ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

Time Series Forecasting Model Based on Mayfly Algorithm Optimization

Yi Chen[1,a,*], Eva Khong[1,b]

[1]*Faculty of Finance, City University of Macau, Macau, China*
[a]*f21092100171@cityu.mo*, [b]*evakhong@cityu.mo*

Abstract: *In order to verify that the model optimized by the mayfly algorithm improves the prediction ability of the stock index, this paper will compare and analyse the prediction results of a single algorithm model as the benchmark model and the model after the optimization of the mayfly algorithm. Therefore, this paper will show the comparative analysis of the prediction results of the optimization algorithm model based on the mayfly algorithm. This paper employs mayfly algorithm (MA), which is a newly developed algorithm recently, to optimize the algorithm models back propagation (BP), extreme learning machine (ELM), and kernel-based extreme learning (KELM) and the long-short term memory (LSTM) model for parameter optimization respectively. These algorithm after optimized would be show that is MA-BPNN model, MA-KELM model, MA-KELM model and MA-LSTM model respectively. From comparative analysis shows that the BPNN model is the better forecasting effect of a single algorithm model, and the model ELM optimized by the mayfly algorithm shows better forecasting ability. Additionally, the prediction ability of LSTM model is better without optimization, compared with itself after optimized by mayfly algorithm. We found that the model after algorithm optimization will basically greatly improve its prediction ability, and the specific performance is that its prediction error will be reduced.*

Keywords: *Mayfly algorithm, Algorithm optimization, Prediction error*

1. Introduction

In academia, methods used by scholars to predict financial time series are mainly divided into statistics and artificial intelligence algorithms [1], which is concluded by linearity and nonlinearity method respectively. Traditionally, the usage of classic prediction tool of financial time series based on certain strict assumptions in the process of researching time series, and the parts that do not meet the assumptions requirement are regarded as interference factors, defined by linearity method [2]. For instance, autoregressive model (AR model), moving average model (MA model), autoregressive moving average model (ARMA model), autoregressive integrated moving average model (ARIMA model), which are frequently used for linear time series forecasting analysis models[3],assuming that the financial time series is fixed and follows a normal distribution [4,5].ARMA model consists of purely AR model and purely MA model, commonly depicting the linearity structure between lagged variables, which is one of classic method without differencing process. The differencing process is proposed by autoregressive integrated moving average (ARIMA) models, is a generalization of an Autoregressive Moving Average (ARMA) model [6], which time series is significantly nonstationary. Additionally, the linearity model mentioned above serves as vehicles for linear forecasting [3]. However, the forecasting performance of these models in financial markets is unsatisfactory effect, surmising that useful information in time series loss somewhat after differencing probably.

Compared with establishing requirement of classic approaches under certain strict assumptions, a data-driven self-adaptive method as an attractive alternative tool to introduced by scholars for improvement of prediction accuracy for financial time series, that is intelligent models, providing flexible approach for financial time series prediction [7], such as smart algorithms or smart algorithms hybrid model provide better predictive performance for nonlinear data, especially for the hybrid methods that predictive power significantly better than single-item[2].

From the above analysis of relevant literature, we can see that traditional time series prediction analysis has its limitations. For time series prediction, the academic world is mainly divided into two categories, that is mainly traditional econometric prediction and algorithmic forecasting models. The former prediction method is based on the setting of some parameters, which sometimes causes a certain

ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

prediction error in the prediction of the sequence, while the latter algorithmic prediction is a popular prediction model in recent years. The algorithm model can preserve the data information to a certain extent based on its internal structural design, so that the error can be reduced when the output is predicted.

