Stock Price Forecasting: Traditional Statistical Methods and Deep Learning Methods

Yihui Xie*

School of Mathematical Science, Zhejiang University, Hangzhou, China

*Corresponding author: 3190103783@zju.edu.cn

Abstract. Statistical method and machine learning methods can be applied to stock forecasting to assist users in making decisions in a variety of practical applications. This paper uses 50 past stock prices to predict future stock prices in this experiment. Both ARIMA and LSTM models are used in this study to predict the price. The close price on transaction days makes up the dataset in this study. The performance of two models is evaluated by MAE, MSE, and RMSE after running them in this research. And the result shows that both the LSTM model and ARIMA model predict the stock price well. To be specific, MSE values of ARIMA model are 1.17151, 1.55678, 6.38663 and 159.58281. MSE values of LSTM model are 1.18464, 1.07799, 4.87162 and 97.59937. In these two kinds of methods, LSTM model has a better performance than ARIMA model.

Keywords: Stock Price Forecasting; ARIMA; LSTM.

1. Introduction

Predicting the stock market is to help investors make decisions and minimize risks as much as possible, thereby obtaining investment returns. If someone can accurately predict the rise and fall of stock market prices, he can buy or sell stocks at higher prices in the market, thereby obtaining higher profits. In addition, predicting market trends can help investors avoid unnecessary losses.

There are many studies about stock price forecasting: Li uses A New dimensional gray Markov forecasting model to improve the accuracy of mid to long term predictions, Jiang uses RBF neural network to improve the adaptive ability of prediction [1-2]. Zhang uses genetic algorithm to train and optimize the initial weights of Elman neural network, an efficient GA-ELman dynamic regression neural network stock price prediction model is proposed [3]. Zhang and Liang based on the dimensionality reduction of many variables affecting target data by principal component analysis method, the BP neural network prediction model of stock market closing price is constructed [4]. Xu uses a back-propagation algorithm with adjustable gain [5]. Liu discusses the efficiency and volatility characteristics of financial markets in the framework of non-linear system theory [6]. Long expands the scope of news used in the stock price prediction task from the perspective of news level and studied the prediction effect of multi-level news system on stock price trend [7]. Based on CAE and GRU models, Huang propose a method to predict stock price by using the correlation between K-line and moving average features [8]. Deng proposes a stock price prediction method based on E-V-ALSTM mixed depth model [9]. Lin proposes a stock price rise and fall prediction model TE-EAN based on technical factor empirical mode decomposition and embedded temporal attention Network (EAN) [10]. These models have good predictive performance; however, relevant investigations are still limited, which makes the price forecasting an interesting question in financial area.

This paper uses two models for prediction: ARIMA model and LSTM model. In terms of how well the two models predict outcomes, LSTM performs better than LSTM. This is because using LSTM to anticipate stock price changes is both practical and successful because it displays the internal law governing stock price movements. However, the short training period and limited number of training parameters for ARIMA make it simpler to derive results.

2. ARIMA

A time series forecasting technique called the ARIMA model, which represent Auto regressive Integrated Moving Average Model, was first put forth by Box in the early 1970s. When a non-stationary time series is changed into a stationary time series, the ARIMA model is created by regressing the dependent variable only on its lag value as well as the present value and lag value of the random error component.

Three variables:p, d, and q make up the ARIMA model.

p: AR/Auto-Regressive item, this term refers to the lags of the time series data utilized in the forecast model.

d: Integrated item, needs to be distinguished in various orders to be stable.

q: The MA/Moving Average term, commonly known as the lags of the forecast error utilized in the forecast model.

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d X_t=(1+\sum_{i=1}^{q}\theta_i L^i)\epsilon_t \tag{1}$$

ϕ_i is the parameter of AR, θ is the parameter of MA, L is the lag operator, and d is a positive integer.

