

Research Article

Forecast of E-Commerce Transactions Trend Using Integration of Enhanced Whale Optimization Algorithm and Support Vector Machine

Suqi Zhang ¹, Hsiung-Cheng Lin ², and Xinxin Wang³

¹School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China

²Department of Electronic Engineering, National Chin-yi University of Technology, Taichung 41170, Taiwan

³School of Science, Tianjin University of Commerce, Tianjin 300134, China

Correspondence should be addressed to Suqi Zhang; zhangsuqilaoshi@126.com

Received 5 April 2021; Revised 21 June 2021; Accepted 8 July 2021; Published 20 July 2021

Academic Editor: Radu-Emil Precup

Copyright © 2021 Suqi Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

E-commerce has become a crucial business model through the Internet around the world. Therefore, its transaction trend forecast can provide important information for the market planning and development in advance. For this purpose, the integrated model of enhanced whale optimization algorithm (EWOA) with support vector machine (SVM) is proposed for forecast of E-commerce transaction trend in this study. First, the global optimization ability of the whale optimization algorithm (WOA) is enhanced by the search updating strategy. Second, multiple factors that may affect the E-commerce transaction trend are analyzed and determined using the gray correlation mechanism. Third, the EWOA algorithm is employed to optimize the SVM random parameters. Finally, the EWOA-SVM model is established for forecasting E-commerce transaction trend. Two representative cases tests confirm that the EWOA-SVM model is superior to other existing methods in terms of fast convergence speed and high prediction accuracy.

1. Introduction

The current digital economy is moving forward much faster than before in recent years. Therefore, E-commerce transactions have become an important core of the digital economy in the global market [1, 2]. The rapid E-commerce development has resulted in certain changes in the logistics, manufacturing and traditional retail industries, etc. Transaction volume is usually regarded as a crucial indicator used for assessment of the E-commerce development level. As a result, E-commerce transactions trend forecast is indispensable to provide a quantitative basis for the long-term planning and strategy formulation for enterprises and governments [3, 4].

At present, the prediction methods applied for E-commerce transactions trend have been focused on machine learning models, regression models, and combination models [5, 6]. Machine learning models include neural

network model, support vector machine (SVM), extreme learning machine (ELM), etc. [7]. They are somehow sensitive to parameters selection to predict the E-commerce transactions trend through the mapping relationship between influencing factors and transactions volume [8]. The regression models mainly refer to moving average model (MA), autoregressive model (AR), autoregressive moving average model (ARMA), and the nonlinear regression models. They may work well in analyzing the stationary series, but they are not suitable for the nonstationary time series analysis. On the other hand, the combination models can provide competitive results and outperform the single model, but their computational cost is relatively higher than that of the single model [9, 10].

Zhang et al. [11] applied the ELM model for forecasting E-commerce transactions and proposed an improved optimization algorithm to optimize the random model parameters. However, only few basic influencing factors were

considered so that it may be not applicable in real circumstances. Ji et al. [12] combined XGBoost and ARIMA models to predict the size of E-commerce transactions, presenting better results than individual XGBoost and ARIMA models, but the complexity and computational cost may post a difficulty for further applications. Alternatively, Chen et al. [13] combined clustering technology and machine learning model to forecast the transaction size. Clustering technology is used to divide the training samples, and then the machine learning model was applied to train groups. The random parameters that may influence machine learning performance were not well considered, resulting in the instability of prediction results. Di Pillo et al. [14] employed SVM to predict the sales scale. Compared with linear regression models, the SVM model has a stronger nonlinear mapping ability. However, it is sensitive to random model parameters. Mao et al. [15] predicted China's E-commerce online transaction volume based on the combination of ARMA model and gray model, and the forecasting result was satisfactory.

SVM model is suitable for small sample prediction, and it has strong generalization ability and less random parameters. Li et al. [16] proposed an improved dragonfly algorithm to optimize SVM's random parameters in short-term wind power prediction. Liu et al. [17] applied SVM to forecast the remaining life of lithium-ion batteries and used chicken swarm optimization algorithm to solve random parameters. Pham et al. [18] employed the SVM model for rainfall prediction and achieved high prediction accuracy. Additionally, Hossain and Muhammad [19] applied SVM in emotion recognition system for emotion classification. Huang and Wang [20] used SVM model for pattern classification, where the genetic algorithm was applied to optimize model's random parameters to improve classification accuracy.

For SVM to be used for forecasting the E-commerce transaction trend, two main problems need to be resolved. The first task is to reduce the influence of random parameters of SVM model, which is an optimization point. The other is to choose crucial factors on E-commerce transaction trend. Consequently, this study proposes an integrated model using enhanced Whale Optimization Algorithm (EWOA) with SVM model on the basis of multiple factors analysis and machine learning. In this approach, EWOA algorithm was used to optimize the random parameters of

SVM model. The modeling process is introduced in Section 2. Section 3 analyzes multiple influencing factors and determines critical points on E-commerce transactions. Section 4 validates the proposed model through two cases. The conclusions with future work are presented in Section 5.

2. Models of E-Commerce Transaction Trend Forecast

2.1. SVM Model. SVM, which is a hot spot model in machine learning models, has the characteristics of simple structure, few adjustable parameters, and strong generalization ability. It is often used in pattern recognition, disease diagnosis, regression prediction, and other fields [21]. The samples are mapped to the high-dimensional space R^N through a mapping function $\varphi(x)$. Given a sample set $\{(x_i, y_i) | x_i \in R^N, y_i \in R, j = 1, 2, \dots, n\}$, the hyperplane g between input x and output y is established in a high-dimensional space as follows [22]:

$$g(x) = w \cdot \varphi(x) + l, \quad (w \in R^N, l \in R), \quad (1)$$

where $g(x)$ represents the output function, w is the weight, and l indicates the offset.

$g(x)$ is transformed into a constrained optimization problem through the principle of structural risk minimization. Taking into account the errors, the slack variable is introduced into the objective function. The constraints minimization is expressed as follows [23]:

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + \rho \sum_{i=1}^n (\zeta_i + \zeta_i^*), \\ & \begin{cases} y_i - l - w \cdot x_i \leq \tau + \zeta_i, \\ l + w \cdot x_i - y_i \leq \tau + \zeta_i^*, \\ \zeta_i^* \geq 0, \\ \zeta_i \geq 0, \end{cases} \end{aligned} \quad (2)$$

where ζ_i and ζ_i^* are slack variables, ρ is a penalty coefficient, and τ is the error.

The optimization problem is transformed into solving the equation by Lagrange multiplier, then the derivation of each variable is performed, and finally the dual form of the optimal problem is obtained [24, 25].

$$\begin{cases} \min \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^\wedge - \alpha_i)(\alpha_j^\wedge - \alpha_j) \kappa(x_i \cdot x_j) + \tau \sum_{i=1}^n (\alpha_i^\wedge - \alpha_i) - \sum_{i=1}^n y_i (\alpha_i^\wedge - \alpha_i) \sum_{i=1}^n (\alpha_i^\wedge - \alpha_i) = 0 \alpha_i^\wedge, \alpha_i \geq 0, & i = 1, 2, 3, \dots, n, \end{cases} \quad (3)$$

where α_i^\wedge and α_i are Lagrange multipliers, $\kappa(x_i \cdot x_j)$ is kernel functions, and the radial basis function (RBF) kernel is used in this study.

$$\kappa(x_i, x_j) = e^{-\left(\frac{\|x_i - x_j\|^2}{2\delta^2}\right)}, \quad (4)$$

where δ ($\delta > 0$) is the size of the kernel parameter.

