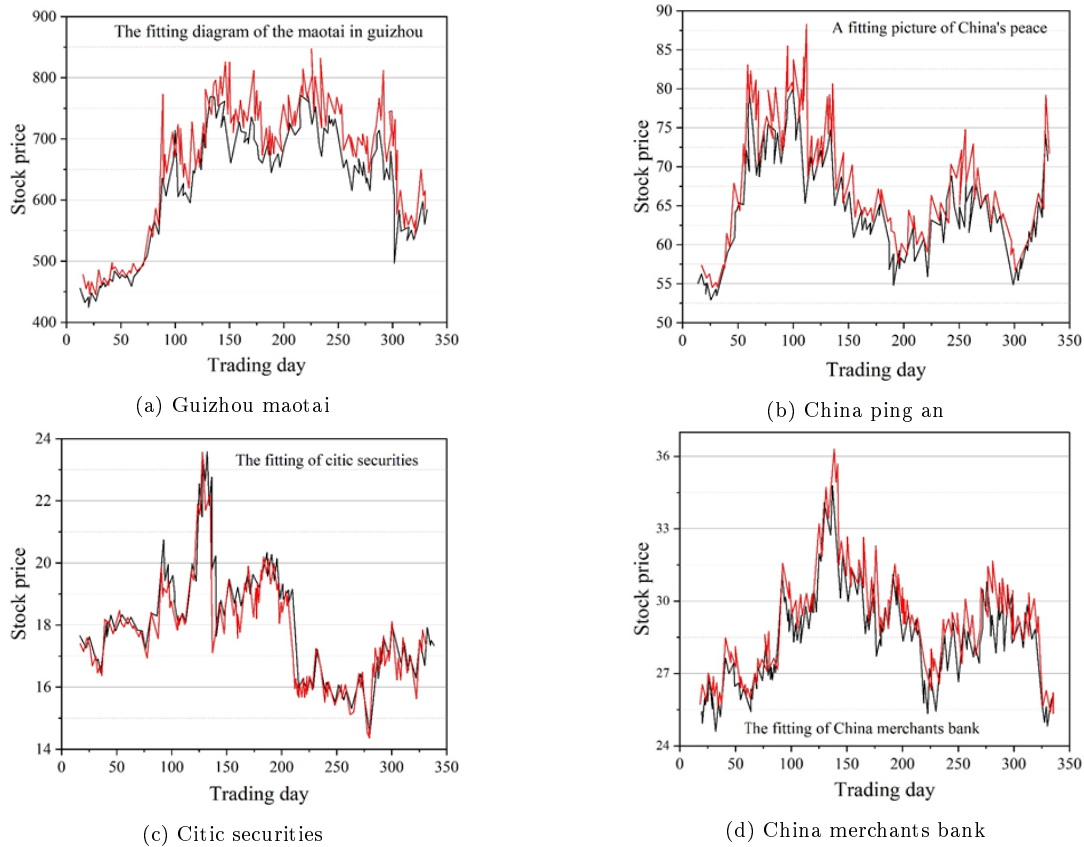
The corresponding experimental results for all 12 stocks are then shown in Table 3.

**Table 3.** The results of all 12 individual stocks

Experimental sample	RMSE	MAPE	MAD
Guizhou maotai 600519	28.6539	6.122	26.078
China ping an 601318	18.9664	17.003	2.99
Grain liquid 000858	1.3651	1.891	0.912
China merchants bank 600036	0.5399	1.424	0.431
Hengrui 600276	1.1243	1.739	0.897
Citic securities 600030	0.3651	0.634	0.102
Gree electric appliance 000651	4.3005	6.528	3.365
Yili shares 600887	0.5872	1.469	0.421
China free 601888	0.4175	1.034	0.365
Societe generale 601166	0.2693	0.728	0.142
Vanke A 000002	7.1694	16.162	5.396
ICBC 601398	0.0981	1.408	0.072

The improved LSTM model has smaller prediction errors on most stock samples. However, there are still some samples with large errors between the predicted and actual values, such as Guizhou Maotai (600519) and Ping An of China (601318). The RMSE values of Guizhou Maotai and Ping An of China reach 28.6539 and 18.9664, respectively.

By analyzing the data sample can be obtained, the above stocks data base is larger compared to other stocks, their closing price data is generally in the thousands. While most of the other stocks sample closing price is a few dollars to a few hundred dollars ranging, the above two stocks and other stocks with a large difference in value. So the fluctuation range of the predicted value is also much higher than most other stock samples.