3. LSTM

Recurrent neural networks, also known as RNN, have the ability to transfer historical data using a chain-like neural network architecture (Zhou et al., 2015). It will examine the previous hidden state's output h(t-1) and current input i(t) in each time period when processing sequential data (Zhou et al., 2015). However, the typical RNNs are unable to acquire long-term reliance when there is a greater gap between the two time steps. Therefore, In order to combat the problem of "long-term dependence," Hochreiter and Schmidhuber (1997) developed LSTM cells, which extend the memory capacity of traditional recurrent cells by adding a "gate" to the cell. It's important to note that LSTM was developed specifically for time series data.

3.1. Principle of LSTM

A forgetting gate was added to the original LSTM by PapeGers, Schmidhuber, and Cummins (2000), as depicted in Fig. 1. An LSTM cell with a forget gate is designed with the core purpose of aligning. Two fundamentally different but complementary purposes for its internal operations: data and data control. The control component controls signals in the 0 to 1 range, whereas the data component is in charge of prospective data signals. The candidate data signal is multiplied by the control signal to determine the fraction of the candidate data that is allowed to transit to a certain node in the cell. Because of this, the forgetting gate stores all information while its value is 1, and discards it when its value is 0. The relevant portion of the data for the control signal's intermediate value (between 0 and 1) will be provided to the following function.

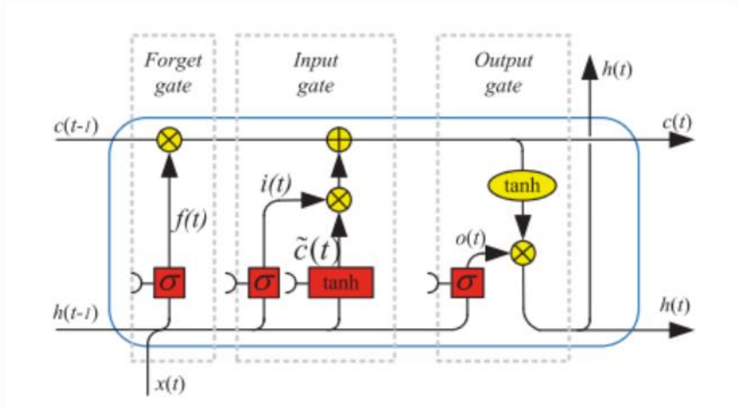

Fig. 1 Architecture of LSTM

3.2. Mathematical Expression

An LSTM cell with a forget gate is expressed mathematically as follows:

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \tag{2}$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \tag{3}$$

$$\tilde{c}_t = \tanh\,(W_{\mathrm{ch}}h_{t-1} + W_{\bar{r}}x_t + b_{\bar{r}}) \tag{4}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \bar{c}_t \tag{5}$$

$$o_t = \sigma(W_{ch}h_{t-1} + W_{ox}x_t + b_o) \tag{6}$$

$$h_t = o_t \cdot \tanh\,(c_t) \tag{7}$$

$\sigma$is a logarithmic sigmoid function with has an output in [0, 1], "·"stands for element multiplication, and tanh, as a hyperbolic function, has an output at [-1, 1].

According to Jozefowicz et al. (2015), In general, the performance of bf and LSTM networks would be enhanced by increasing the forget gate's bias. Additionally, Schmidhuber suggested that, occasionally, LSTM can be trained using a hybrid approach that combines evolutionary algorithms and other methods as opposed to using only gradient descent.

4. Results

4.1. Data Importing

As the research object for this article, 4 firm stocks were chosen from finance.yahoo.com: Apple, Google, Netflix and Amazon. The closing price on business days, which from January 1, 2015, to December 31, 2018, is the dataset used in this study. This study used 60 data from both models to forecast the following day. Hence, there would be 60 inputs and 1 output at once.

4.2. Parameter Estimation

Three parameters make up the ARIMA model: p, d, and q. Through several statistical indicators, the parameters for ARIMA are summarized in the following Table 1.

