An Optimized Hybrid Deep Learning Model with Dung Beetle Optimizer for Stock Price Prediction

Yafei Wu

2021410055@sdtbu.edu.cn

Shandong Technology and Business University, Yantai, Shandong, China

Abstract. Stock prices are known to vary nonlinearly, which makes stock price forecasting quite difficult. Therefore linear models cannot accurately predict frequently fluctuating stock prices; instead, nonlinear models such as gated recurrent unit (GRU) and temporal convolutional network (TCN) tend to outperform linear models in stock price prediction. And yet, nonlinear models that are not well optimized to forecast unstable stock data generally result in poor fit and instability problems. To improve the fitting and stability of a model for share price forecasting, it is essential to optimize the parameters of a model. Dung beetle optimizer (DBO) is a novel bee colony intelligent optimization algorithm that can be used to optimize nonlinear models to enhance the precision of stock price forecasting. In the paper, we first present the GRU-TCN hybrid model with the dung beetle optimization algorithm for stock price prediction, which alleviates the poor fitting and instability problems of stock price prediction to a certain extent and improves the accuracy of stock price prediction. We conducted extensive experiments to show that GRU-TCN-DBO has better performance on MSE and $\boldsymbol{R^2}$ evaluation metrics compared to GRU-TCN, GRU, TCN, and LSTM by using 30 stocks of the Dow Jones Industrial Average.

Keywords: hybrid model; dung beetle optimizer; GRU-TCN; stock price prediction.

1 Introduction

Stock markets improve the efficiency of the use of social capital and play a crucial role in the whole society's advancement. Stock prices are correctly predicted in the stock market, which can help investors make the right decisions to minimize risk and thus gain a return on their investment. However, it is challenging to precisely forecast stock prices because stock prices are affected by international markets, human manipulation, political factors, and future policies leading to frequent fluctuations in stock prices.

Stock price forecasting is a canonical question at the intersection of finance and computers, and many experts in this field have created various models to increase the precision of stock price forecasts. In general, stock price prediction models can be categorized into linear and nonlinear models. For linear models, they are able to forecast smooth data with high accuracy [1][2][3]. However, linear models predict sharply fluctuating stock price data with lower forecasting accuracy. For nonlinear models, they can be classified into conventional machine learning models and deep learning models. For conventional machine learning models, it can improve the accuracy of stock price prediction to some extent [4][5][6].

ICPDI 2023, September 01-03, Chongqing, People's Republic of China

DOI 10.4108/eai.1-9-2023.2338789

Deep learning is a new research area in machine learning that has attracted much attention. Deep learning models could increase the precision of stock price predictions even further compared to conventional machine learning models. Zaheer[7] et al. proposed a deep hybrid model CNN-LSTM-RNN and compared it with RNN[8], LSTM[9], CNN-LSTM[10], and so on, and the results of the experiments proved that the presented model is superior to other models. Chen[11] et al. proposed a share forecasting model with a gated recursive unit (GRU) and reorganized dataset to solve the overfitting problem and improve the forecasting accuracy. Yuan[12] et al. presented a hybrid stock index forecasting model to improve forecasting accuracy, which is built on a multi-variate empirical modal division and temporal convolutional network (TCN). But they did not investigate how to combine GRU and TCN for stock price prediction and optimization of deep learning models to raise the forecasting accuracy. For this reason, we propose for the first time a hybrid model GRU-TCN for stock price prediction and optimize it using the dung beetle optimization (DBO) [13] algorithm.

The major contributions of this research are as follows.

a) We first use the dung beetle optimization algorithm for hybrid model GRU-TCN with and apply it to stock price prediction.

b) We conduct numerous experiments to demonstrate that the present hybrid model in the paper has high forecasting precision and fitting ability compared with other models.

The next part of this paper is arranged below. We present the research methods in Section 2 and describe the experiments in Section 3. Finally, we will summarize our work in Section 4.