**Fig. 2.** Compared with the four stock price forecasts of the improved LSTM model

Through the analysis of the above experimental results, it can be concluded that the performance of the short-term and long-term memory network model in the problem of stock closing price prediction has achieved the expected results, but in some sample data need to make the corresponding normalization of the data and the adjustment of the model parameters so that the model can get a more ideal effect.

### 3. Stock portfolios based on improved LSTM return prediction

#### 3.1. LSTM-based quantitative investment modeling

3.1.1. Overall framework for quantitative investment modeling. In order to realize real-time automated trading of stocks, this project proposes a quantitative investment model for stocks based on an improved LSTM model.

The model aims to achieve the following objectives:

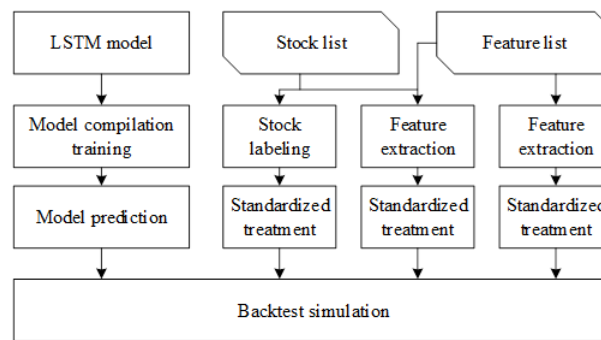
On the one hand, the model predicts the return of the stock and visualizes the potential trend of the stock, so as to provide investors with a reference for decision-making.

On the other hand, through the risk indicators of the model, the risk level of the transaction is assessed so that investors can better manage the risk and maximize the return.

The overall framework of the model is shown in Figure 3. In this framework, the stock data are first predicted using the LSTM model to obtain the predicted returns, and this step takes advantage of the LSTM model's strengths in handling time series data, which can capture the long-term dependencies in the data to more accurately predict the future returns. Next, through risk assessment of the model output returns, corresponding risk indicators can be obtained, which can reflect the risk level

of the trading strategy, helping investors to consider the risk factors when making decisions and make corresponding adjustments according to their own risk tolerance. Finally, based on the results of the model prediction and risk assessment, investors can formulate the corresponding trading strategy and conduct live trading. In this way, investors can trade according to the guidance of the model in a real-time market environment in order to realize the value-added of their investment portfolios.

The improved LSTM-based quantitative stock investment model plays an important role in the establishment of the LSTM-based quantitative stock investment methodology, the automated trading function of the live market, and helps investors to make informed investment decisions and maximize their returns during the investment process by providing yield prediction and risk assessment.



**Fig. 3.** A comprehensive prediction model based on LSTM

3.1.2. Quantitative investment modeling process. The factor data selected in this paper can be categorized into two types, fundamental and technical factor data. Fundamental factors use monthly data, intercepting data from cross-sectional dimension, used for multi-factor regression stock selection session, containing basic earnings per share, operating income per share, operating profit per share, retained earnings per share, net profit growth rate. The technical factors use weekly data, intercepted from the time series dimension, and are used in the LSTM model for stock price prediction, containing the maximum price per share and the closing price. In this paper, a total of seven factors are selected, five fundamental factors and two technical factors. In order to carry out the initial screening of the factors, it is necessary to carry out factor validity analysis, and the selected valid factors are used in the training fitting process of the model in this paper to provide protection for the subsequent process. The basic workflow of the model is as follows: first, all stocks in the A-share market are automatically obtained and a certain number of factors are calculated as sample features. These factors can include company fundamental data, technical indicators, market sentiment and so on. Next, calculate the return of each stock within a specified date and normalize the return data to ensure comparability among different stocks. Then, an LSTM model is introduced to predict the future returns of a stock by inputting historical factor data and the corresponding standardized returns. The prediction results can help to determine the trend of the stock and the potential upward or downward trend. Finally, based on the predicted yields, stocks are sorted in descending order and buy and sell operations are executed to maximize portfolio returns.

### 3.2. Combinatorial optimization model construction based on improved LSTM

3.2.1. Portfolio. The basic assumptions of modern portfolio theory include that all investors are in the same single investment period. Investors are risk averse and seek to maximize expected utility. Investors make portfolio selection allocations based only on the return and variance of the underlying



investments.