In recent years, there is a trend that algorithm models have begun to introduce time series prediction research. The overall characteristics of stock financial time series are volatility, noise, and nonlinearity. For the prediction and analysis of this time series, it is necessary to find a suitable model for prediction and analysis. Thus, this research has its significance. This research based on the nonlinear and noisy characteristics of financial time series, we consider to make a comparative analysis of mayfly algorithm optimization algorithm to find a suitable model to predict financial time series, using Mean-Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) as predictive effect evaluate indicators. In this paper, we choose mayfly algorithm (MA), which is proposed in 2020 and it has advantage of good optimization performance and fast convergence speed [33]. Therefore, it is selected to optimize the BP, ELM, KELM and LSTM models respectively. Therefore, a suitable algorithm model optimized by the mayfly algorithm with better prediction effect could be found.

The common models of traditional research time series are autoregressive models (AR model), moving average model (MA model), and autoregressive moving average model (ARMA model). Compared with traditional models, the currently commonly used algorithm models for time series prediction include BP neural network, extreme learning machine, kernel extreme learning machine, LSTM model, etc. These models have advantages in processing time series with nonlinear characteristics when analyzing time series, especially in forecasting. For example, in time forecasting, algorithmic models can improve their forecasting accuracy. But a single algorithm model has certain drawbacks. For example, the training of the BP neural network model is time-consuming, which will affect the efficiency of the compensation implemented. The parameters of a single LSTM model have many uncertainties, such as the number of hidden layers, the number of neurons in the hidden layer, the learning rate, and the number of iterations. Compared with the first two single algorithm models, the ELM model has the advantages of automatic parameter setting, efficient learning speed, and good generalization performance. Pre- and post-optimization comparisons with other common algorithm models.

This paper will present the newer mayfly algorithm (MA) and back propagation (BP), extreme learning machine (ELM), kernel extreme learning machine (KELM)and long-short term memory (LSTM) model are combined into MA-BPNN model, MA-KELM model, MA-KELM model and MA-LSTM model respectively For financial time series financial forecasting, in order to verify the predictive ability of the stock index improved by the modulo optimized by the mayfly algorithm, this paper presents the BPNN, ELM, KELM and LSTM models as benchmark models for comparative research.

2. Methodology

2.1. Mayfly algorithm (MA)

The mayfly algorithm was proposed by Zervoudakis and Tsafarakis in 2020. The principle of the algorithm is based on the flight behaviour and mating behaviour of mayflies. The algorithm has advantage of good optimization performance and fast convergence speed. Since the principle of the algorithm originates from the mating behaviour of mayflies, it is designed according to the behavioural characteristics of mayflies.

2.2. Back propagation (BP)

The BP arithmetic, which was put forward by several researchers [28], primarily comprises 2 components, that is, the forward signal transmission and the reversed error dissemination respectively. It primarily comprises 3 components: input tier, multiple hidden tiers, and output tier. It's a monitored training arithmetic with potent self-adaptation, self-training, non-linearity mapping abilities, and can better tackle the problems of insufficient data, poorer informational contents, and uncertainties. It's not restricted by non-linear modeling methods.

The input signal transmits from the input tier to the output tier across every hidden tier, and the real responsive result is acquired at the output tier during forward dissemination.

$$I_j = \sum_i w_{ij} O_i + \theta_j$$

ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

In which w_{ij} denotes the connection weight with the previous tier neuron i and unit j; O_i denotes the output of unit i, and θ_j denotes the bias of unit j, which is utilized as a liminal value to alter the unity activity.

The connection weights and liminal values of every neuron are persistently modified from the output tier across every hidden tier as per the approach of error gradient descent, and repeated the iterative process till the error of the neural networks output is decreased to a satisfactory degree or proceed to a preset number of training times during back dissemination.

For the learning tuple (x_k, y_k), we assume that the output of neural networks is $\widehat{y_k}=(\hat{y}_1^k, \hat{y}_2^k, \ldots, \hat{y}_l^k)$, then the mean squared error of the neuronetwork on (x_k, y_k) is:

$$E_k = \frac{1}{2}\sum_{j=1}^{l}(\hat{y}_1^k - y_j^k)^2$$

$$minE_k = \frac{1}{2}\sum_{j=1}^{l}(\hat{y}_1^k - y_j^k)^2$$

In which $\hat{y}_1^k$ denotes the learning group value fitting the model, y_j^k denotes the testing group observation result.