2 Methods

2.1 Gated recursive unit (GRU)

A recurrent neural network is slightly unlike a conventional neural network in that it comes with a loop to itself, indicating that the information it is currently processing can be passed on to be used at the next point in time. GRU is a variant of the recurrent neural network, which solves the problems of not being able to memorize for any length of time in the recurrent neural network and the gradient in the back-propagation, and it is similar to the role of the LSTM, but it is simpler than the LSTM, which is easy to train.

The formula for GRU is derived as follows:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{1}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{2}$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \tag{3}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{4}$$

where x_t is the input information at the current moment and h_{t-1} is the hidden state at the previous moment. h_t is the hidden state passed to the next moment and $\tilde{h}_t$ is the candidate hidden state. r_t is the reset gate and z_t is the update gate. σ is the sigmoid function through which the data can be changed to the range 0-1, and $tanh$ is the $tanh$ function through which the data can change the data to a value in the range [-1,1].

2.2 Temporal convolutional network (TCN)

The TCN model (**Fig. 1.** Diagram of TCN model structure.(**Fig. 1.** a) is based on the CNN model with the below enhancements:

1) Applicable Sequence Model: Causal Convolution.

2) Dilated Convolution, Residual block (**Fig. 1.** b).

Causal convolution can handle sequential problems. The sequential problem can be transformed into: predict $y_1, y_2, \dots, y_t$ based on $x_1, x_2, \dots, x_t$. Here is the definition of Causal Convolution at x_t:

$$(\mathrm{F} * \mathrm{X})_{(x_t)} = \sum_{k=1}^{K} f_k \, x_{t-K+k} \tag{5}$$

where filter $\mathrm{F} = (f_1, f_2, \dots, f_t)$ and sequence $X = (x_1, x_2, \dots, x_T)$.

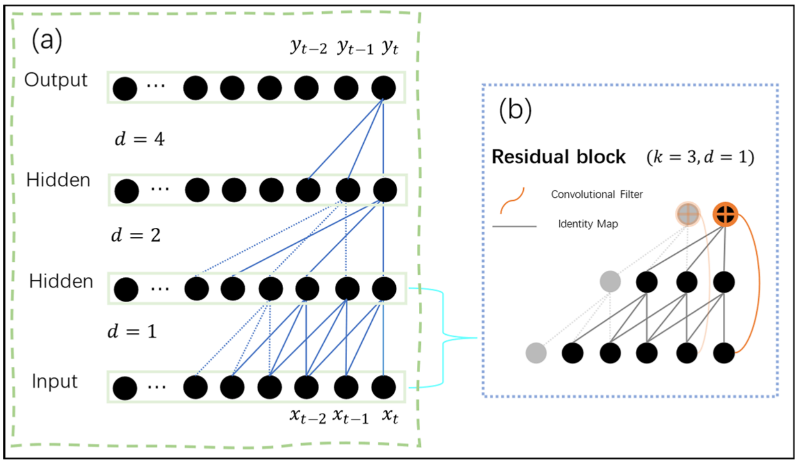

Fig. 1. Diagram of TCN model structure.

Standard CNN obtain a larger sensory field by increasing the pooling layer, which is always followed by a loss of information. Dilated Convolution increases the receptive field by injecting holes into the standard convolution. The definition of the x_t unfolding convolution is given below:

$$(\mathrm{F} *_d \mathrm{X})_{(x_t)} = \sum_{k=1}^{K} f_k \, x_{t-(K-k)d} \tag{6}$$

Where filter $\mathrm{F} = (f_1, f_2, \dots, f_t)$ and sequence $X = (x_1, x_2, \dots, x_T)$.