Using  $E$  to denote the expected return and  $\delta^2$  to denote the variance, the mean-variance model expression is shown in Eq. (13). Namely:

$$\begin{cases} \min \delta^2(r_p) = \Sigma \Sigma w_i w_j \text{cov}(r_i, r_j), \\ E(r_p) = \Sigma w_i r_i, \end{cases} \tag{13}$$

where  $r_p$  is the portfolio return.  $w_i$  is the weight of asset  $i$  in the portfolio, and  $r_i$  is the return on asset  $i$ .  $\text{cov}(r_i, r_j)$  is the covariance between asset  $i$  and asset  $j$ .

The model can be solved by the Lagrange method, which solves for the weights of the assets that minimize the risk of the portfolio when the expected rate of return is determined.

Economically speaking, an investor can minimize the overall investment risk by determining the expected rate of return before investing, and then obtaining the weights of each asset. Each different expected return corresponds to a different portfolio weighting solution, which together form the efficient portfolio, i.e., the portfolio with the lowest variance. The curve formed between the expected return of the efficient portfolio and the corresponding minimum variance is called the efficient frontier of the portfolio. Investors will choose the portfolio solution with the highest utility on the efficient portfolio frontier based on different return expectations and risk-taking levels.

(1) *Capital asset pricing model:* Based on the assumptions in modern portfolio theory, other assumptions in CAPM show a complete capital market, i.e., the absence of any frictions that discourage investment, including the ability of investors to borrow or lend any funds without restriction at the level of the risk-free rate of interest, the absence of fees and taxes for trading securities, and the possibility of unlimited splitting of security shares.

The CAPM expression for a single stock or portfolio is shown in Eq. (14). Namely:

$$\bar{r}_t = r_f + \beta_i(\bar{r}_m - r_f), \tag{14}$$

where  $\bar{r}_i$  is the expected rate of return on a single stock  $i$  or portfolio  $i$ .  $r_f$  is the risk-free rate of return,  $\beta_i$  is the  $\beta$  coefficient of the asset  $i$  or portfolio  $i$ .  $\bar{r}_m$  is the expected return of the market portfolio.

The CAPM gives a simple conclusion that there is only one factor that will lead to a higher return on investment, and that is investing in risky stocks.

(2) *Sparse and stable portfolio selection:* For a standard portfolio selection problem with  $N$  risky asset, at moment  $t$ , the excess return  $R_t$  obeys a multivariate normal distribution with mean  $\mu$  and variance-covariance matrix  $\Sigma$ , where  $\mu$  is a  $N \times 1$ -dimensional column vector and  $\Sigma$  is an  $N \times N$ -dimensional matrix. At moment  $t$ , the investor needs to determine the portfolio weights  $w$  to maximize the mean-variance objective function as shown in Eq. (15). Namely:

$$U(w) = w^T \mu - \frac{\gamma}{2} w^T \Sigma w, \tag{15}$$

where  $U$  is the utility obtained by the investor.  $\gamma$  is the investor risk aversion coefficient. The optimal portfolio weight is  $w = \gamma^{-1} \Sigma^{-1} \mu$ .

And for the multiple linear regression with  $N$  independent variables and  $N$  observations  $y = Xw + e$ ,  $e$  is the random error term that makes:

$$\begin{cases} X = \sqrt{\gamma} \Sigma^{-\frac{1}{2}}, \\ y = \frac{1}{\sqrt{\gamma}} \Sigma^{-\frac{1}{2}} \mu. \end{cases} \tag{16}$$

Then the least squares estimator of the multivariate linear regression,  $\hat{w}_{oLS} = (X^T X)^{-1}(X^T y)$ , equals the optimal portfolio weights,  $\hat{w} = \gamma^{-1}\Sigma^{-1}\mu$ . In other words, the least squares estimator solves the portfolio selection problem.

Since investors do not know the real  $\mu$  and  $\Sigma$ , among the estimators based on historical data, the great likelihood estimators  $\hat{\mu}$  and  $\hat{\Sigma}$  are widely used, and then the optimal portfolios are found by Eq. (17). Namely:

$$U(w) = w^T \hat{\mu} - \frac{\gamma}{2} w^T \hat{\Sigma} w. \tag{17}$$

There are three sources of estimation error in Eq. Estimated mean  $\hat{\mu}$ , estimated covariance matrix  $\hat{\Sigma}$ , and inverse matrix of estimated covariance matrix.