2.3. Kernel-based extreme learning machine (KELM)

ELM is featured by its rapid training and potent generalization abilities [29], which is an easy training approach of single hidden layer feedforward networks (SLFNs) [30]. Such arithmetic runs under a learning set following 3 steps:

Step 1: Action authors can stochastically assign input weights and biases appropriate

Step 2: The operator computes the hidden tier and output the matrix

Step 3: Compute the outcome of output weights

2.4. Long-short term memory (LSTM)

LSTM pertains to one kind of Recurrent Neural Network (RNN) [31]. Such neuronetwork has a certain memory capability, particularly such capability is displayed in long-term and short-term memory. Such modeling method is often utilized in time-series analyses. Its applications in prediction analyses, financial time series analyses, etc. To be specific, because of the beneficial modeling architecture, the structure features of the modeling method, i.e., easy unit comprises units with input gates, output gates and forgetting gates. When data stream across every gate, in which desirable results are collected and preserved [9,32].

3. Evaluation indicators of prediction results

In this paper, four evaluation indicators, MSE (Mean-Square Error), RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error), are selected to test the prediction effects of different models. The calculation formulas are:

$$MAE = 1/n\sum_{i=1}^{n}|Y_i - \widehat{Y_i}|$$

$$MAPE = 1/n\sum_{i=1}^{n}|Y_i - \widehat{Y_i}|/Y_i$$

$$MSE = [1/n\sum_{i=1}^{n}(Y_i - \widehat{Y_i})]^2$$

$$RMSE = \sqrt{[1/n\sum_{i=1}^{n}(Y_i - \widehat{Y_i})]^2}$$

Where Y_i represents the observation value, $\widehat{Y_i}$ represents the prediction value.

4. Experimental results and discussions

The empirical time series of the model in this paper is the Shanghai Composite Index, one of the representatives of the broader market index. The Shanghai Composite Index contains a total of 1,459 transaction data from January 4, 2016 to December 31, 2021. The data comes from Yahoo Finance.

ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

In order to verify whether the accuracy of the time series prediction results is improved after the mayfly algorithm optimizes the model, this paper introduces BPNN, ELM, KELM and LSTM as benchmarks. Therefore, Table 1 and Table 2 respectively give the prediction results of a single model and a single model after the mayfly algorithm. The prediction results of the optimized model are compared.

Table 1: Predictive performance comparison of single model

Model	MAE	MSE	RMSE	MAPE
BPNN	**35.6573**	**1823.95**	**42.7078**	**1.01%**
ELM	65.6454	6062.86	77.8644	1.84%
KELM	39.8189	2271.66	47.6619	1.12%
LSTM	45.1944	2952.47	54.3367	1.27%

Table 2: Predictive performance comparison of optimization model

Model	MAE	MSE	RMSE	MAPE
MA-BPNN	28.0186	1179.07	34.3376	0.80%
MA-ELM	23.8829	925.976	30.4299	0.68%
MA-KELM	27.977	1181.75	34.3765	0.79%
MA-LSTM	81.5884	8549.68	92.4645	2.30%

From the analysis results of the single algorithm model of Table 1, it can be seen from Table 1 that the BPNN model is dominant in the prediction compared to the ELM model, the KELM model and the LSTM model. Specifically, you can see the prediction evaluation indicators of the BPNN model. MAE, MSE, RMSE, and MAPE are 35.6573, 1823.952, 42.7078, and 1.0078%, respectively. This obviously reflects the advantage of a single algorithm model prediction, and its error indictor result is relatively small compared to other models, indicating that the BPNN model prediction performance is satisfied among these several single algorithm models.