2.3 Dung beetle optimizer (DBO)

The dung beetle optimization algorithm focuses on simulating the ball-rolling, ball-dancing, food-finding, stealing, and reproductive actions of dung beetles. The specific formulas are given below:

$$x_i(t+1) = x_i(t) + \alpha k x_i(t-1) + b\Delta x,$$

$$\Delta x = |x_i(t) - X^w| \tag{7}$$

$$x_i(t+1) = x_i(t) + \tan(\theta)\,|x_i(t) - x_i(t-1)|,$$
$$\theta \in [0, \pi] \tag{8}$$

$$Lb^* = \max(X^*(1-R), Lb),$$
$$Ub^* = \max(X^*(1+R), Ub),$$
$$R = 1 - \frac{t}{T_{max}} \tag{9}$$

$$B_i(t+1) = X^* + b_1(B_i(t) - Lb^*) + b_2(B_i(t) - Ub^*) \tag{10}$$

$$Lb^b = \max(X^b(1-R), Lb),$$
$$Ub^b = \max(X^b(1+R), Ub),$$
$$x_i(t+1) = x_i(t) + C_1(x_i(t) - Lb^b) + C_2(x_i(t) - Ub^b) \tag{11}$$

$$x_i(t+1) = X^b + Sg(|x_i(t) - X^*| + |x_i(t) - X^b|) \tag{12}$$

where t denotes the number of current alternations and $x_i(t)$ denotes the location of the i_th dung beetle at the t_th alternation. $k \in (0,0.2]$ is a deflection coefficient denoted by a constant and $b \in (0,1)$. α is a nature factor taken as -1 or 1, and X^w is the global worst position. Δx simulates the variation of light intensity and X^* represents the current local optimal location. $Lb^*$and Ub^* denote the bottom and top boundaries of the calving region respectively. T_{max} represents the largest iteration number. Lb and Ub denote the bottom and top boundaries of the optimization problem respectively. $B_i(t)$ is the position of the i_th sphere at the t_th alternation. b_1 and b_2 are two unrelated stochastic vectors of size $1 \times D$, and D is the dimension of the optimization problem. X^b denotes the global optimal position. Lb^b and Ub^b denote the lower and upper bounds of the optimal foraging area, respectively. C_1 denotes a random number obeying normal distribution and C_2 is a random vector in the range of $(0,1)$. g is a stochastic vector of size $1 \times D$ obeying a normal distribution and S is a constant.

2.4 GRU-TCN-DBO

The GRU model is simple, has fewer parameters, and is faster to train, so it is more suitable for building larger networks. However, GRU still cannot completely solve the gradient vanishing problem and cannot be computed in parallel. TCN can perform convolution in parallel, and can also adjust the sense field by parameters such as the number of layers, expansion factor, and filter size (**Fig. 1.**). However, TCN can only receive the original sequence up to the desired length.

The hybrid model GRU-TCN-DBO combines two different structures, GRU and TCN, taking into account the advantages of both models and finally uses DBO to optimize the parameters of the hybrid model, which in turn improves the accuracy of time series prediction. **Fig. 2.** demonstrates the flow of the hybrid model GRU-TCN-DBO for forecasting stock prices.

3 Experiments

3.1 Dataset preparation

The Dow Jones Industrial Average (DJIA) is a relatively large United States stock market index, indicating the long-term share market trading and company values of 30 large United States public companies.

Initially, all models have a sliding window of 10 days to forecast the close price for the next 10 days. The experimental dataset is 30 stocks in the Dow Jones Industrial Stock Market for the time period (2018/7/31-2023/7/31). Data division: 70% for the training set and 30% for the testing set. Data features include open, high, low, change, and volume.