Finally, the SVM regression function is defined as follows:

$$g(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \kappa(x_i, x_j) + l. \quad (5)$$

Generally, in SVM model, the penalty coefficient ρ and kernel coefficient δ are random parameters, which bring uncertainty to the prediction results under the complexity of the data. To solve this problem, these random parameters need to be optimized. For this reason, a whale optimization algorithm based on search updating strategy is proposed to optimize SVM parameters and achieve the prediction accuracy for E-commerce transaction trend.

2.2. WOA Algorithm. To date, a variety of intelligent algorithms have been developed and applied, such as PSO algorithm [26], crow search algorithm (CSA) [27], and a series of hybrid swarm algorithms [28–33]. Each algorithm uses a different method or strategy to fit some specific purposes or applications. For example, Zapata et al. [34] developed a hybrid swarm algorithm for collective construction of 3D structures, and Precup [35] proposed slime mould algorithm-based tuning of cost-effective fuzzy controllers for servo systems. Alternatively, WOA, which has a strong optimization ability, is a new swarm intelligence optimization algorithm [36]. It can simulate the predatory behavior of whales in nature, including foraging, encircling, bubble hunting, and food searching [37]. During the foraging phase, the information is exchanged between individuals in the whale group, and the food location is determined through the information communication. Usually, the initial optimal target position is used as the food position, and the whale can approach the food by updating its position. The whale location update strategy is described as follows [38, 39]:

$$\mathbf{x}(m+1) = \mathbf{x} \wedge (m) - \mathbf{A} \bullet \mathbf{B}, \quad (6)$$

$$\mathbf{B} = |\mathbf{C} \bullet \mathbf{x} \wedge (m) - \mathbf{x}(m)|, \quad (7)$$

where m is the current iteration number; \mathbf{A} and \mathbf{C} are coefficient matrices; \mathbf{B} represents the distance between the whale and the food; \mathbf{x} is the whale position, and $\mathbf{x} \wedge$ represents the optimal position in the whale group.

The coefficient matrices \mathbf{A} and \mathbf{C} in equations (6) and (7) are calculated as follows:

$$\mathbf{A} = 2\mathbf{a} \bullet \mathbf{u} - \mathbf{a}, \quad (8)$$

$$\mathbf{C} = 2\mathbf{u},$$

where \mathbf{u} is a random number between 0 and 1; \mathbf{a} decreases linearly from 2 to 0 in the iterative process.

The whales adopt enveloping and spiraling behaviors in the predation stage. To realize the contraction encirclement, \mathbf{A} decays with \mathbf{a} that decreases from 2 to 0. The prey is attacked through spiraling model when the food location is locked. At this time, the location search updating strategy of whales is shown as follows [40, 41]:

$$\mathbf{x}(m+1) = \mathbf{x} \wedge (m) + \mathbf{B} \wedge \bullet e^{bo} \bullet \cos(2\pi o), \quad (9)$$

$$\mathbf{B} \wedge = |\mathbf{x} \wedge (m) - \mathbf{x}(m)|$$

where b ($b = 1$) as a constant is the spiral shape; o is a random number in the interval $[-1, 1]$; $\mathbf{B} \wedge$ represents the distance between the whale and the locked food.

Assume that the probability of the whale taking the action of shrinking encirclement and spiral attack is 50%, and the position updating strategy is expressed as follows:

$$\begin{cases} \mathbf{x}(m+1) = \mathbf{x} \wedge (m) - \mathbf{A} \bullet \mathbf{B}, & p < 0.5, \\ \mathbf{x}(m+1) = \mathbf{x} \wedge (m) + \mathbf{B} \wedge \bullet e^{bo} \bullet \cos(2\pi o), & p \geq 0.5, \end{cases} \quad (10)$$

where p is a random number in the interval $[0, 1]$.

In addition, whales randomly searching for food can succeed through updating \mathbf{A} . The whale can search for food in a larger area when $|A| > 1$ and search for food in a smaller area when $|A| < 1$.

$$\mathbf{B} = |\mathbf{C} \bullet \mathbf{x}_{\text{rand}}(m) - \mathbf{x}(m)|, \quad (11)$$

$$\mathbf{x}(m+1) = \mathbf{x}_{\text{rand}}(m) - \mathbf{A} \bullet \mathbf{B}, \quad (12)$$

where \mathbf{x}_{rand} represents a random position vector.

In the WOA algorithm, most of the parameters are random, and only the maximum number of iterations and population size need to be set, which is one of the advantages of the algorithm.

2.3. EWOA Algorithm. Wolpert and Macready [42] believed that no optimization algorithm can solve all optimization problems according to the “no free lunch” theory. It means that different optimization algorithms may obtain different solutions under the same issue. Therefore, developing new algorithms may achieve better results. The traditional WOA algorithm may suffer from some disadvantages even it has stronger optimization capability than Particle Swarm Optimization (PSO) and differential evolution algorithms, etc. For example, its coefficient a decreases linearly to achieve shrinking encirclement, but the dynamic changes during the iteration process are not convincing. The population diversity is also limited in the later iteration, resulting in being trapped into local minimum. The EWOA model is thus developed to resolve above problems. First, the dynamic attenuation coefficient d_a is introduced to simulate the dynamic change situation in the shrinking and enveloping behaviors of whales during the iterative process. The mathematical model of the dynamic attenuation coefficient is defined as follows:

$$d_a = 2 * \left(\frac{(m_{\text{max}} - m)}{m_{\text{max}}} \right)^3, \quad (13)$$

where m_{max} represents the maximum number of iterations and m is the current iteration coefficient.

The value of dynamic attenuation coefficient over iterations is depicted in Figure 1. It can be seen that the coefficient (d_a) value declines faster in the initial stage of the

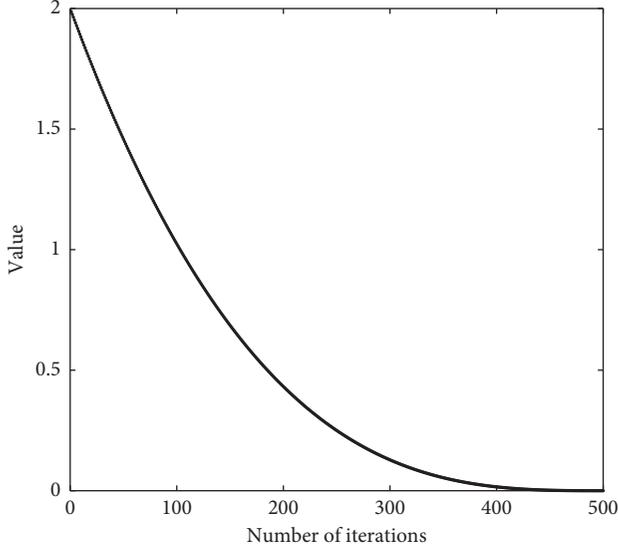


FIGURE 1: Dynamic attenuation coefficient over iterations.

iteration, which enables locking the food position shortly. During the later iteration period, it decays slower but strengthens the algorithm's local exploration ability instead.

Aiming at the deficiency of population diversity weakening in the later iteration, an area search updating strategy is proposed. The whales migrate to other regions to search for food by the regional update frequency M during the optimization process. It can promise the population diversity reaching to a large extent, thus improving the algorithm's optimization ability. The whale's update position is shown as follows:

$$\mathbf{x}(m+1) = \text{rand}n(0, \sigma^2) \bullet \mathbf{x}(m) + \mathbf{x}(m), \quad (14)$$

where $\text{rand}n(0, \sigma^2)$ obeys Gaussian distribution.

Similar to WOA algorithm, most parameters in EWOA algorithm are random, but the maximum number of iterations, population size, and migration frequency need to be set in advance.

The process flowchart of the EWOA algorithm for searching the global optima is shown in Figure 2.

