The Progress of Stock Price Forecasting Model Based on AI Techniques: ARIMA, Neural Networks and LSTM

Tianhao Gao[1, *, †] Yuwen Huo[2, *, †] Mengmeng Yu[3, *, †]

[*]Sgtgao2@liverpool.ac.uk[1], [*]1694680640@qq.com[2], [*]yumengmengaaaa@163.com[3]

Financial Mathematic, University of Liverpool, Liverpool, United Kingdom[1]
School of Economics and Management, North University of China, Taiyuan, China[2]
School of Finance, Nanjing University of Aeronautics and Astronautics, Nanjing, China[3]
[†]These authors contributed equally.

Abstract—In contemporary society, under great demand for effective investing, stock price forecasting is always a hot topic in financial research. On this basis, this study investigates the price forecasting models based on three common used measures and approaches, i.e., ARIMA, Neural networks and LSTM models regarding information retrieval and literature review methods. Specifically, for Artificial neural networks, we present a traditional three-layer model consisting of input, hidden and output layers with six variables. As for ARIMA, we mainly determine the main application scope of ARIMA through the brief comparison between ARIMA and other models in stock forecasting and the advantages and disadvantages of ARIMA. With regard to LSTM, we discuss the study of its various doors and functions and the analysis of its advantages and disadvantages. According to our analysis, among the three selected scenarios, LSTM shows pretty well performances among various approaches. Overall, these results shed light on future stock predicting model improvement.

Keywords-Price forecasting; ARIMA; machine learning; Neural Networks; LSTM.

1. INTRODUCTION

Before the advent of computers, the stock market analysis is generally separated as 2 types, i.e., the technical and fundamental analysis. The former one focused on analysing price time-series of underlying assets and volume time series to predict, evaluate, estimate and analysis future stock values. The fundamental analysis relied on the analysis of open source message, e.g., financial news and earnings reports [1]. The stock market prediction has entered the field of machine learning after computers developed, where traditional statistic models like ARIMA and machine learning (e.g., CNN, LSTM) have been proposed. Due to the development of modelling algorithms, e.g., ANN and GA, machine learning methods have been used to predict stock prices.

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2327116

Although the ARIMA has excellent performances with high accuracy, the factors it contained are limited. The implementation of machine learning models addresses this issue. Contemporarily, Machine learning, as one of the well-known technological methods, has been widely used in machine stock prices forecast. Patel and Yalamalle (2014) stated that ANN are being used in many fields because it is an irrefutably effective tool that helps the scientific community to predict possible outcomes [2].

The rest of the paper is organized as follow. Sec. 2 will introduce the basic description of ANN and its application in stock price prediction. Subsequently, Sec. 3 will discuss the performances of LSTM. Afterwards, the results of ARIMA are demonstrated and evaluated. Eventually, Sec. 5 gives a brief summary.

2. ARTIFICIAL NEURAL NETWORK

A neural network, also known as an ANN, comes from biological neural networks. Similar to a synapse in a biological brain, each connection can transmit signals to other neurons and process them accordingly. Neural networks have opened a new field of efficient and usable forecasting of financial products to optimize profits.

ANN is an intelligent data mining technique that identifies fundamental trends from huge data and generalizes them. Compared with most traditional methods, ANN can simulate and analyse complex changes in unstructured data. The input layer includes 6 variables: High-Low, Open-Close, and 7-day mean average, 14-day mean average, 21-day mean average, 7-day Standard Deviation and Volume [3]. Figure 1 illustrates a typical example of ANN

There are plenty of metrics and evaluation indicators for analysing the effectiveness of models, including RMSE, MAPE and MBE [4]. The description expression for them are given as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Oi-Fi)^2}{n}} \tag{1}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{(Oi-Fi)}{Oi} * 100 \tag{2}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(Oi - Fi) \tag{3}$$

where 'Oi' is the original closing price, 'Fi' is the predicted closing price, and 'n' is the total window size. MAPE is also used to evaluate the performance of the model and is calculated as

Other scholars demonstrate the original stock closing prices against the closing prices of the stocks of five different companies predicted using ANN [5]. Based on their analysis, ANN has ability for forecasting. According to the results, DNN can integrate news and fundamental factors, e.g., closing prices, trading volumes, and income statements to obtain better results in future work.