When the number of assets is large, a sparse portfolio needs to be constructed. First, a sparse portfolio reduces transaction and management costs. Second, by setting the weights of smaller portfolios to zero, the estimated portfolio weights are no longer unbiased, but their variance and mean-square prediction errors can be reduced. Sparse portfolio weights can be obtained by imposing Lasso constraints on the portfolio weights. The portfolio weights are estimated through Eq. (18). Namely:

$$\hat{w}_{L1} = \operatorname{argmax} \left\{ w^T \hat{\mu} - \frac{\gamma}{2} w^T \hat{\Sigma} w \right\} \text{ s.t. } \|w\|_1 < s_1, \tag{18}$$

where  $s_1$  is a constant greater than zero.

When  $s_1 > 0$ , the portfolio weights  $\hat{w}_{L1} = (\hat{w}_{L1,1}, \dots, \hat{w}_{L1,N})^T$  shrink toward zero. If the OLS estimates are small enough in absolute value, the penalized least squares  $\hat{w}_{L1,j}$  is exactly zero.

In mean-variance efficient portfolios, extreme weights usually occur and portfolio weights can change significantly when new return information is used and when a set of assets is not available for trading. This is caused by large estimation errors in the inverse matrices of  $\Sigma$  and  $\hat{\Sigma}$ . When the returns of the two assets are highly correlated, the inverse matrix of  $\hat{\Sigma}$  becomes highly unstable and causes the weights of the two assets to fluctuate substantially over time. Therefore, imposing stability constraints is expected to reduce the estimation risk due to parameter uncertainty and multicollinearity. Replace  $\hat{\Sigma}$  with  $\hat{\Sigma}_s = v\hat{\Sigma} + (1 - v)\hat{\Sigma}_g$ , where  $\hat{\Sigma}_g$  is a contraction target with low variance and  $v$  is the contraction intensity. This is equivalent to imposing a constraint on the sum of squares of the portfolio weights.

The portfolio weights can be estimated by Eq. (19). Namely:

$$\hat{w}_{L1L2} = \operatorname{argmax} \left\{ w^T \hat{\mu} - \frac{\gamma}{2} w^T \hat{\Sigma} w \right\} \text{ s.t. } \|w\|_1 < s_1, \|w\|_2^2 < s_2. \tag{19}$$

When  $s_1 > 0$  and  $s_2 > 0$ , the portfolio weights are first scaled and then contracted towards zero, thus improving the sparsity and stability of the constructed portfolio.

3.2.2. Equity portfolio modeling. (1) *MCA-LSTM-BL stock portfolio modeling*: The Black-Litterman model has been widely used for asset allocation decisions within the investment field. Its core idea uses Bayesian statistics to introduce asset returns based on market equilibrium assumptions. Combined with the investor’s prediction of the return of the investment product, thus obtaining the return of the investment product. And this is incorporated into the mean-variance portfolio theory, which in turn confirms the asset allocation that meets the investor’s viewpoint to obtain the best portfolio solution.

In this study, we optimize the LSTM to predict the future returns of stocks to achieve the goal of obtaining the subjective viewpoint parameters of the BL model, and input them into the BL model

to obtain the allocation weights of the stock portfolio. And the assets are allocated in accordance with the model output, constituting the MCA-LSTM-BL stock portfolio model.

(2) *Portfolio optimization model based on improved LSTM network return prediction*: Based on the framework of the mean semi-absolute deviation (MSAD) portfolio optimization model, a new portfolio optimization model based on return prediction is established. That is, the portfolio optimization model based on improved LSTM network return prediction (MCA-LSTM +MSAD).

### 3.3. Simulated trading with consideration of risk hedging

In order to test the performance of the MCA-LSTM stock portfolio model in terms of arbitrage and generalization, this section applies the analysis of the MCA-LSTM-BL model in terms of data sources, indicator factor selection and parameter setting, model evaluation indexes, and comparative analysis between the model and the data to further demonstrate the practical value of the model.

The research data are still selected from the same stocks as in the previous section, the constituent stocks in the January 2022 stock pool of CSI 300 index. The daily frequency data is from January 16, 2012 to December 31, 2022 by applying the analysis to the stocks.