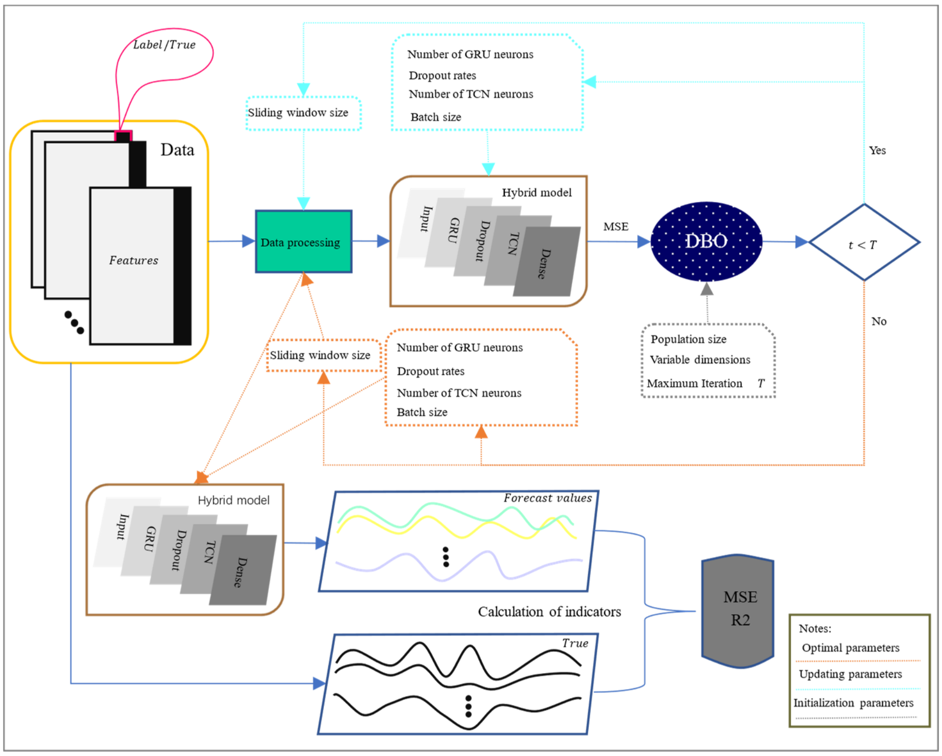

Fig. 2. A hybrid model GRU-TCN-DBO stock price prediction process framework optimized with the dung beetle optimization algorithm.

3.2 Evaluation indicators

In order to compare the results more adequately, Mean Squared Error (MSE) and R^2 were chosen as the evaluation metrics to judge the prediction performance (**Table 1**.). The smaller the MSE the smaller the error and the larger the R^2 the better the fit.

Table 1. MSE and R^2 formulas.

Indicators	MSE	$\boldsymbol{R^2}$
Formulas	$\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$	$1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}$

3.3 Model parameters

To demonstrate the effectiveness of the hybrid depth model GRU-TCN-DBO optimized based on the dung beetle optimization algorithm proposed in this paper, the hybrid models GRU-TCN, GRU, TCN, and LSTM are used for comparison experiments. All model parameters are shown in **Table 2**. Optimization parameters are x_1, x_2, x_3, sliding window and batch size.

Table 2. Model parameters.

Models	Layers	Neurons/ Filters	Kernels/Dilations
GRU-TCN	GRU	64	
	Dropout	0.1	
	TCN	64	
	Dense	1	
GRU	GRU	64	
	Dropout	0.1	
	Dense	1	
TCN	TCN	64	
	Dropout	0.1	3/[2, 4, 6, 8, 10]
	Dense	1	
LSTM	LSTM	64	
	Dropout	0.1	
	Dense	1	
GRU-TCN-DBO	GRU	x_1	
	Dropout	x_2	
	TCN	x_3	3/[2, 4, 6, 8, 10]
	Dense	1	

All models use relu activation function, adam optimizer, learning rate 0.001, batch size 50, and MSE as loss function.

3.4 Experimental results and analysis

In this experiment, we use 30 stocks on the Dow Jones Industrial Average as the dataset. We compare the proposed model GRU-TCN-DBO with the hybrid model GRU-TCN, TCN, GRU, LSTM and use MSE and R^2 as evaluation metrics.

The proposed model GRU-TCN-DBO has the best prediction results. **Fig. 3.** randomly shows the prediction result graph for two stocks out of 30 stocks. From the figure, it can be seen that the model GRU-TCN-DBO has the highest fit of five models.