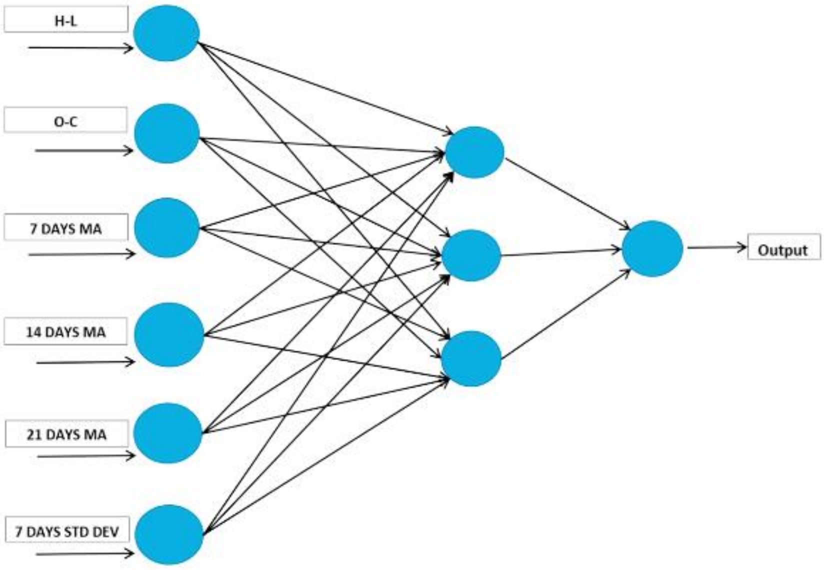

Figure 1. A sketch of ANN [4].

Through deeper research and optimization of the ANN model, Sahoo and Mohanty (2019), Chandar (2021) respectively used the Gray wolf optimization (GWO) technology and the Elman neural network (ENN) technology to predict the data of the BSE from August 25, 2004, to October 24, 2018, and the New York Stock Exchange and Nasdaq stock data. Both verified the accuracy and superiority of the ANN method through various metrics and comparing with a different model [5, 6].

3. THE MODEL OF LSTM

LSTM is a special kind of RNN that mainly solves the problem of gradient disappearance and gradient explosion in the process of long sequence training. There are many complicated financial factors ascribed to the complexity and stochastic of the stock market. However, with the technology developing, the opportunities to acquire steady fortune from the stock market is increasing, which is be conducive. Some of the important descriptions functions and expressions for LSTM are given as follows [7]:

$$h^t = f(h^{t-1}, x^t; \theta) \tag{4}$$

$$f_t = \sigma(w_f \cdot (h_{t-1}, x_t) + b_f) \tag{5}$$

$$i_t = \sigma(w_i \cdot (h_{t-1}, x_t) + b_i) \tag{6}$$

$$c_t = \tan h(w_c \cdot (h_{t-1}, x_t) + b_c) \tag{7}$$

$$o_t = \sigma(w_o(h_{t-1}, x_t) + b_o) \quad (8)$$

$$h_t = o_t * \tan h(c_t) \quad (9)$$

where x_t: input vector, h_t: output vector, c_t: cell state vector, f_t: forget gate vector, i_t: input gate vector o_t: output gate vector and W, b are the parameter matrix and vector [7]. LSTM introduces the memory cell, which can replace traditional ANN in the network's hidden layer. According to these memory cells, the net can connect with memory and input remote effectively and timely. Therefore, it is suitable to grasp the data structure [8, 9]. A typical visualization is given in Fig. 2.