3.3.1. *Indicator factor selection and parameterization*. In this section, four models are constructed, MCA-LSTM-BL, market capitalization weighting, historical BL, and LSTM-BL models. Among them, the market capitalization weighting model uses the market capitalization of 120 stocks to calculate the weights, construct the portfolio and hold it for a long period of time to calculate the returns. For the historical BL model, the BL model is used to determine the weights of the portfolio and change the positions daily. In the BL model, the historical average return of  $T=145$  days is used as the viewpoint return,  $\tau$  set to 0.035. The weights calculated by the BL model may be negative, based on the restrictions on shorting in the Chinese stock market, this paper sets the negative weights predicted by the BL model to zero, and normalizes the final weights as the final weights.

The two models, MCA-LSTM-BL and LSTM-BL, mainly use the returns predicted by the MCA-LSTM model and the LSTM model as the viewpoint returns input into the BL model. Since the LSTM model belongs to the time-series model, which is more sensitive to recent data and is not suitable for predicting too long a time period in the future, this paper adopts a rolling training approach, where every 50 days, the model is trained using data from the past 100 days, and the results of the yields for the day's backward 50 days are predicted, and so on, until the end of the training set.

3.3.2. *Comparative analysis of data*. The data is selected as stock data of trading days during January 16, 2020 January 20, 2022 as a comparative experiment to test the application effectiveness of the model. The dataset includes daily frequency data of 485 trading days, and the historical data from January 12, 2021 to January 16, 2022 is backtested by adopting a rolling training approach. The comparison of asset values of different portfolio models is shown in Figure 4. In the portfolio asset value graph, the vertical axis indicates that the initial asset is 1, and the asset value formed by accumulating according to the rate of return. It is easy to see from the graph that the MCA-LSTM-BL model has the highest portfolio asset value and the best return from the portfolio, which better validates the stability of the model. The portfolio constructed by the LSTM-BL model follows closely in the comparison of model results, and the worst performer is still the market capitalization weighting method. Once again, it confirms that the new model has good effectiveness in stock

portfolios. At the same time, the trend of the curve in the figure shows that the overall asset values have a fluctuating downward trend, which is in line with the general environment of the stock market in the dataset area, reflecting the reliability of the study. A comparison of the effectiveness of yield prediction using LSTM and MCA-LSTM in terms of model prediction accuracy under the new data set is shown in Figure 5.