As shown in Figure 2, EWOA algorithm optimization process includes the following steps:

- (1) Initialize EWOA algorithm parameters.
- (2) Determine whether to implement area search updating strategy.
- (3) If the area search updating strategy is implemented, the location is updated according to equation (14); otherwise the location is updated according to equations (10) and (12) [36].
- (4) Update the optimal location of whale population [37].
- (5) Determine whether to terminate the iteration. If the iteration is terminated, the optimization is completed. Otherwise, return to step (2).

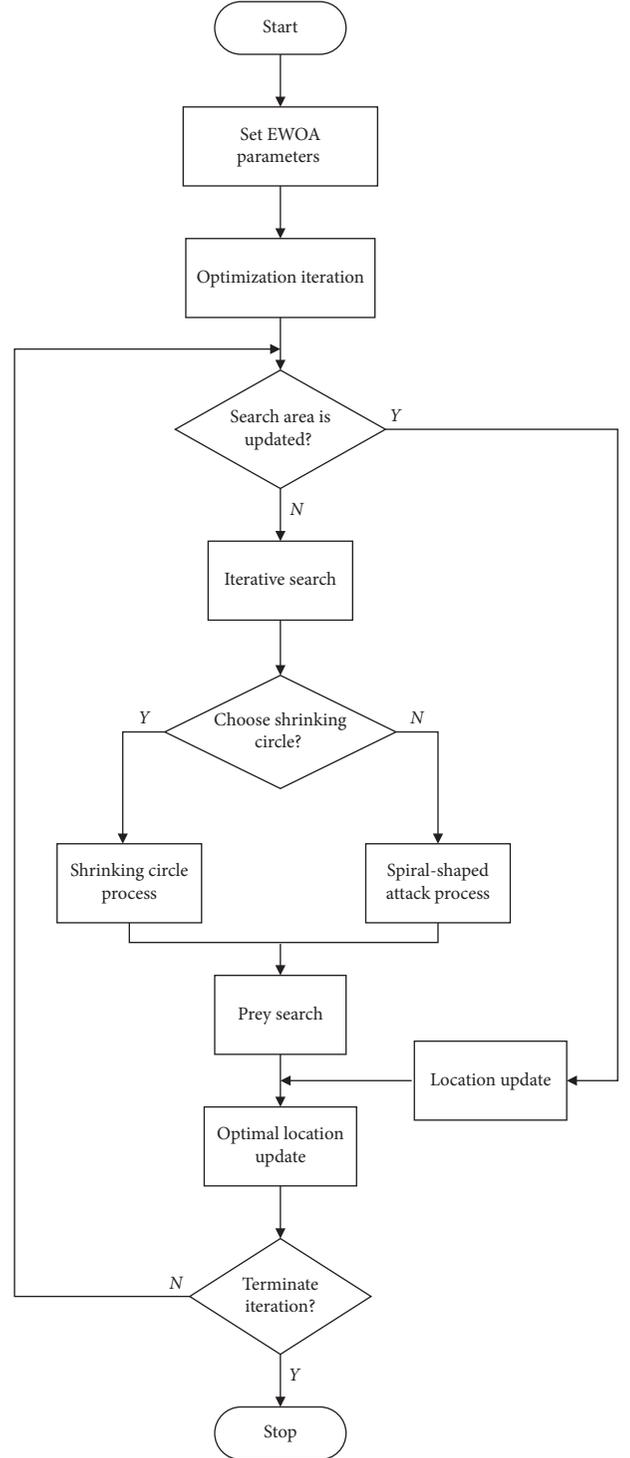


FIGURE 2: Flowchart of EWOA optimization process.

2.4. *Convergence Analysis.* There are five standard test functions used to analyze the model convergence efficiency. The f_1 , f_2 , and f_3 are unimodal functions, where the local extremum is the global optima, while f_4 and f_5 are multimodal functions. The variable ranges in functions are listed in Table 1.

TABLE 1: Standard test functions and variable ranges [43–45].

Equations	Variable ranges	Global optima	dim
$f_1(v) = \sum_{i=1}^{\dim} v_i^2$	$[-100, 100]$	0	30
$f_2(v) = \sum_{i=1}^{\dim} (\sum_{j=1}^i v_j)^2$	$[-100, 100]$	0	30
$f_3(v) = \sum_{i=1}^{\dim} v_i + \prod_{i=1}^{\dim} v_i $	$[-10, 10]$	0	30
$f_4(v) = (1/4000) \sum_{i=1}^{\dim} v_i^2 - \prod_{i=1}^{\dim} \cos(v_i/\sqrt{i}) + 1$	$[-600, 600]$	0	30
$f_5(v) = 10 * \dim + \sum_{i=1}^{\dim} (v_i^2 - 10 * \cos(2\pi v_i))$	$[-5.12, 5.12]$	0	30

Note: “dim” denotes the test dimension.

In addition to EWOA algorithm, PSO algorithm [26], crow search algorithm (CSA) [27], and classic WOA algorithm were chosen tests for comparison. Note that CSA algorithm is a new type of swarm intelligence optimization algorithm with better convergence performance and is suitable as a comparison algorithm. PSO algorithm is a classic optimization algorithm and is usually used as a comparison algorithm. Simultaneously, the traditional WOA algorithm is used as a comparison algorithm to compare the convergence results with the EWOA algorithm. Algorithms’ parameters are set as shown in Table 2.

In PSO, w_{\max} and w_{\min} are the maximum and minimum values, respectively. $C1$ and $C2$ are the learning coefficients. In CSA, AP is the consciousness probability. FL is the flight length. b is used to define the spiral shape in both WOA and EWOA algorithms. M represents the search update frequency in EWOA. The population size is 30, and the maximum number of iterations is 500. The algorithms are tested under a unified platform, and each optimization algorithm is repeated 30 times for optimizing each test function. The average convergence value, the best convergence value, and the worst convergence value from every test function are concluded in Table 3.

As can be seen, the EWOA algorithm presents better search outcomes than others. Obviously, the results obtained from unimodal functions such as f_1 , f_2 , and f_3 consistently outperform those of multimodal functions like f_4 and f_5 . Furthermore, PSO and CSA algorithms showed poor optimization results in multimodal functions, where the worst value is up to 28.85 in PSO. For all unimodal functions, the WOA algorithm can achieve very low optimal value close to 0, while the EWOA algorithm reaches the optimal value, i.e., 0.

The fitness function to evaluate the convergence process is defined as follows:

$$\text{fitness} = \sqrt{\frac{\sum_{i=1}^n (P_{\text{train},i} - P_{\text{train},i}^*)^2}{n}}, \quad (15)$$

where n is the number of training samples; $P_{\text{train},i}$ represents the E-commerce transaction training value; and $P_{\text{train},i}^*$ indicates the E-commerce transaction prediction value.

The iterative convergence curves (log (Fitness)) using WOA and EWOA algorithms in various test functions are shown in Figure 3. The convergence speed of EWOA algorithm is considerably faster than WOA algorithm in all test functions, requiring much shorter iterations to converge.

TABLE 2: Optimization algorithm parameters.

Algorithms	Parameters
PSO	$w_{\max} = 0.9$; $w_{\min} = 0.3$; $C1 = C2 = 1.49$
CSA	$AP = 0.1$; $FL = 2$
WOA	$b = 1$
EWOA	$b = 1$; $M = 20$

TABLE 3: Achievement of convergence results.