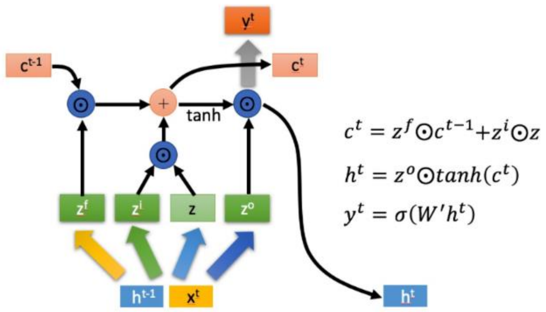

Figure 2. A sketch of LSTM [10]

In the financial field, most of the applications of the ISTM model are for stock prediction. The usual steps of LSTM neuron and network design are to set the number of input neurons to the number of input features, output neurons corresponding to the final predicted closing price, initial chemical rate and network weights of forgetting gate, updating gate and output gate [11]. To analyse the system's efficiency, RMSE is widely used to minimize the error or difference. It can be seen from the experiment that the prediction value of stock closing price can be more accurate and stable for regions with data-intensive or obvious trends under the LSTM model. The prediction effect is limited for the wave crest and trough regions with large fluctuations and the turning point. In the experimental process, the deviation value can improve its performance.

LSTM solves the problem of gradient disappearance and gradient explosion in the process of gradient backpropagation, which is very suitable for dealing with the problem of high correlation with time series. For the stock prediction problem, by adjusting the specific parameters of LSTM, such as the number of nerves and the number of network layers, the prediction accuracy reaches a high level in theory. LSTM network training time is too long, so that the actual operation rate is extremely slow. According to experimental results, LSTM is difficult to show its advantages in the normal network construction and operation environment.

4. ARIMA

An ARIMA (p, d, Q), AR is autoregression, p is an autoregression term, MA is a moving average, Q is a moving average, d is the difference (order). It is a time series data affected by many factors, such as company operating situation, financial situation, macroeconomic situation and institutional policy [11]. However, the ARMA model only considers the characteristics of the time series itself to predict and does not consider other factors that affect the change of stock price, i.e., it chooses the time series model, which eliminates other factors. The model results predict Nokia stock price in the short term, and the effect is very good and acceptable [12].

The ARIMA model is viewed from the perspective of the statistical model. There are literature reports that can be predicted from two perspectives: statistics and AI technology. The full name of the ARIMA model is the autoregressive moving average model, which is a commonly used statistical model used to predict time series. It is usually recorded as ARIMA (p, d, q): where p is the number of lags of the data itself and the parameters in the AR model, the autoregressive model. d is that time-series data requires several differences to obtain stable data. q is the number of lags in the prediction error and the parameters in the MA model, namely the sliding averaging model.

ARIMA is a financial analysis and prediction model, using ARIMA based on past values of the series, coupled with previous partial errors, to predict stock motion. It is more accurate and efficient than those that use more complex structural models for prediction. ARIMA assumes that the digital columns are linearly generated, so real-world prediction problems may not apply. While another predictive model, ANNs, as soft computing technology, is the most commonly used prediction model in the fields of society, engineering, economy, business, finance, foreign exchange, stock, etc. [13]. First, one obtains the time series data; second, checks whether the observation data is stable and whether the time series data is relatively stable. There is basically no certain rise or downward trend. Suppose the instability can be stabilized differentially. If unstable, converted to stable timing data, q and p, determined by observing the autocorrelation coefficient ACF generally use partial autocorrelation coefficient PACF view software. However, Nonlinear relationships ARIMA does not handle very well, i.e., ARIMA can only handle linear relationships. When using it, time data after differential differentiation requires stability or stability, i.e., the model cannot predict well when unstable. Meanwhile, ARIMA is not used for price forecasting and requires additional model assistance. Besides, in the prediction inventory, ARIMA outperforms other models used primarily for linear time series prediction and works for short-term directional trends. ARIMA is more robust and more effective in financial time series prediction. To determine the best choice in practising the ARIMA model, one needs to use it right below.

The ARIMA (p, d, q) model of the Nokia stock index illustrates how best can be found in practice ARIMA's model [14]. The Nokia stock data used in this study covered 25 periods from April to 25 February 2011, totalling 3,990 observations. Fig. 3 describes the original pattern. The model was tested using the enhanced Dickey Fuller (A D F) unit root test of the Nokia stock index "DCLOSE". It is confirmed that the sequence becomes stationary after the first difference of the sequence. If the model performs well, the model's residual value is a series of stochastic errors. Therefore, it is unnecessary to consider any AR (p) and MA (q) further. Finally, in the prediction form, the best model is:

$$Y_t = \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \tag{10}$$

Since ACF and PACFs have no significant peaks, comparing the predicted prices to the horizontal prices to the actual stock prices, we finally verify that this is the best ARIMA model in multiple experiments.