For the MSE evaluation indicator, the smaller the MSE the smaller the model error. We compare five models and the results of the comparison are presented in **Table 3**. For 30 stocks, our proposed model GRU-TCN-DBO has the smallest MSE on 19 stocks, followed by GRU-TCN with 5. This shows that it is the most accurate among the five models.

For the R^2 indicator, the larger the R^2 the higher the fit of the model. We compare five models and the comparison results are shown in **Table 4**. For 30 stocks, our proposed model GRU-TCN-DBO has the largest R^2 on 19 stocks, followed by GRU-TCN with 5. This demonstrates that it has the best fit across the five models.

In conclusion, our proposed deep hybrid model GRU-TCN-DBO based on the dung beetle optimization algorithm outperforms the hybrid deep model (GRU-TCN) and individual models (GRU, TCN, LSTM) in terms of fitting and MSE on the Dow Jones Industrial Average dataset.

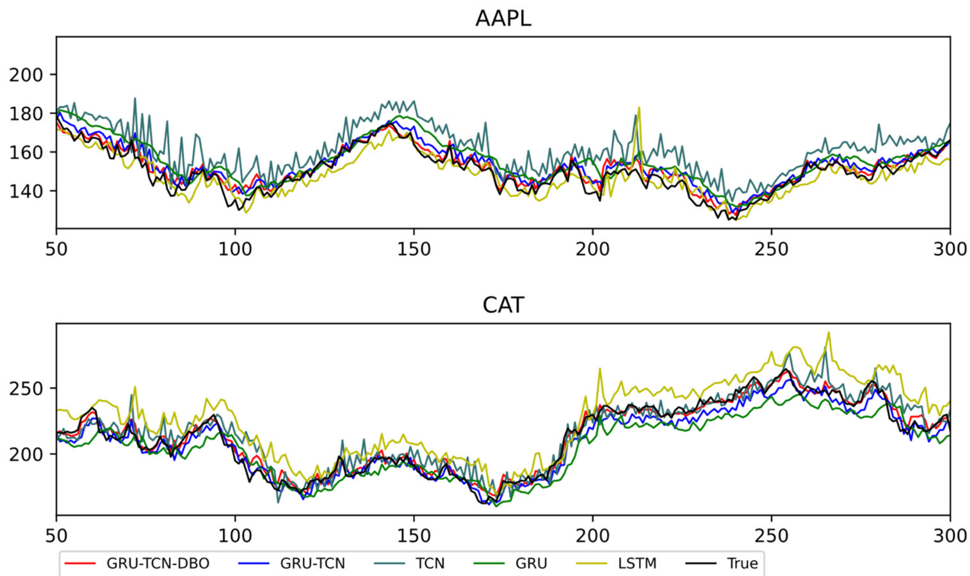

Fig. 3. Comparison of five model predictions for any two data presented on the Dow Jones Industrial Average.

Table 3. Comparative results of MSE for experimental models.

Stocks \ Models	GRU-TCN-DBO	GRU-TCN	TCN	GRU	LSTM
AAPL	**7.641**	20.852	176.736	49.960	38.885
AMGN	59.867	**47.325**	211.104	67.755	93.197
AXP	**7.349**	24.361	113.505	64.862	66.787
BA	**69.987**	78.483	331.087	100.568	187.517
CAT	**15.682**	51.455	81.812	148.682	272.748
CRM	**32.971**	35.729	88.631	44.050	159.019
CSCO	**0.539**	2.501	4.763	0.918	1.034
CVX	**66.686**	140.768	487.870	114.721	173.028
DIS	42.593	**30.001**	423.082	23.700	84.224