Test function	Algorithm	Average result	Best result	Worst result
f_1	PSO	0.1083	$5.51e-07$	2.41
	CSA	$2.49e-04$	$4.54e-06$	$9.01e-04$
	WOA	$1.49e-72$	$6.38e-85$	$4.34e-71$
	EWOA	$1.51e-266$	$8.82e-313$	$4.52e-265$
f_2	PSO	24.66	$2.38e-02$	373.84
	CSA	0.16	$2.70e-03$	1.75
	WOA	$3.98e-09$	$3.95e-11$	$4.72e-06$
	EWOA	$1.56e-217$	$1.89e-248$	$4.46e-216$
f_3	PSO	0.81	$4.98e-02$	3.09
	CSA	$6.85e-02$	$1.90e-03$	0.81
	WOA	$5.54e-51$	$7.70e-59$	$1.01e-49$
	EWOA	$1.50e-154$	$2.38e-173$	$4.49e-153$
f_4	PSO	0.31	$6.74e-02$	0.75
	CSA	0.17	$4.79e-02$	0.47
	WOA	$5.8e-03$	0	0.17
	EWOA	0	0	0
f_5	PSO	15.52	7.96	28.85
	CSA	9.25	3.97	17.90
	WOA	$3.78e-15$	0	$5.68e-14$
	EWOA	0	0	0

3. Analysis of Multiple Influencing Factors in E-Commerce Transactions Trend

E-commerce transaction volume indicates the average value of E-commerce sales volume and E-commerce procurement volume. The E-commerce transactions sample set was collected from 2005 to 2019 in China, including the annual E-commerce transaction volume and the multiple influencing factors. The influencing factors in E-commerce transactions mainly consist of basic resource, transaction level, and economic development level. The basic resources include the Internet penetration rate (A1; unit: %), number of websites (A2; unit: ten thousand), number of CN domain names (A3, unit: ten thousand), and number of Internet

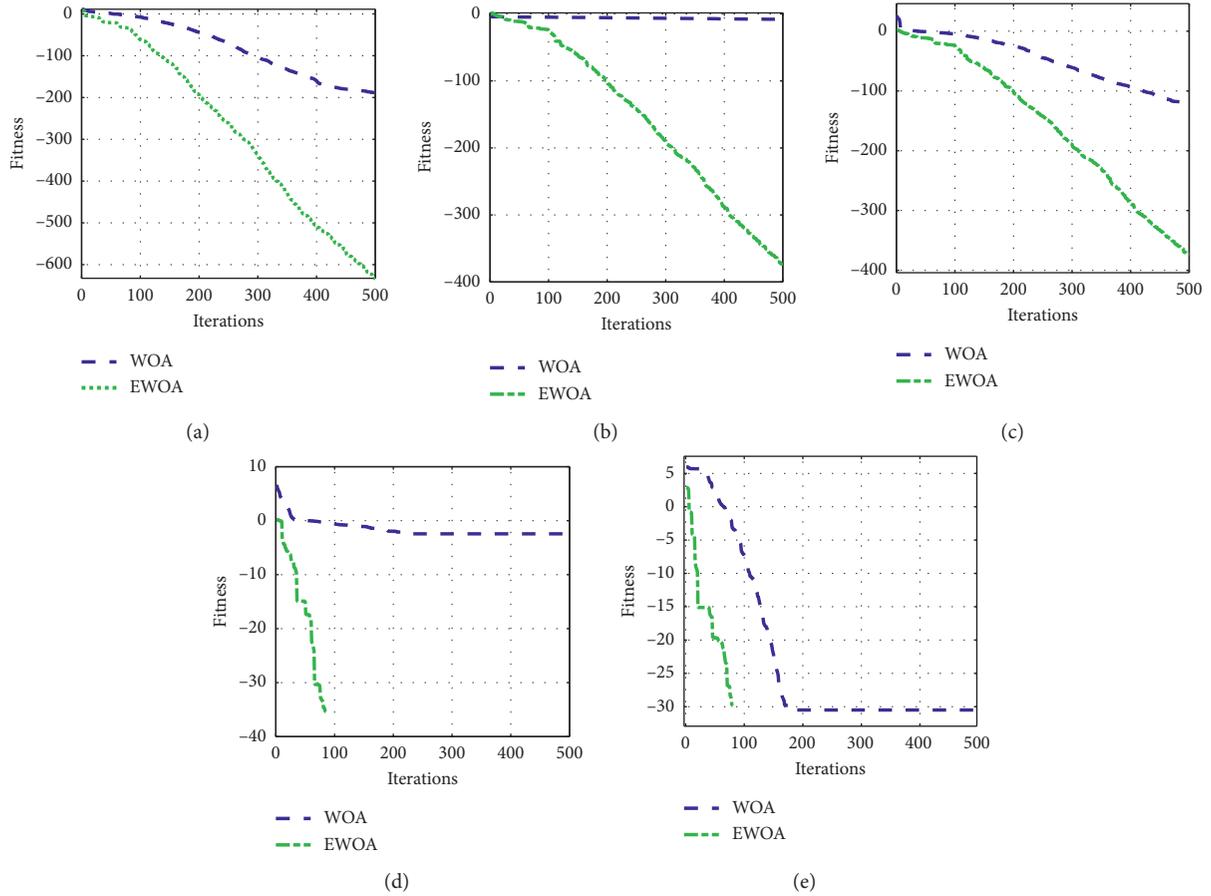


FIGURE 3: Convergence curves of WOA and EWOA algorithms. (a) f_1 . (b) f_2 . (c) f_3 . (d) f_4 . (e) f_5 .

users (A4; unit: 100 million). Indeed, the Internet penetration rate is to reflect the sharing degree of basic resources. Websites applied as a transaction platform are an important index on E-commerce transactions; the number of CN domain names can reflect the number of Internet service companies. The number of Internet users is to reflect the demand for shopping services via Internet.

In transaction level, the factors including express delivery business volume (A5; unit: 100 million) and express delivery business revenue (A6; unit: 100 million yuan) have a crucial impact on transactions level. Among them, express delivery business plays an important role in the E-commerce sales. On the other hand, the express delivery business revenue can reflect the level of E-commerce transactions in the express delivery industry. In economic development level, Gross Domestic Product (GDP) regarded as a macroeconomic factor is considered a key factor in the economic activity. It can reflect the situation of the E-commerce development. Therefore, GDP (A7; unit: trillion yuan) is used to evaluate the E-commerce transactions in this study.

The statistics on China's E-commerce transactions volume and the influence factors from 2005 to 2019 are presented in Table 4. It indicates that E-commerce transaction volume (T) increases every year, i.e., from 1.29 trillion in 2005 to 34.81 trillion in 2019. Similarly, other influencing factors, e.g., A1–A7, in E-commerce transactions have a

TABLE 4: Statistics of E-commerce transactions.

Year	A1	A2	A3	A4	A5	A6	A7	T
2005	8.5	69.4	109	1.11	8.66	239.7	18.58	1.29
2006	10.5	84	180	1.37	10.6	299.7	21.76	1.54
2007	16	150	900	2.1	12.02	342.6	26.8	2.17
2008	22.6	287	1357	2.98	15.13	408.4	31.67	3.14
2009	28.9	323	1368	3.84	18.58	479	34.56	3.67
2010	34.3	191	435	4.57	23.39	574.6	40.89	4.55
2011	38.3	250	353	5.13	36.73	758	48.41	6.09
2012	42.1	255	751	5.64	56.85	1005.3	53.41	8.11
2013	45.8	320	1082	6.17	91.87	1441.7	58.8	10.4
2014	47.9	355	1129	6.48	139.59	2045.4	63.59	16.39
2015	50.3	425	1636	6.88	206.7	2760	68.9	21.79
2016	53.2	482	2061	7.51	313.5	4005	74.41	26.1
2017	55.8	533	2084	7.72	400.6	4957	82.07	29.16
2018	59.6	523	2124	8.29	507.1	6010	91.92	31.63
2019	61.2	518	2185	8.85	630	7450	99.08	34.81

Note: " T " represents E-commerce transaction volume (trillion yuan; 1 yuan \approx 0.15 USD).

growing trend, in which A5 and A6 increase much more than others.