Table 1. Parameters of ARIMA model

	Apple	Google	Amazon	Netflix
p	2	1	1	2
q	0	0	1	1
d	1	1	1	1

Normalizing the stock price data is important before creating an LSTM model. This study reformats the dataset using the "min-max" normalization strategy, which retains the pattern of the data because the LSTM neural network needs the stock patterns during training:

$$X_t^n = \frac{X_t - \min\,(X_t)}{\max\,(X_t) - \min\,(X_t)} \tag{8}$$

Where X_t^n is the normalized data. De-normalization is therefore necessary when conclusion of the forecast process to obtain the original pricing, which is provided by

$$\hat{X}_t = \hat{X}_t^n[\max\,(X_t) - \min\,(X_t)] + \min\,(X_t) \tag{9}$$

Where $\hat{X}_t$ represents the anticipated data. The compound score is not standardized because it ranges at [-1, 1], which indicates that all of the scale for the compound score data is the same and does not need to be processed through normalization. Related results are shown in the Following Table 2, Table 3 and Table 4, respectively.

Table 2. Deviation results of different Dropout

	Training Points	Validation Points	Test Points
0.1 Dropout	0.00025767239	0.0006482924	0.00517316022
0.2 Dropout	0.00065637490	0.0011541662	0.0082821311
0.5 Dropout	0.00046212933	0.0013630003	0.00868191849

Table 3. The deviation outcomes from several LSTM layers

	Training Points	Validation Points	Test Points
3 Layers	0.00025767239	0.0006482924	0.00517316022
4 Layers	0.00079648674	0.0015846733	0.00606228038

Table 4. The deviation outcomes from several LSTM layers

	Training Points	Validation Points	Test Points
100 Units	0.00025767239	0.0006482924	0.00517316022
200 Units	0.00031754944	0.0006133969	0.00494382111

4.3. Predictions

In this part of article, the forecast results for the stocks of the four companies are shown in a line chart below. In the line chart, this paper shows the real value, the predicted value of ARIMA model and the predicted value of LSTM model all in one figure (See Fig. 2, Fig. 3, Fig 4, and Fig. 5)