DOW	**2.094**	2.405	13.395	3.823	4.668
GS	**37.652**	74.335	179.167	298.155	294.178
HD	**15.600**	49.226	181.093	214.962	155.206
HON	**6.249**	74.445	54.814	39.946	60.241
IBM	**3.520**	11.093	30.129	40.201	66.907
INTC	214.806	228.177	671.561	87.821	**86.754**
JNJ	29.552	**18.401**	213.854	24.195	55.808
JPM	**6.922**	8.067	23.386	29.524	19.561
KO	8.615	4.366	37.818	4.995	**1.822**
MCD	57.752	**44.556**	156.021	139.313	85.780
MMM	876.206	984.154	3446.826	**232.625**	749.793
MRK	146.164	148.729	675.279	**8.062**	347.254
MSFT	**87.041**	358.166	321.010	313.866	1002.290
NKE	**12.283**	13.998	47.782	37.322	36.024
PG	**5.795**	9.107	77.556	10.333	31.679
TRV	40.261	86.876	131.284	**18.271**	1668.538
UNH	207.977	**155.149**	391.528	195.752	596.225
VZ	**5.286**	24.164	44.905	6.105	15.267
V	**13.502**	24.986	170.337	52.863	76.417
WBA	1.726	10.859	14.507	1.559	**1.523**
WMT	**16.480**	42.435	42.012	38.989	72.841

Table 4. Comparative results of R^2 for experimental models.

Models / Stocks	GRU-TCN-DBO	GRU-TCN	TCN	GRU	LSTM
AAPL	**0.966**	0.910	0.234	0.783	0.831
AMGN	0.793	**0.836**	0.269	0.765	0.677
AXP	**0.967**	0.890	0.488	0.708	0.699
BA	**0.921**	0.912	0.628	0.887	0.789
CAT	**0.972**	0.910	0.857	0.740	0.524
CRM	**0.950**	0.946	0.867	0.934	0.761
CSCO	**0.971**	0.852	0.719	0.946	0.939
CVX	**0.611**	0.133	-2.005	0.293	-0.066
DIS	0.858	**0.895**	-0.481	0.917	0.705
DOW	**0.949**	0.939	0.660	0.903	0.882
GS	**0.932**	0.862	0.668	0.447	0.455
HD	**0.965**	0.879	0.554	0.470	0.618
HON	**0.954**	0.448	0.593	0.704	0.553
IBM	**0.928**	0.773	0.384	0.178	-0.368
INTC	-2.396	-2.830	-10.272	-0.474	**-0.456**
JNJ	0.548	**0.722**	-2.230	0.635	0.157
JPM	**0.952**	0.943	0.835	0.792	0.862
KO	-0.708	0.134	-6.503	0.009	**0.639**

MCD	0.810	**0.863**	0.520	0.572	0.736
MMM	-1.257	-1.631	-8.216	**0.378**	-1.005
MRK	0.066	0.050	-3.315	**0.948**	-1.219
MSFT	**0.909**	0.629	0.667	0.675	-0.038
NKE	**0.932**	0.923	0.736	0.794	0.801
PG	**0.920**	0.871	-0.099	0.854	0.551
TRV	0.557	0.031	-0.465	**0.796**	-17.616
UNH	0.665	**0.750**	0.370	0.685	0.041
VZ	**0.878**	0.453	-0.017	0.862	0.654
V	**0.929**	0.877	0.159	0.739	0.623
WBA	0.944	0.637	0.515	0.948	**0.949**
WMT	**0.830**	0.568	0.573	0.603	0.259

4 Conclusions

In the paper, we present a robust and effective deep hybrid model GRU-TCN with the dung beetle optimization (DBO) algorithm for stock price prediction, called GRU-TCN-DBO, with the aim of improving the stock price prediction accuracy. The hybrid model GRU-TCN balances the advantages of GRU and TCN with powerful time series prediction and outstanding feature extraction. The dung beetle optimization algorithm is a new swarm smart optimization algorithm that can be used for the optimization of neural network parameters. We use 30 stocks on the Dow Jones Industrial Average as experimental data, each with a time horizon of 2018/7/31-2023/7/31, and compare the model proposed in this paper with other models used for stock price prediction. The experimental results demonstrate that GRU-TCN-DBO achieves the best prediction accuracy and fitting ability compared to GRU-TCN, TCN, GRU, and LSTM. Highly accurate stock price prediction using only a single deep learning model is challenging, but the deep hybrid model GRU-TCN based on dung beetle optimization can achieve better prediction accuracy than a single model or a hybrid model and provide the right direction for investment.