The gray correlation is employed to analyze the correlation degree between multiple influencing factors and E-commerce transaction volume. Initially, the dimensionless process in E-commerce transaction volume and

influencing factors is implemented to reduce the difference between the numerical values. Then, the correlation coefficient is calculated. The E-commerce transaction volume is denoted as the reference sequence $\{x_0(t)\}$, and the influencing factors are denoted as the comparison subsequence $\{x_i(t)\}$. The correlation coefficients at time k are expressed as follows. The difference $(\Delta_i(k))$ between the reference sequence $\{x_0(t)\}$ and the comparison sequence $\{x_i(t)\}$ is as follows:

$$\Delta_i(k) = |x_0(k) - x_i(k)|. \quad (16)$$

The correlation coefficient $C_i(k)$ between the i_{th} comparison sequence and the reference sequence is as follows:

$$C_i(k) = \frac{\Delta_{\min} + 0.5\Delta_{\max}}{\Delta_i(k) + 0.5\Delta_{\max}}, \quad (17)$$

where $\Delta_i(k)$ represents the absolute difference between the two sequences at time k , Δ_{\min} is the minimum absolute difference of the comparison sequences, and Δ_{\max} is the maximum absolute difference of the comparison sequences.

The correlation degree (cr_i) between the i_{th} comparison sequence and the reference sequence is defined in the following equation:

$$cr_i = \frac{1}{n} \sum_{k=1}^n C_i(k), \quad (18)$$

where n denotes the number of the sequence data.

The correlation degree between E-commerce transaction volume and multiple influencing factors is presented in Table 5.

It reveals that the highest correlation degree in the express business revenue (A6) reaches 0.91; the correlation degrees in Internet penetration rate (A1), number of CN domain names (A3), and number of Internet users (A4) exceed 0.8; the correlation degrees in website number (A2), express delivery business volume (A5), and GDP (A7) are below 0.8. As above, the collected data from A1, A3, and A4, A6 are considered as the input variables for the forecasting models.

4. Model Implementation with Case Analysis

4.1. E-Commerce Transaction Prediction Using the EWOA-SVM Model. Based on the integration of EWOA and SVM models, the proposed EWOA-SVM model is established to forecast E-commerce transactions trend. The architecture of prediction process is depicted in Figure 4, being demonstrated as follows:

- (1) Analyze the impact of multiple influencing factors on E-commerce transaction volume
- (2) Calculate the correlation degree between different influencing factors and E-commerce transactions through gray correlation using equation (18)
- (3) Select strongly related factors with E-commerce transactions as model input variables

TABLE 5: Gray correlation degree in various influencing factors.

Gray correlation degree (cr_i)			
A1	0.81	A5	0.79
A2	0.79	A6	0.91
A3	0.85	A7	0.77
A4	0.81		

- (4) Classify the training set and test set, and normalize the data
- (5) Construct E-commerce transaction trend prediction model using EWOA-SVM
- (6) Set the parameters of EWOA algorithm and SVM model [40]
- (7) Train EWOA-SVM model using training set
- (8) Use EWOA to optimize the random parameters of SVM
- (9) Calculate the fitness values of EWOA algorithm through equation (15) [41]
- (10) Output the optimal parameters of SVM after the training process is complete
- (11) Verify EWOA-SVM model using test set
- (12) Employ trained SVM to predict E-commerce transaction trends
- (13) Evaluate the forecast results of E-commerce transaction

The performance of all algorithms throughout this study was carried out using MATLAB software, and the code of core programs and datasets can be freely accessed on the web page https://drive.google.com/drive/folders/1OPlt_W_u8XHrvT_PW-wOwUBZtUL2mind?usp=sharing.

The root mean square error (rmse) [46] and fitting coefficient r_2 [47] are used to evaluate the model performance, where rmse represents the prediction error, and r_2 indicates the changing trend of the predictive values. When the fluctuating trend of the predictive value is closer to the real one, the r_2 is closer to 1.

$$\begin{aligned} \text{rmse} &= \sqrt{\frac{1}{\text{num}} \sum_{i=1}^{\text{num}} (P_{\text{test},i} - P_{\text{test},i}^*)^2}, \\ r_2 &= 1 - \frac{\sum_{i=1}^{\text{num}} (P_{\text{test},i} - P_{\text{test},i}^*)^2}{\sum_{i=1}^{\text{num}} (P_{\text{test},i} - \bar{P}_{\text{test},i})^2}, \end{aligned} \quad (19)$$

where num is the number of testing samples; $P_{\text{test},i}$ is the test value of E-commerce transaction; $P_{\text{test},i}^*$ represents the predicted value of E-commerce transaction; and $\bar{P}_{\text{test},i}$ is the average test value of E-commerce transaction.

4.2. Case 1. Two cases were used to test the effectiveness of the proposed EWOA-SVM model, also including the SVM and WOA-SVM model for comparison. SVM model is selected to analyze the influence of random parameters on the prediction results, and WOA-SVM model is chosen to

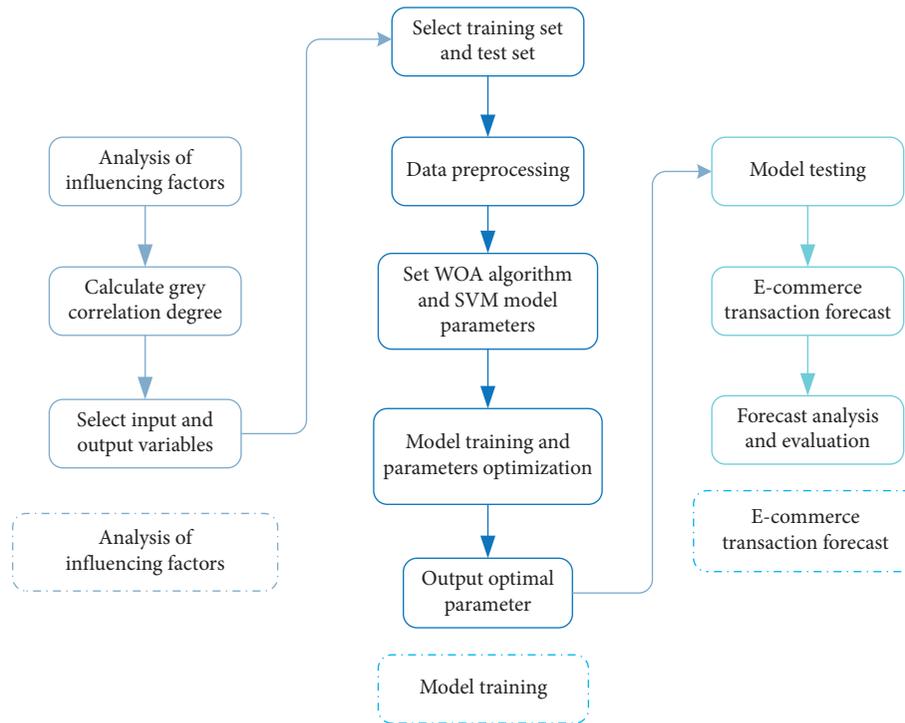


FIGURE 4: Flowchart of E-commerce transaction forecast.

compare with the mining capability of the EWOA algorithm. The E-commerce transaction data collected from 2005 to 2014 was chosen as the training set, and the data collected between 2015 and 2019 was used as the test set. The training convergence curves in both WOA-SVM and EWOA-SVM models are presented in Figure 5. It indicates that the convergence speed of the EWOA algorithm is significantly faster than WOA algorithm. Moreover, the fitness value of the EWOA algorithm is obviously smaller than that of the WOA algorithm during the iteration process.

The test results from the performance of SVM, WOA-SVM, and EWOA-SVM models are presented in Figure 6. The specific prediction values are listed in Table 6. It is found that the prediction of WOA-SVM and EWOA-SVM models fits well with the actual E-commerce transaction curve. On the contrast, the SVM model shows a relatively higher error, particularly in 2015.