We can see the prediction results of the algorithm model optimized by the mayfly algorithm given by Table 2. From the result given by Table 2, the MA-ELM model is dominant in the prediction compared to MA-BPNN, MA-KELM and MA-LSTM. From Table 2, we can see that the prediction evaluation indicators MAE, MSE, RMSE and MAPE of the MA-KELM model are 23.8829, 925.9763, 30.4299 and 0.68143%, respectively. From the comparison given by the error indicators result, the ELM model optimized by the mayfly algorithm prediction performance is satisfied. In addition, it can be found that after the individual algorithm models BPNN, ELM, and KELM are optimized by the mayfly algorithm, the value of their prediction error index is smaller than before, which means that their prediction The effect is improved. However, the LSTM model is an exception. The prediction ability of the LSTM model without algorithm optimization is better than that after algorithm optimization.

5. Conclusion

From the above empirical analysis, this paper uses the newer mayfly algorithm (MA), Back Propagation (BP), extreme learning machine (ELM), kernel extreme learning machine KELM (Kernel-based extreme learning machine (KELM)) and long-short term memory (LSTM) models are combined into MA-BPNN model, MA-KELM model, MA-KELM model and MA-LSTM model respectively for financial time series forecasting test and make comparative analysis. This paper draws the conclusion that the BPNN model is the better forecasting effect of a single algorithm model, and the model ELM optimized by the mayfly algorithm has better forecasting ability. In addition, the prediction ability of LSTM model is better before it has been optimized, compared with itself after optimized by mayfly algorithm. One thing is for sure that, the model after algorithm optimization will basically greatly improve its prediction ability, and the specific performance is that its prediction error will be reduced.

References

[1] Lin, Y., Yan, Y., Xu, J., Liao, Y., & Ma, F. (2021). Forecasting stock index price using the CEEMDAN-LSTM model. The North American Journal of Economics And Finance, 57, 101421. https://doi.org/ 10.1016/j.najef. 2021.101421

[2] Lin, H., Sun, Q., & Chen, S. (2020). Reducing Exchange Rate Risks in International Trade: A Hybrid Forecasting Approach of CEEMDAN and Multilayer LSTM. Sustainability, 12(6), 2451. https://doi.org/ 10.3390/su12062451

ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

[3] Fan, J., & Yao, Q. (2008). Nonlinear time series: nonparametric and parametric methods. Springer Science & Business Media.

[4] Lin, H., & Sun, Q. (2020). Crude Oil Prices Forecasting: An Approach of Using CEEMDAN-Based Multi-Layer Gated Recurrent Unit Networks. Energies, 13(7), 1543. https: //doi.org/ 10.3390/ en13071543

[5] Chen, L., Chi, Y., Guan, Y., & Fan, J. (2019, May). A hybrid attention-based EMD-LSTM model for financial time series prediction. In 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 113-118). IEEE.

[6] Riesgo García, M., Krzemień, A., Manzanedo del Campo, M., Escanciano García-Miranda, C., & Sánchez Lasheras, F. (2018). Rare earth elements price forecasting by means of transgenic time series developed with ARIMA models. Resources Policy, 59, 95-102. https://doi.org/10.1016/ j.resourpol. 2018.06.003

[7] Zhang, G., Eddy Patuwo, B., & Y. Hu, M. (1998). Forecasting with artificial neural networks:. International Journal of Forecasting, 14(1), 35-62. https://doi.org/10.1016/s0169-2070(97) 00044-7

[8] Lv, P., Wu, Q., Xu, J., & Shu, Y. (2022). Stock Index Prediction Based on Time Series Decomposition and Hybrid Model. Entropy, 24(2), 146. https://doi.org/10.3390/e24020146

[9] Sezer, O., Gudelek, M., & Ozbayoglu, A. (2020). Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181. https://doi.org/10. 1016/j.asoc.2020.106181

[10] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). IEEE.

[11] Karmiani, D., Kazi, R., Nambisan, A., Shah, A., & Kamble, V. (2019, February). Comparison of predictive algorithms: backpropagation, SVM, LSTM and Kalman Filter for stock market. In 2019 Amity International Conference on Artificial Intelligence (AICAI) (pp. 228-234). IEEE.