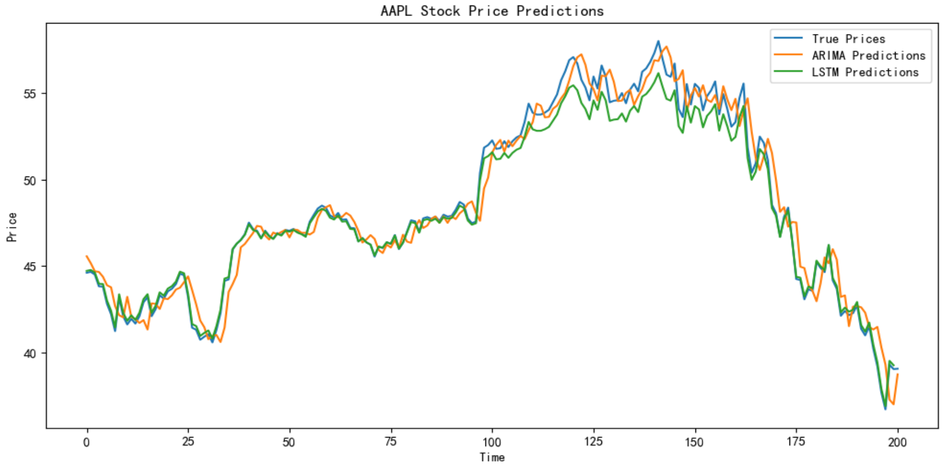

Fig. 2 Predictions of Apple

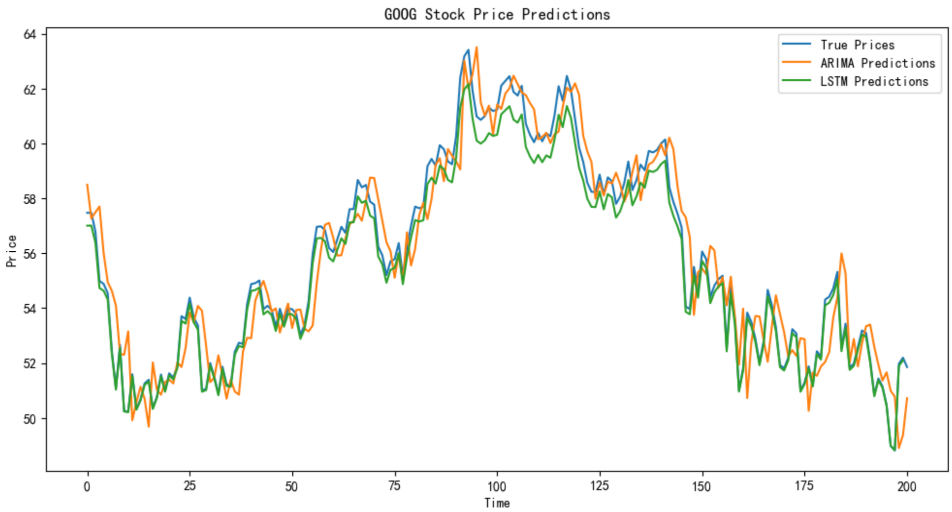

Fig. 3 Predictions of Google

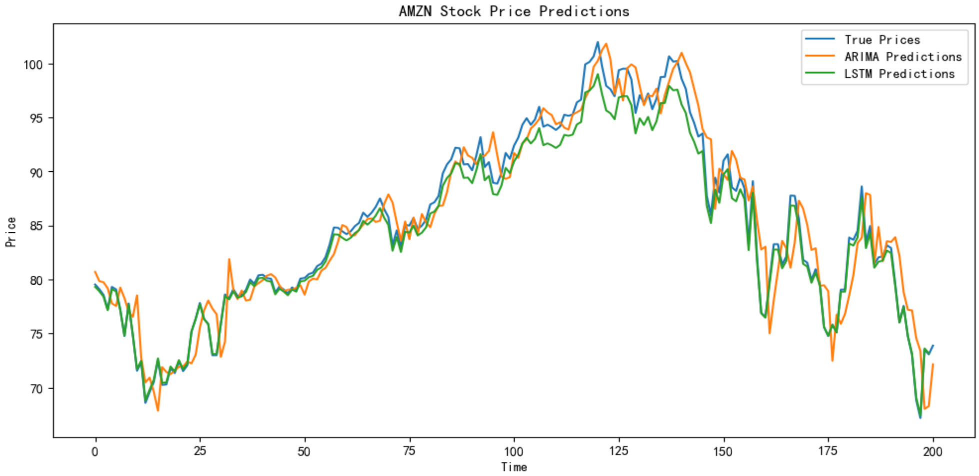

Fig. 4 Predictions of Amazon

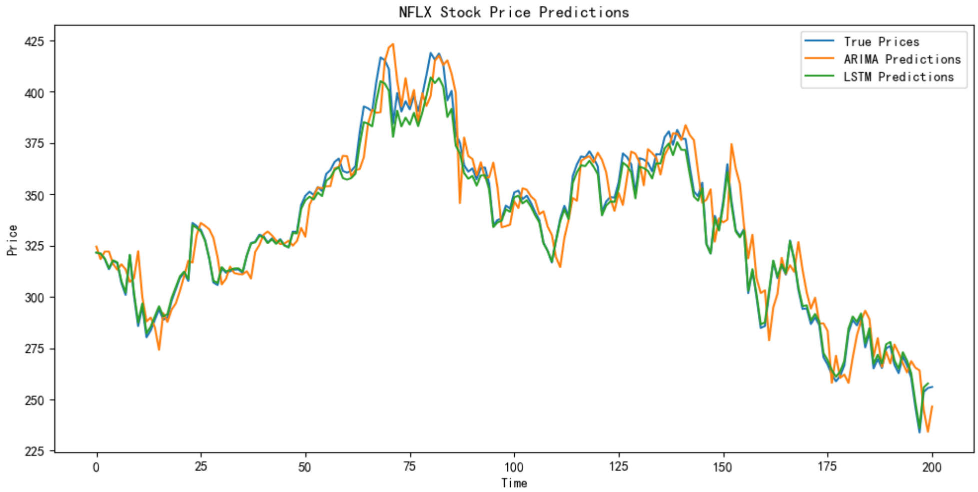

Fig. 5 Predictions of Netflix

4.4. Performance Evaluation

In this study, mean absolute error (MAE), mean square error (MSE), and root-mean square error (RMSE) are utilized, respectively, to evaluate the performance of the ARIMA model and the LSTM model (See Table 5).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2 \tag{10}$$

Table 5. MSE comparative of several models for various equities

	Apple	Google	Amazon	Netflix
ARIMA	1.17151	1.55678	6.38663	159.58281
LSTM	1.18464	1.07799	4.87162	97.59937

These indications show that LSTM performs better than LSTM in terms of the two models' ability to predict outcomes.

5. Conclusion

In this paper, 50 past stock prices are used to predict one future stock price. In the process of forecasting, ARIMA model and LSTM model are both used to predict the price after the data are checked for stationarity. After the estimation of parameters, both of the two models run well and give satisfactory results. As the basis of evaluation model, MSE is used to judge whether a model performs well. The outcomes demonstrate that the LSTM model performs better in predicting stock prices.

However, there are many limits in this research. For example, there are many time series model like ARCH and GARCH model. There are also many machine learning model like random forest model and RNN model. These models are not used in this study, and the prediction problem using other models is worthy of further study.