4.3. Case 2. The E-commerce transaction data collected between 2005 and 2010 was selected as the training set, and the data collected between 2011 and 2019 was used as the test set. The training convergence curves of the WOA-SVM and EWOA-SVM models are presented in Figure 7. It reveals that the convergence speed of the EWOA algorithm is faster than that of the WOA algorithm. Besides, the fitness value of the EWOA algorithm is considerably smaller than that of the WOA algorithm during iteration process. The results from the prediction performance of SVM, WOA-SVM, and EWOA-SVM models are shown in Figure 8. It indicates that the prediction accuracy of WOA-SVM and EWOA-SVM models is relatively higher than that of SVM model in

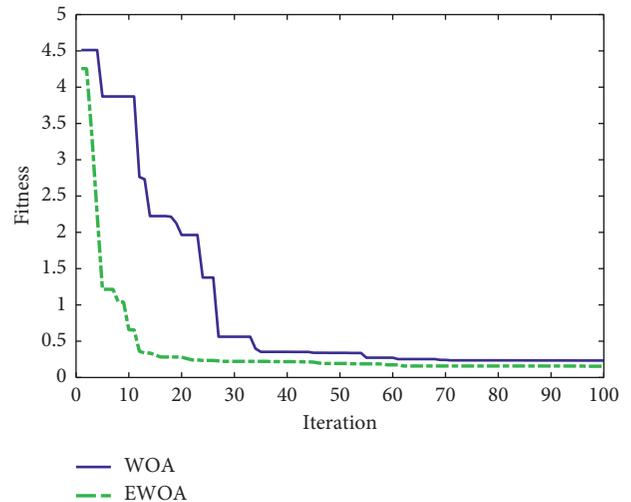


FIGURE 5: Training convergence curves of WOA and EWOA algorithms between 2005 and 2014.

general, well fitting with the actual E-commerce transactions trend. Nevertheless, the predicted result in SVM model is satisfactory except 2014 to 2015, showing more deviation from the actual values. The detailed prediction outcomes are listed in Table 7.

4.4. Evaluation of Test Results. The SVM, WOA-SVM, and EWOA-SVM models were used to predict the trend of E-commerce transactions in Cases 1 and 2. The prediction results of the model were evaluated in this section. For Cases 1 and 2, the relative error (Re%) curves from SVM,

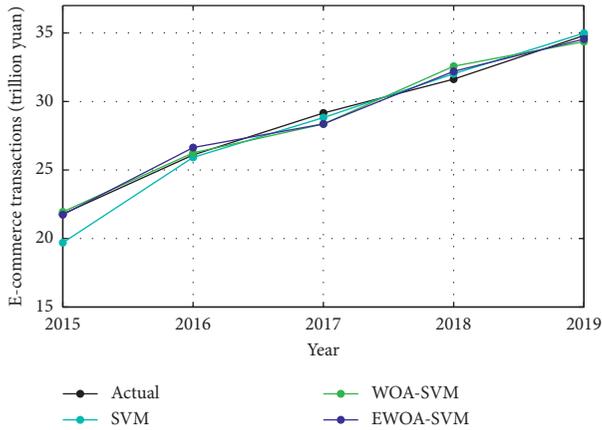


FIGURE 6: Test results of E-commerce transactions between 2015 and 2019.

TABLE 6: Results of E-commerce transactions forecast from 2015 to 2019.

Year	Actual	SVM	WOA-SVM	EWOA-SVM
2015	21.79	19.69616	21.94652	21.73803
2016	26.1	25.92237	26.24643	26.63792
2017	29.16	28.82402	28.35714	28.36315
2018	31.63	32.02076	32.57347	32.20004
2019	34.81	34.98426	34.36644	34.55087

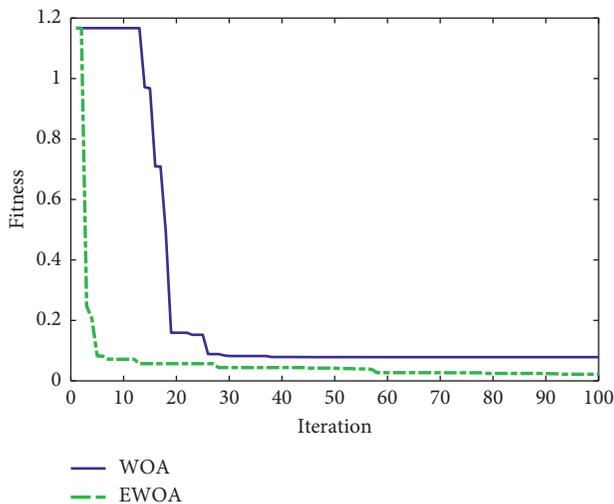


FIGURE 7: Training convergence curves of WOA and EWOA algorithms between 2005 and 2010.

WOA-SVM, and EWOA-SVM models are shown in Figure 9, where the $Re\%$ values are concluded in Table 8.

For Case 1, the Re interval of the SVM model was $[-9.61\%, 1.23\%]$; the Re interval of the WOA-SVM model was $[-2.75\%, -2.98\%]$; the Re interval of the EWOA-SVM model was $[-2.73\%, 2.06\%]$. The fluctuation range of the EWOA-SVM model was the smallest, and the Re error of the model was significantly smaller than the other two models. For Case 2, the maximum Re value for WOA-SVM and

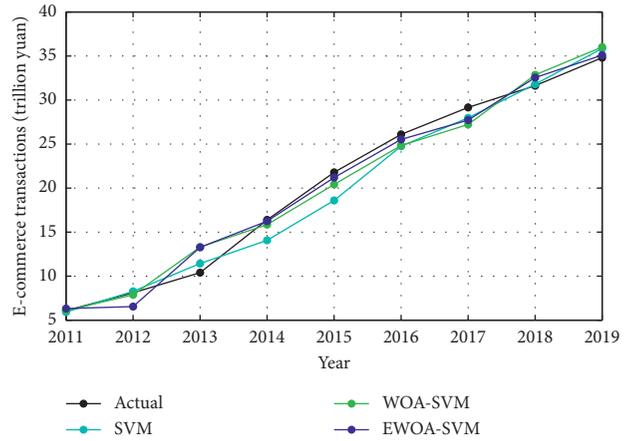


FIGURE 8: Test results of E-commerce transactions from 2011 to 2019.

TABLE 7: Results of E-commerce transactions forecast from 2011 to 2019.

Year	Actual	SVM	WOA-SVM	EWOA-SVM
2011	6.09	5.923791	6.09362	6.322908
2012	8.11	8.275549	7.878242	6.545946
2013	10.4	11.43486	13.30829	13.27391
2014	16.39	14.07154	15.8526	16.24195
2015	21.79	18.59799	20.41689	21.19814
2016	26.1	24.75947	24.83894	25.52773
2017	29.16	27.97686	27.22712	27.7213
2018	31.63	31.79973	32.85884	32.54913
2019	34.81	35.85496	36.01635	35.09898

EWOA-SVM exceeded 27%. The prediction error of E-commerce in 2013 was relatively large, but the remaining errors were less than 20%. The overall prediction effects of WOA-SVM and EWOA-SVM models were better than that of SVM model.

The prediction evaluation results using the $rmse$ and r_2 from the SVM, WOA-SVM, and EWOA-SVM models are listed in Table 9. In Case 1, the $rmse$ values of WOA-SVM and EWOA-SVM models are contained below 0.6, where the minimum $rmse$ value as 0.51 is obtained in EWOA-SVM model. It verifies that the $rmse$ of EWOA-SVM model is 13.56%, 47.42% smaller than that of WOA-SVM and SVM models, respectively. However, the fitting result of the SVM model is better than others, showing its r_2 value up to 99%. In Case 2, the $rmse$ values in all three models significantly increase, being compared with Case 1. The minimum $rmse$ of the EWOA-SVM model is 1.26, which is 14.86% and 17.64% smaller than that of the WOA-SVM and SVM models, respectively. Additionally, the EWOA-SVM model reaches the highest r_2 value up to 98.42% among all models.