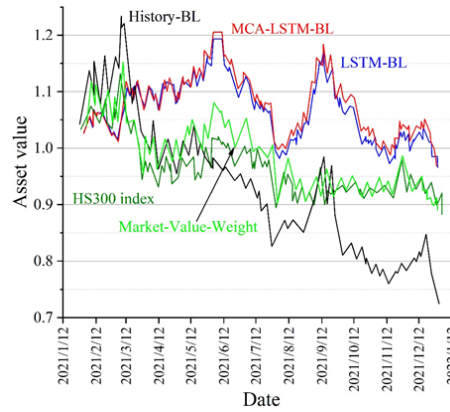


Fig. 4. Comparison of asset values of different portfolio models

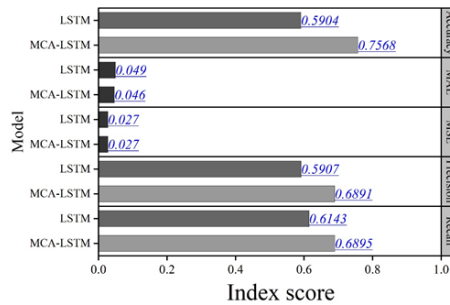
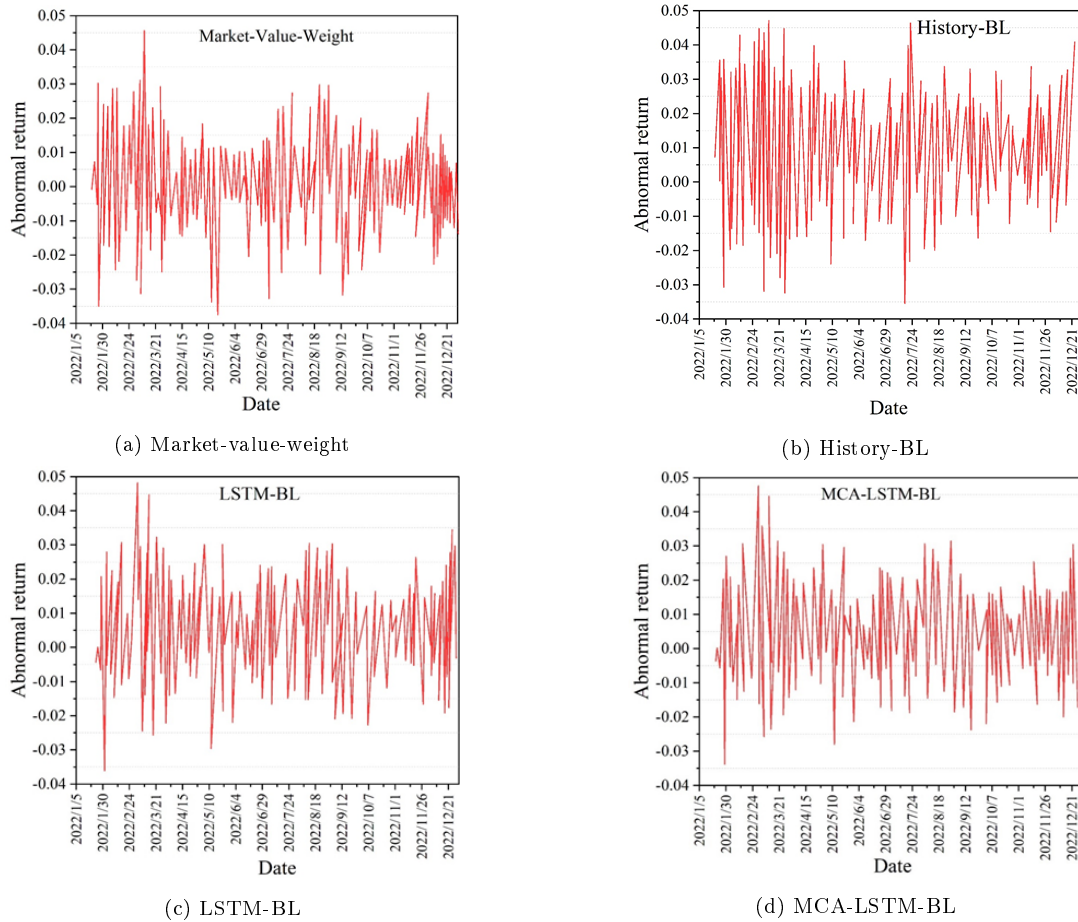


Fig. 5. Predictive performance contrast

The errors of the two prediction methods in the comparison of the datasets are slightly higher compared to the previous ones, but they also reflect a good performance within the acceptable range. The five comprehensive measures of the model’s predictive ability in the table reflect that MCA-LSTM outperforms LSTM, with the Accuracy of MCA-LSTM being 0.7568. This indicates that the MCA-LSTM model has high accuracy and generalization ability in yield prediction. The differences between the different model portfolio returns and the market stock index returns are compared under the new data set. The different model abnormal return curves are shown in Figure 6. The difference between different portfolios and the HS300 index market is reflected by the different modeled abnormal return curve plots in the figure. Also in the comparison, it can be concluded that the MCA-LSTM-BL model has the smallest fluctuation and the smoothest amplitude of the portfolio abnormal return (which is equivalent to the excess return here), which still shows a good stability, and similarly in contrast to the market capitalization weight methodology, although there are sporadic and more prominent highs in the test set, there is a negative return in high frequency, and there is less fluctuation near the zero, which shows that the portfolio represented by its model cannot bring stable excess returns and the investment effect is not satisfactory. Through the experimental comparison, it can be concluded that the MCA-LSTM-BL stock portfolio model has a good return.



**Fig. 6.** Abnormal yield curve of different models

3.4. Simulated transactions taking into account transaction costs

Among all DNNs, Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) are the most commonly used models for stock price prediction. Therefore, the purpose of this section is to further improve the out-of-sample performance of return prediction-based portfolio optimization models by using LSTM networks and CNNs.

Based on the framework of the mean semi-absolute deviation (MSAD) portfolio optimization model, a new return prediction-based portfolio optimization model is developed. In order to illustrate the merits of these models, three equally weighted portfolio models are selected in this section for comparison, who select their stocks using improved LSTM, LSTM network and CNN, respectively. And two portfolio models based on Support Vector Regression (SVR) return prediction are used as benchmarks, and they use SVR instead of DNNs for return prediction.

It is well known that trading costs can have a significant impact on the returns of trading strategies. And a high turnover rate can lead to high trading costs. Therefore, it is meaningful to test the actual performance of the portfolio optimization model based on return prediction after deducting the trading costs. For simplicity, this section only considers a 0.06% per unit turnover rate to study the performance of different models.