However, the deep hybrid model GRU-TCN based on dung beetle optimization proposed in this paper may be a preliminary exploration, and in the future, we can focus on the following aspects: a) Increase the parameters of dung beetle optimization, such as kernel size, etc. b) Optimize other models using dung beetle optimization algorithm.

References

[1] Xu, W., Fu, Z., Li, H., Huang, J., Xu, W., Luo, Y. (2023) A study of the impact of COVID-19 on the Chinese stock market based on a new textual multiple ARMA model. Stat. Anal. Data Min.: ASA Data Sci. J. 16, 5–15. https://doi.org/10.1002/sam.11582.

[2] Saleh A., Tahira N., Marriam N., Aun I. (2023) An improved DenseNet model for prediction of stock market using stock technical indicators. Expert Systems with Applications, 232: 120-903. https://doi.org/10.1016/j.eswa.2023.120903.

[3] Miklesh P., Taimur S., Shruti A., Deepika D., Mohammad Z. (2023) Investigating volatility spillover of energy commodities in the context of the Chinese and European stock markets. Research in International Business and Finance, 65: 101-948. https://doi.org/10.1016/j.ribaf.2023.101948.

[4] Kurani, A., Doshi, P., Vakharia, A., Shah, M. (2023) A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting. Ann. Data. Sci. 10, 183–208. https://doi.org/10.1007/s40745-021-00344-x.

[5] Ma T., Hu Y., Wang J., Beckline M., Pang D., Chen L., Ni X., Li X. (2023) A Novel Vegetation Index Approach Using Sentinel-2 Data and Random Forest Algorithm for Estimating Forest Stock Volume in the Helan Mountains, Ningxia, China. Remote Sensing. 15(7):18-53. https://doi.org/10.3390/rs15071853.

[6] Mohit B., Archana S., Nand K. (2023) Forecasting long-term stock prices of global indices: A forward-validating Genetic Algorithm optimization approach for Support Vector Regression. Applied Soft Computing, 145:110-566. https://doi.org/10.1016/j.asoc.2023.110566.

[7] Zaheer S, Anjum N, Hussain S, Algarni AD, Iqbal J, Bourouis S, Ullah SS. (2023) A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model. Mathematics. 11(3):590. https://doi.org/10.3390/math11030590.

[8] Vanguri N. Y., Pazhanirajan S., Kumar T. A. Tversky-RideNN based Feature Fusion and Optimized Deep RNN for Stock Market Prediction. 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2022, 1056-1063. https://doi.org/10.1109/ICIRCA54612.2022.9985572.

[9] Abdul Q. M., Sanjit K., Chris J., Arun K. S., Kong F. T., Sabireen H., Janakiraman N. (2023) Novel optimization approach for stock price forecasting using multi-layered sequential LSTM. Applied Soft Computing, 134: 109-830. https://doi.org/10.1016/j.asoc.2022.109830.

[10] Song H., Choi H. (2023) Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. Applied Sciences. 13(7):4644. https://doi.org/10.3390/app13074644.

[11] Chen C., Xue L., Xing W. (2023) Research on Improved GRU-Based Stock Price Prediction Method. Applied Sciences. 13(15):8813. https://doi.org/10.3390/app13158813.

[12] Yuan Y., Zhao Z., Yang Z. (2023) Stock index forecasting based on multivariate empirical mode decomposition and temporal convolutional networks. Applied Soft Computing, 142: 110-356. https://doi.org/10.1016/j.asoc.2023.110356.

[13] Xue J., Shen B. (2023) Dung beetle optimizer: a new meta-heuristic algorithm for global optimization. J Supercomput. 79:7305–7336 (2023). https://doi.org/10.1007/s11227-022-04959-6.