4.5. Discussion about Theory and Real Applications. At present, the global economy has entered a brand-new information network era. E-commerce is a new type of business operation model, which has been fast inspired with

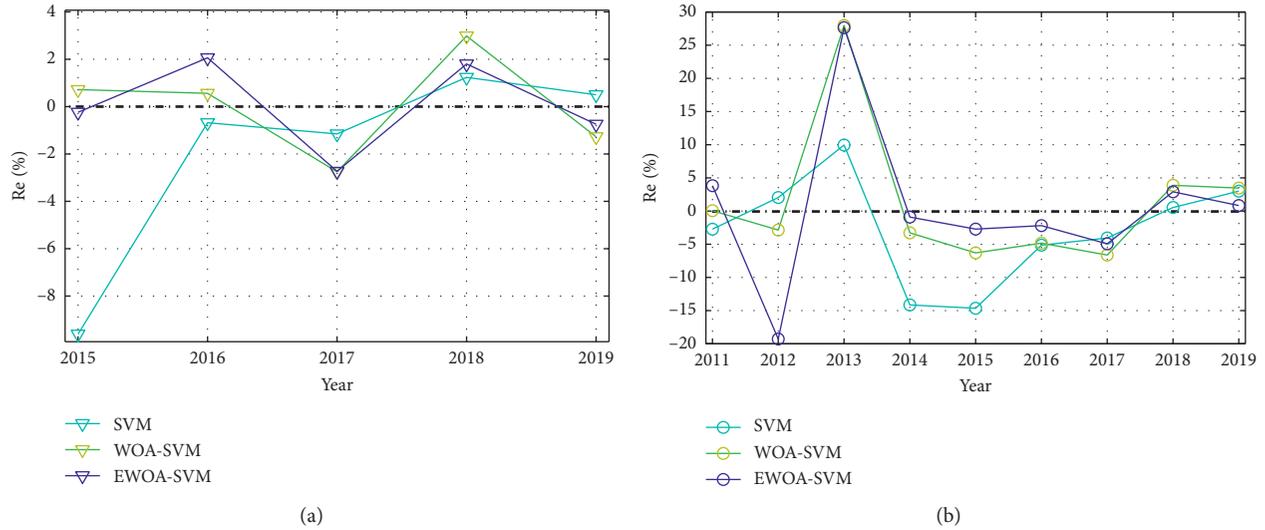


FIGURE 9: Prediction error curves. (a) Prediction relative errors in Case 1. (b) Prediction relative errors in Case 2.

TABLE 8: Forecast error values.

Case	Year	Re (%)		
		SVM	WOA-SVM	EWOA-SVM
Case 1	2015	-9.61	0.71	-0.23
	2016	-0.68	0.56	2.06
	2017	-1.15	-2.75	-2.73
	2018	1.23	2.98	1.80
	2019	0.50	-1.27	-0.74
Case 2	2011	-2.72	0.10	3.82
	2012	2.04	-2.85	-19.28
	2013	9.95	27.96	27.63
	2014	-14.14	-3.27	-0.90
	2015	-14.64	-6.30	-2.71
	2016	-5.13	-4.83	-2.19
	2017	-4.05	-6.62	-4.93
	2018	0.53	3.88	2.91
	2019	3.01	3.46	0.83

TABLE 9: Evaluation results of each model.

Cases	Models	rmse	r_2 (%)
Case 1	SVM	0.97	99.27
	WOA-SVM	0.59	98.23
	EWOA-SVM	0.51	98.69
Case 2	SVM	1.53	97.79
	WOA-SVM	1.48	98.30
	EWOA-SVM	1.26	98.42

the impetus of the information technology. As E-commerce has the characteristics of wide transaction coverage, low cost, fast information circulation, and high work flow coordination, it has become a new engine for economic development. Accordingly, E-commerce transaction trend is becoming an important indicator to measure the business or economic activity level. To this end, this study proposes the EWOA-SVM model to predict the trend of E-commerce

transactions, which provides a theoretical and effective tool for E-commerce development.

In real applications, a precise E-commerce transaction trend prediction can provide a decision-making basis for the governments or enterprises to formulate relevant development policies or plans in future business or industrial investments. The proposed model in this study can mine the crucial factors with high correlation degree in E-commerce transactions and construct E-commerce transaction trend correlation indexes. Importantly, it can be applied to logistics enterprises, Internet enterprises, and other information network companies in their business behavior.

5. Conclusions

In this study, the model training data was collected from E-commerce transactions volume and the influence factors, e.g., A1–A7, between 2005 and 2019. This sufficient data support and robust EWOA network structure can effectively alleviate the overfitting problem. The evaluation results of each model have been given more evidence with discussion to clarify this issue. The main contributions in this paper are concluded as follows:

- (1) A dynamic search coefficient and search updating strategy are combined to solve WOA algorithm's limitations. Accordingly, the EWOA algorithm can reach the global optima, i.e., 0, for multimodal functions, indicating a strong ability to escape from local minima.
- (2) The express delivery, Internet penetration rate, number of CN domain names, and number of Internet users are confirmed as the most crucial factors in the E-commerce transactions trend.
- (3) The evaluation results demonstrate that the EWOA-SVM model is superior to existing algorithms in the prediction of E-commerce transaction trend. For example, rmse of the EWOA-SVM model for Case 1

is 13.56% smaller than that of the WOA-SVM model and 47.42% smaller than that of SVM model. In Case 2, the rmse of the EWOA-SVM model is 14.86% smaller than that of the WOA-SVM and 17.64% smaller than that of the SVM model.