[12] Makala, D., & Li, Z. (2021). Prediction of gold price with ARIMA and SVM. Journal Of Physics: Conference Series, 1767(1), 012022. https://doi.org/10.1088/1742-6596/1767/1/012022

[13] Kumar, M., & Thenmozhi, M. (2014). Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models. International Journal of Banking, Accounting And Finance, 5(3), 284. https://doi.org/10.1504/ijbaaf.2014.064307

[14] Büyükşahin, Ü., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing, 361, 151-163. https://doi.org/10.1016/j.neucom.2019.05.099

[15] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669. https://doi.org /10.1016/j.ejor.2017.11.054

[16] Ding, G., & Qin, L. (2019). Study on the prediction of stock price based on the associated network model of LSTM. International Journal of Machine Learning And Cybernetics, 11(6), 1307-1317. https://doi.org/10.1007/s13042-019-01041-1

[17] Qin, J., Tao, Z., Huang, S., & Gupta, G. (2021, March). Stock price forecast based on ARIMA model and BP neural network model. In 2021 IEEE 2nd International Conference on Big Data,

[18] Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536. https://doi.org/10.1038/323533a0

[19] Wu, D., Wang, X., & Wu, S. (2021). A Hybrid Method Based on Extreme Learning Machine and Wavelet Transform Denoising for Stock Prediction. Entropy, 23(4), 440. https://doi.org/ 10.3390/ e23040440

[20] Wang, F., Zhang, Y., Rao, Q., Li, K., & Zhang, H. (2016). Exploring mutual information-based sentimental analysis with kernel-based extreme learning machine for stock prediction. Soft Computing, 21(12), 3193-3205. https://doi.org/10.1007/s00500-015-2003-z

[21] Das, S., & Padhy, S. (2015). A novel hybrid model using teaching–learning-based optimization and a support vector machine for commodity futures index forecasting. International Journal of Machine Learning And Cybernetics, 9(1), 97-111. https://doi.org/10.1007/s13042-015-0359-0

[22] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLOS ONE, 12(7), e0180944. https://doi.org/ 10.1371/journal.pone.0180944

[23] Büyükşahin, Ü., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing, 361, 151-163. https://doi.org/10.1016/j.neucom. 2019.05.099

[24] Liu, M., Ding, L., & Bai, Y. (2021). Application of hybrid model based on empirical mode

ISSN 2616-5902 Vol. 4, Issue 10: 94-99, DOI: 10.25236/AJBM.2022.041016

decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction. Energy Conversion And Management, 233, 113917. https://doi.org/10.1016/j.enconman. 2021.113917

[25] Zhou, F., Zhou, H., Yang, Z., & Yang, L. (2019). EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. Expert Systems With Applications, 115, 136-151. https://doi.org/10.1016/j.eswa.2018.07.065

[26] Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., & Zheng, Q. et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings Of The Royal Society Of London. Series A: Mathematical, Physical and Engineering Sciences, 454(1971), 903-995. https://doi.org/10.1098/rspa.1998.0193

[27] Zhou, Y., Li, T., Shi, J., & Qian, Z. (2019). A CEEMDAN and XGBOOST-Based Approach to Forecast Crude Oil Prices. Complexity, 2019, 1-15. https://doi.org/10.1155/2019/4392785

[28] Witaszek, J. (1995). Backpropagation: Theory, architectures, applications. Neurocomputing, 9(3), 358-359. https://doi.org/10.1016/0925-2312(95)90002-0

[29] Yang, Z., Ce, L., & Lian, L. (2017). Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. Applied Energy, 190, 291-305. https://doi.org/10.1016/j.apenergy.2016.12.130

[30] Huang, G., Zhu, Q., & Siew, C. (2006). Extreme learning machine: Theory and applications. Neurocomputing, 70(1-3), 489-501. https://doi.org/10.1016/j.neucom.2005.12.126

[31] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

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

[33] Zervoudakis, K., & Tsafarakis, S. (2020). A mayfly optimization algorithm. Computers &Amp; Industrial Engineering, 145, 106559. https://doi.org/10.1016/j.cie.2020.106559