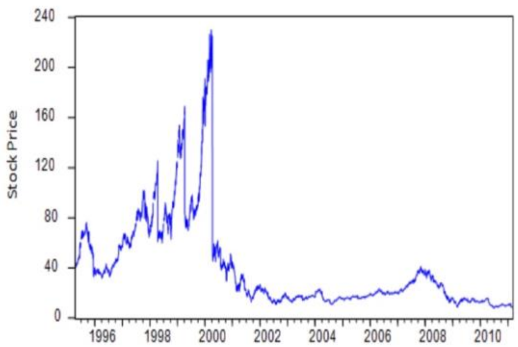

Figure 3. Graphical representation of a specific underlying assets.

5. Conclusion

In summary, we discuss the feasibility and performances of different stock price forecasting models. Specifically, the ANN, LSTM and ARIMA models are selected as investigation targets. According to our analysis, all the models show excellent performances, at least in some cases, though they also have drawbacks needed to be improved. In general, the ANN model has been developed over a long period and has become accurate in predicting stock prices. Hence, it is believed that the ANN model will have more room for development in the future. As for the ARIMA model has the potential to predict stock prices in the short term. Based on the obtained results, the ARIMA model can compete fairly with the emerging prediction technology in short-term prediction. Using ARIMA is based on past values in the series to predict inventory movement, combined with past errors. It is more accurate and effective than those that use more complex structural models for prediction. Regarding LSTM, it is very suitable for dealing with the problem of high correlation with time series. Stock data is in good agreement with the LSTM model. Compared with the other two models, LSTM can better predict the stock trend. This forecasting technique is helpful for investors and anyone. A prediction model with high accuracy is needed in the future. All in all, these offer a guideline for future stock forecasting model development.

References

[1] Faten Subhi Alzazah and Xiaochun Cheng, 'Recent Advances in Stock Market Prediction Using Text Mining: A Survey', DOI: 10.5772/intechopen.92253, https://www.intechopen.com/chapters/72381

[2] Mayankkumar B Patel, Sunil R Yalamalle, 'Stock Price Prediction Using Artificial Neural Network', International Journal of Innovative Research in Science, Engineering and Technology, ISSN: 2319-8753 vol. 3, Issue 6, June 2014
[3] Mehar Vijh, Deeksha Chandola, Vinay Anand Tikkiwal, Arun Kumar, Stock Closing Price Prediction using Machine Learning Techniques, Procedia Computer Science, Volume 167, 2020, Pages 599-606, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.326. (https://www.sciencedirect.com/science/article/pii/S1877050920307924)
[4] Sahoo S., Mohanty M.N. (2020) Stock Market Price Prediction Employing Artificial Neural Network Optimized by Gray Wolf Optimization. In: Patnaik S., Ip A., Tavana M., Jain V. (eds) New Paradigm in Decision Science and Management. Advances in Intelligent Systems and Computing, vol 1005. Springer, Singapore. https://doi.org/10.1007/978-981-13-9330-3_8
[5] Kumar Chandar, S. Grey Wolf optimization-Elman neural network model for stock price prediction. Soft Comput 25, 649–658 (2021). https://doi.org/10.1007/s00500-020-05174-2
[6] https://www.researchgate.net/publication/321503983_Stock_price_prediction_using_LSTM_RNN_and_CNN-sliding_window_model
[7] https://www.researchgate.net/publication/327967988_Predicting_Stock_Prices_Using_LSTM
[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997
[9] https://zhuanlan.zhihu.com/p/32085405
[10] Ding Wenjuan. Comparison of ARIMA model and LSTM model based on stock prediction [J]. Industrial control computer, 2021,34 (07): 109-112 + 116
[11] https://mo.mbd.baidu.com/r/rJj0aFkF8Y?f=cp&u=b42f0d14ecfa4c47
[12] https://mo.mbd.baidu.com/r/rJj6bD7ccg?f=cp&u=f8adeae3c4115c79
[13] http://dx.doi.org/10.1 155/2014/614342
[14] 78- 1-4799-4923-6/14 $31.00 O 2014 IEEE DOI 10.1 109/UKSim.2014.67