In the future work, it suggests that additional influencing factors in E-commerce transaction trend may be extended in practical circumstances. Besides, the generalization ability for various data prediction may be improved further.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- [1] C. Wang, W. Fu, J. Jin, Q. Shang, X. Luo, and X. Zhang, "Differential effects of monetary and social rewards on product online rating decisions in e-commerce in China," *Frontiers in Psychology*, vol. 11, p. 1440, 2020.
- [2] C. Leong, S. L. Pan, S. L. Pan, S. Newell, and L. Cui, "The emergence of self-organizing e-commerce ecosystems in remote villages of China: a tale of digital empowerment for rural development," *MIS Quarterly*, vol. 40, no. 2, pp. 475–484, 2016.
- [3] T. Teo and J. Liu, "Consumer trust in e-commerce in the United States, Singapore and China," *Omega*, vol. 35, no. 1, pp. 22–38, 2007.
- [4] C. Yoon, "The effects of national culture values on consumer acceptance of e-commerce: online shoppers in China," *Information & Management*, vol. 46, no. 5, pp. 294–301, 2009.
- [5] X. Yang, S. Mao, H. Gao, Y. Duan, and Q. Zou, "Novel financial capital flow forecast framework using time series theory and deep learning: a case study analysis of Yu'e Bao transaction data," *IEEE Access*, vol. 7, pp. 70662–70672, 2019.
- [6] T. Huang, R. Fildes, and D. Soopramanien, "The value of competitive information in forecasting FMCG retail product sales and the variable selection problem," *European Journal of Operational Research*, vol. 237, no. 2, pp. 738–748, 2014.
- [7] Ö. G. Ali, S. Sayın, T. Van Woensel, and J. Fransoo, "SKU demand forecasting in the presence of promotions," *Expert Systems with Applications*, vol. 36, no. 10, pp. 12340–12348, 2009.
- [8] I. Guven and F. Simsir, "Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM) methods," *Computers & Industrial Engineering*, vol. 147, 2020.
- [9] P. Liu, W. Ming, and B. Hu, "Sales forecasting in rapid market changes using a minimum description length neural network," *Neural Computing and Applications*, vol. 33, 2020.
- [10] R. Y. K. Lau, W. Zhang, and W. Xu, "Parallel aspect-oriented sentiment analysis for sales forecasting with big data," *Production and Operations Management*, vol. 27, no. 10, pp. 1775–1794, 2018.
- [11] B. Zhang, R. Tan, and C. J. Lin, "Forecasting of e-commerce transaction volume using a hybrid of extreme learning machine and improved moth-flame optimization algorithm," *Applied Intelligence*, vol. 51, 2020.
- [12] S. W. Ji, X. J. Wang, W. P. Zhao, and D. Guo, "An application of a three-stage XGBoost-based model to sales forecasting of a cross-border e-commerce enterprise," *Mathematical Problems in Engineering*, vol. 15, 2019.
- [13] I.-F. Chen and C.-J. Lu, "Sales forecasting by combining clustering and machine-learning techniques for computer retailing," *Neural Computing and Applications*, vol. 28, no. 9, pp. 2633–2647, 2017.
- [14] G. Di Pillo, V. Latorre, S. Lucidi, and E. Procacci, "An application of support vector machines to sales forecasting under promotions," *A Quarterly Journal of Operations Research*, vol. 14, no. 3, pp. 309–325, 2016.
- [15] S. Mao, X. Wang, and M. Zhu, "A new coupled ARMA-FGM model and its application in the internet third-party payment forecasting in China," *Grey Systems: Theory and Application*, vol. 8, no. 2, pp. 181–198, 2018.
- [16] L. L. Li, X. Zhao, M. L. Tseng, and R. R. Tan, "Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm," *Journal of Cleaner Production*, vol. 242, pp. 1–12, 2020.
- [17] Z. F. Liu, L. L. Li, M. L. Tseng, R. R. Tan, and K. B. Aviso, "Improving the reliability of photovoltaic and wind power storage systems using least squares support vector machine optimized by improved chicken swarm algorithm," *Applied Sciences-Basel*, vol. 9, no. 18, pp. 1–18, 2019.
- [18] B. T. Pham, L. M. Le, T. T. Le et al., "Development of advanced artificial intelligence models for daily rainfall prediction," *Atmospheric Research*, vol. 237, pp. 1–15, 2020.
- [19] M. S. Hossain and G. Muhammad, "Emotion recognition using deep learning approach from audio-visual emotional big data," *Information Fusion*, vol. 49, pp. 69–78, 2019.
- [20] C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Systems with Applications*, vol. 31, no. 2, pp. 231–240, 2006.
- [21] L. L. Li, Z. Y. Cen, M. L. Tseng, Q. Shen, and M. H. Ali, "Improving short-term wind power prediction using hybrid improved cuckoo search arithmetic-support vector regression machine," *Journal of Cleaner Production*, vol. 279, pp. 1–15, 2021.
- [22] C.-C. Chang and C.-J. Lin, "Libsvm," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.
- [23] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
- [24] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: a review," *Isprs Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, 2011.
- [25] L. L. Li, Y. B. Chang, M. L. Tseng, J. Q. Liu, and M. K. Lim, "Wind power prediction using a novel model on wavelet decomposition-support vector machines-improved atomic search algorithm," *Journal of Cleaner Production*, vol. 270, 2020.
- [26] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.
- [27] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm," *Computers & Structures*, vol. 169, pp. 1–12, 2016.

- [28] B. Abed-Alguni, "Island-based cuckoo search with highly disruptive polynomial mutation," *International Journal of Artificial Intelligence*, vol. 17, no. 1, pp. 57–82, 2019.
- [29] A. Soares, R. Rábalo, and A. Delbem, "Optimization based on phylogram analysis," *Expert Systems with Applications*, vol. 78, pp. 32–50, 2017.
- [30] R.-E. Precup, E.-L. Hedrea, R.-C. Roman, E. M. Petriu, A.-I. Szedlak-Stinean, and C.-A. Bojan-Dragos, "Experiment-based approach to teach optimization techniques," *IEEE Transactions on Education*, vol. 64, no. 2, pp. 88–94, 2021.
- [31] H. J. Zhang, Z. Xie, H. C. Lin, and S. Y. Li, "Power capacity optimization in a photovoltaics-based microgrid using the improved artificial bee colony algorithm," *Applied Sciences-Basel*, vol. 10, pp. 1–21, 2020.
- [32] Y. D. Ren, H. Li, and H. C. Lin, "Optimization of feedforward neural networks using an improved flower pollination algorithm for short-term wind speed prediction," *Energies*, vol. 12, pp. 1–17, 2019.
- [33] C. Fu, Q. Lv, and H.-C. Lin, "Development of deep convolutional neural network with adaptive batch normalization algorithm for bearing fault diagnosis," *Shock and Vibration*, vol. 2020, Article ID 8837958, 10 pages, 2020.
- [34] H. Zapata, N. Perozo, W. Angulo, and J. Contreras, "A hybrid swarm algorithm for collective construction of 3D structures," *International Journal of Artificial Intelligence*, vol. 18, no. 1, pp. 1–18, 2020.
- [35] R. E. Precup, R. C. David, R. C. Roman, E. M. Petriu, and A. I. Szedlak-Stinean, "Slime mould algorithm-based tuning of cost-effective fuzzy controllers for servo systems," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, 2021.
- [36] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [37] M. M. Mafarja and S. Mirjalili, "Hybrid whale optimization algorithm with simulated annealing for feature selection," *Neurocomputing*, vol. 260, pp. 302–312, 2017.
- [38] I. Aljarah, H. Faris, and S. Mirjalili, "Optimizing connection weights in neural networks using the whale optimization algorithm," *Soft Computing*, vol. 22, no. 1, pp. 1–15, 2018.
- [39] M. A. E. Aziz, A. A. Ewees, and A. E. Hassanien, "Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation," *Expert Systems with Applications*, vol. 83, pp. 242–256, 2017.
- [40] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," *Applied Soft Computing*, vol. 62, pp. 441–453, 2018.
- [41] L. L. Li, Y. W. Liu, M. L. Tseng, G. Q. Lin, and M. H. Ali, "Reducing environmental pollution and fuel consumption using optimization algorithm to develop combined cooling heating and power system operation strategies," *Journal of Cleaner Production*, vol. 247, pp. 1–13, 2020.
- [42] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997.
- [43] Z. F. Liu, S. F. Luo, M. L. Tseng, H. M. Liu, L. L. Li, and A. H. M. D. Mashud, "Short-term photovoltaic power prediction on modal reconstruction: a novel hybrid model approach," *Sustainable Energy Technologies and Assessments*, vol. 45, 2021.
- [44] M.-Y. Cheng and D. Prayogo, "Symbiotic organisms search: a new metaheuristic optimization algorithm," *Computers & Structures*, vol. 139, pp. 98–112, 2014.
- [45] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems," *Engineering with Computers*, vol. 29, no. 1, pp. 17–35, 2013.
- [46] I. M. Alecu, J. Zheng, Y. Zhao, and D. G. Truhlar, "Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries," *Journal of Chemical Theory and Computation*, vol. 6, no. 9, pp. 2872–2887, 2010.
- [47] H. Nguyen, X.-N. Bui, H.-B. Bui, and N.-L. Mai, "A comparative study of artificial neural networks in predicting blast-induced air-blast overpressure at Deo Nai open-pit coal mine, Vietnam," *Neural Computing and Applications*, vol. 32, no. 8, pp. 3939–3955, 2020.