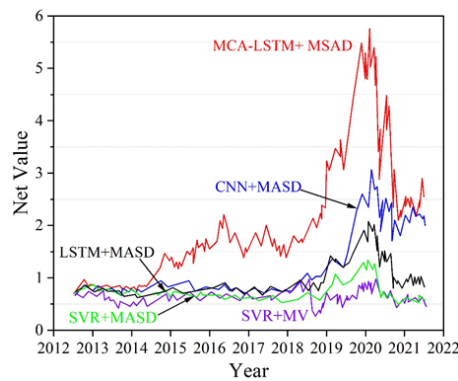
This section applies  $R_p = 0.03$  to discuss the performance of different portfolio optimization models based on return forecasting. The performance of different models considering transaction costs is shown in Table 4. ER, SD, IR, TR, and MD are denoted as excess return, standard deviation, informatization rate, total return, and maximum retracement, respectively. The improved LSTM

algorithm with mean semi-absolute deviation (MSAD) portfolio optimization model proposed in this paper is able to achieve 56.98% excess return.

**Table 4.** Consider the performance of different models of transaction costs

Model	ER	SD	IR	TR	MD
MCA-LSTM+ MSAD	0.5698	0.5962	0.9175	1.9637	0.7542
LSTM+ MSAD	0.1025	0.2028	0.4938	1.4604	0.4138
CNN+ MSAD	-0.0517	0.0905	-0.4614	0.2976	0.5446
SVR+ MSAD	-0.1493	0.1834	0.8993	-0.0428	0.5299
SVR+MV	-0.2115	0.4617	0.4352	-0.2013	0.5867

The net values of the different models considering transaction costs are shown in Figure 7, which directly shows the performance of the different models on the out-of-sample test set. The MCA-LSTM+MASD model is able to maintain the maximum value consistently over the stock trading day period 2012-2022 and with an  $R_p$  value of 0.03. This paper concludes that MCA-LSTM+MASD is a promising model for portfolio optimization in real investment.



**Fig. 7.** The net worth of different models of transaction costs

## 4. Conclusion

In this paper, a new stock price prediction model MCA-LSTM is proposed by improving the long and short-term memory network algorithm, adding multi-scale feature extraction and attention mechanism, screening and processing the stock dataset, and utilizing the new prediction model to make the time series prediction of stock price. Considering the stock return prediction to form a stock portfolio, the investment efficiency optimization analysis is performed with the help of evaluation indexes.

The improved LSTM stock price prediction model has smaller prediction errors for the four samples of Guizhou Moutai, Ping An of China, CITIC Securities, and China Merchants Bank, which are in line with the demand for predicting price changes in the stock market.

Combining the quantitative investment framework of LSTM and substituting it into the improved LSTM stock price prediction model, it constitutes a portfolio optimization model based on the improved LSTM. In the simulated trading considering risk hedging, the MCA-LSTM-BL model has the highest portfolio asset value and the best return generated by the portfolio, which better validates the stability of the model. The composite measure of predictive ability index of each prediction model

reflects that MCA-LSTM outperforms LSTM in all cases, with the Accuracy of MCA-LSTM being 0.7568. This indicates that the MCA-LSTM model has a high accuracy and generalization ability in yield prediction.

Also under the condition of considering the transaction cost, the portfolio strategy based on the improved LSTM model proposed in this paper can achieve the maximum excess return, which verifies that the MCA-LSTM portfolio model proposed in this paper can be established out-of-sample. Thus, this paper concludes that the MCA-LSTM portfolio model is capable of optimizing portfolio returns in real investment.