References

[1] Dong Li, Xiaohong SU, & Shunagyu Ma. Stock price prediction algorithm based on new dimension grey Markov model. 2002. Available at: 10.3321/j.issn:0367-6234.2003.02.029

[2] Yi Jiang, & Yongpeng Lin. The application of RBF neural network in stock price prediction. 2007. Available at: CNKI:SUN:XIZH.0.2007-04-005

[3] Lijun Zhang, & Di Yuan. Research on stock price prediction model based on GA-ELMAN dynamic regression neural network. 2008. Available at: 10.3969/j.issn.1007-5097.2008.09.018

[4] Jigang Zhang, & Na Liang. Stock price prediction based on som network - Principal component -BP network. 2008. Available at: CNKI:SUN:TJJC.0.2008-06-060

[5] Di Xu, Dajun Ma, & Yuanxi Li. Application of neural network in stock price prediction. 1998. Available at:10.3321/j.issn:1000-6788.1998.11.022

[6] Haibo Liu, & Dongyun Yi. Stock price prediction method based on wavelet analysis and fractal theory. 2007. Available at: CNKI:SUN:TJJC.0.2007-05-048

[7] Wen Long, Jiaqi Tian, & Yuanfeng Mao. Research on stock price prediction and trading strategy based on multi-level news. 2023. Available at: 10.3778/j.issn.1002-8331.2109-0388

[8] Rui Huang, Zijie Liu, & Ji Cui. CAE and GRU model based on K-line and moving average for stock price prediction. 2023. Available at: 10.12677/AAM.2023.121041

[9] Dejun Deng, Hongzhen Xu, & Shiyue Wei. Stock price prediction of E-V-ALSTM model. 2023. Available at: 10.3778/j.issn.1002-8331.2207-0482

[10] Yu Lin, Jinyuan Chang, & Yanyong Huang. Stock price prediction by combining empirical mode decomposition and deep time series model. 2022. Available at: 10.12011/SETP2021-3002