## References

- [1] S. Abri and R. Abri. Deep learning methods for lstm-based personalized search: a comparative analysis. *International Journal of Machine Learning and Cybernetics*:1–13, 2024. <https://doi.org/10.1007/s13042-024-02418-7>.
- [2] G.-Y. Ban, N. El Karoui, and A. E. Lim. Machine learning and portfolio optimization. *Management Science*, 64(3):1136–1154, 2018. <https://doi.org/10.1287/mnsc.2016.2644>.
- [3] M. J. Best. *Portfolio Optimization*. CRC Press, 2010.
- [4] O. Bustos and A. Pomares-Quimbaya. Stock market movement forecast: a systematic review. *Expert Systems with Applications*, 156:113464, 2020. <https://doi.org/10.1016/j.eswa.2020.113464>.
- [5] Y. E. Cakra and B. D. Trisedya. Stock price prediction using linear regression based on sentiment analysis. In *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pages 147–154. IEEE, 2015. <https://doi.org/10.1109/ICACSIS.2015.7415179>.
- [6] J.-S. Chou and T.-K. Nguyen. Forward forecast of stock price using sliding-window metaheuristic-optimized machine-learning regression. *IEEE Transactions on Industrial Informatics*, 14(7):3132–3142, 2018. <https://doi.org/10.1109/TII.2018.2794389>.
- [7] A. Gunjan and S. Bhattacharyya. A brief review of portfolio optimization techniques. *Artificial Intelligence Review*, 56(5):3847–3886, 2023. <https://doi.org/10.1007/s10462-022-10273-7>.
- [8] K. Khare, O. Darekar, P. Gupta, and V. Attar. Short term stock price prediction using deep learning. In *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 482–486. IEEE, 2017. <https://doi.org/10.1109/RTEICT.2017.8256643>.
- [9] C. K.-S. Leung, R. K. MacKinnon, and Y. Wang. A machine learning approach for stock price prediction. In *Proceedings of the 18th International Database Engineering & Applications Symposium*, pages 274–277, 2014. <https://doi.org/10.1145/2628194.2628211>.
- [10] K.-P. Lim and R. Brooks. The evolution of stock market efficiency over time: a survey of the empirical literature. *Journal of Economic Surveys*, 25(1):69–108, 2011. <https://doi.org/10.1111/j.1467-6419.2009.00611.x>.
- [11] X. Liu, Y. Wu, M. Luo, and Z. Chen. Stock price prediction for new energy vehicle companies based on multi-source data and hybrid attention structure. *Expert Systems with Applications*, 255:124787, 2024. <https://doi.org/10.1016/j.eswa.2024.124787>.
- [12] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang. A cnn-lstm-based model to forecast stock prices. *Complexity*, 2020(1):6622927, 2020. <https://doi.org/10.1155/2020/6622927>.
- [13] W. Lu, J. Li, J. Wang, and L. Qin. A cnn-bilstm-am method for stock price prediction. *Neural Computing and Applications*, 33(10):4741–4753, 2021. <https://doi.org/10.1007/s00521-020-05532-z>.

- 
- [14] B. Manujakshi, M. G. Kabadi, and N. Naik. A hybrid stock price prediction model based on pre and deep neural network. *Data*, 7(5):51, 2022. <https://doi.org/10.3390/data7050051>.
- [15] S. Mehtab, J. Sen, and A. Dutta. 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*, pages 88–106. Springer, 2021. [https://doi.org/10.1007/978-981-16-0419-5\\_8](https://doi.org/10.1007/978-981-16-0419-5_8).
- [16] P. Mondal, L. Shit, and S. Goswami. Study of effectiveness of time series modeling (arima) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications*, 4(2):13, 2014. <https://doi.org/10.5121/ijcsea.2014.4202>.
- [17] S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. K. Menon, and K. Soman. Stock price prediction using lstm, rnn and cnn-sliding window model. In *2017 International Conference on Advances in Computing, Communications and Informatics (icacci)*, pages 1643–1647. IEEE, 2017. <https://doi.org/10.1109/ICACCI.2017.8126078>.
- [18] D. Shah, H. Isah, and F. Zulkernine. Stock market analysis: a review and taxonomy of prediction techniques. *International Journal of Financial Studies*, 7(2):26, 2019. <https://doi.org/10.3390/ijfs7020026>.
- [19] M. A. I. Sunny, M. M. S. Maswood, and A. G. Alharbi. Deep learning-based stock price prediction using lstm and bi-directional lstm model. In *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, pages 87–92. IEEE, 2020. <https://doi.org/10.1109/NILES50944.2020.9257950>.
- [20] S. Yadav. Stock market volatility-a study of indian stock market. *Global Journal for Research Analysis*, 6(4):629–632, 2017.
- [21] H. Zhan, X. Meng, and M. Asif. Risk early warning of a dynamic ideological and political education system based on lstm-mlp: online education data processing and optimization. *Mobile Networks and Applications*, 29(2):1, 2024. <https://doi.org/10.1007/s11036-024-02439-0>.
- [22] L. Zhang, C. Aggarwal, and G.-J. Qi. Stock price prediction via discovering multi-frequency trading patterns. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2141–2149, 2017. <https://doi.org/10.1145/3097983.3098117>.