IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 23, 2019, accepted December 29, 2019, date of publication January 1, 2020, date of current version January 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963540

Research on Regional Logistics Demand Forecast Based on Improved Support Vector Machine: A Case Study of Qingdao City under the New Free Trade Zone Strategy

NAN YU [1], WEI XU [1], AND KAI-LI YU [2]
[1]College of Transportation, Shandong University of Science and Technology, Qingdao 266590, China
[2]Department of Economics and Management, Technology Vocational College of Dezhou, Qingdao 266232, China

Corresponding author: Wei Xu (xuwei972@163.com)

This work was supported in part by the Joint Funding of the Shandong Social Science Planning Research Project under Grant 18CCXJ25, in part by the Qingdao Social Science Planning Research Project under Grant QDSKL1801134, and in part by the Qingdao Shuangbai Research Project under Grant 2019-B-10.

ABSTRACT Based on the analysis on the influencing factors of urban logistics demand, this paper, taking into account the logistics demand with non-linear and small sample modeling characteristics from the perspective of urban freight volume, introduces the ant colony algorithm into the modeling process to optimize the penalty parameter "c" and "g" parameter of Radial Basis Function in support vector machine, and has made a prediction to the logistics demand of Qingdao with the optimized support vector machine model. The experimental results show that the prediction results of the improved support vector machine can bring the prediction closer to the reality with their more accuracy, stronger stability and less error rate, thus providing a guarantee for the logistics demand forecast of Qingdao.

INDEX TERMS Ant colony algorithm, logistics demand forecast, logistics engineering, support vector machine, urban logistics demand.

I. INTRODUCTION

Urban logistics demand forecasting is based on the analysis of relationship between supply and demand of logistics. By analyzing relevant factors affecting logistics demand, We have made a prediction on the indicators reflecting logistics demand according to objective information and data. Under the new situation, scientific forecasting of urban logistics demand is an important basis for rationally formulating logistics development policies, strengthening logistics infrastructure construction, and building a logistics service system. Therefore, timely mastering the information of effective logistics demand's development trend is beneficial to the optimization and improvement of the modern logistics system of the city for realizing the goal of reducing costs and increasing efficiency of the logistics, which has an important guiding significance for improving the economic competitiveness of the city and promoting the sustainable development of the city.

The associate editor coordinating the review of this manuscript and approving it for publication was Guanjun Liu.

The research on logistics in foreign countries started earlier, and foreign scholars have made certain researching achievements on the logistics demand forecasting. Samvedi and Jain [1] conducted a simulation experiment on beer by comparing the performance of three established forecasting methods and gray forecasting methods in such a situation that is interrupted and stable. The results show that the gray prediction method has the best stability. Jaipuria and Mahapatra [2]used the DWT-ANN model in the forecasting study of regional logistics demand to reduce the inventory cost resulted by the bullwhip effect, and the DWT-ANN model and ARIMA were taken as examples of three different local manufacturing companies. The results of the model predictions are compared and the results verify the validity of the DWT-ANN model. Zendehboudi *et al.* [3] proposed a new hybrid support vector machine model to predict solar and wind energy resources. Experimental results show that this method is more accurate than those of other models. Currently, domestic research on urban logistics demand forecasting is in its infancy. As localities pay more attention to the development of logistics industry, some scholars realize the impact of urban logistics development on economic growth. In the research, Ge *et al.* [4] uses a new R-BFGS optimization method based on Riemannian manifold to estimate the GMM parameters, and verifies the applicability of the model. Liu [5] proposed a combination forecasting method based on Shapley value, and analyzed the index data of Shandong Province from 2005 to 2014 by giving reasonable weights and comparing three quantitative analysis models. Guo *et al.* [6] proposed a SVM-LSTM combined forecasting model based on real-time prediction error to predict the passenger flow of urban rail transit. The experimental results show that the combined forecasting model not only accurately reflects the abnormal fluctuation of passenger flow, but also has higher Forecast accuracy. Sun [7] optimized the initial parameters of gray neural network by using particle swarm optimization algorithm, and established a gray neural network prediction model based on particle swarm optimization. Taking Beijing as an example, combining the characteristics of logistics demand is to analyze Beijing's Logistics demands. Li *et al.* [8] first used wavelet to decompose the sequence and sequence to reduce the unsteady load order. In order to obtain the optimal parameters of the second-order gray prediction model, the paper uses the neural network mapping method to establish a second-order gray neural network prediction model. It shows that the proposed method can effectively improve the accuracy of load prediction. Wang *et al.* [9] adopts fuzzy analytic hierarchy process (AHP), first optimizes parameters σ and c, improves the SVM-based prediction model, and then uses the optimal parameters to train SVM and establish a prediction model to predict network traffic. Sujjaviriyasup [10] uses the fractional auto-regressive integral moving average method to predict vibration trends and tire spectral characteristics, and this method can not only be used to monitor driving performance but also improve driving comfort. In conclusion, studies home and abroad focus mainly on the mathematical models used in the analysis of logistics systems but are limited to the improvement of the forecasting methods of urban logistics. As to the selection of research objects and research methods, few studies concentrates on algorithms based on the specific urban logistics development to analyze the logistics demand forecast.

Based on the above research achievements, it is considered that the logistics system is a complex nonlinear system. The article focuses on the characteristics of fewer data samples, lacking of internal connection and regularity. By comparing various prediction methods, combination with support vector machine has unique advantages in solving problems such as sample finite, nonlinear function and multidimensional pattern recognition. The support vector machine algorithm is applied to the the forecasting research of urban logistics demand by detecting the classification accuracy of four different kernel functions, the RBF kernel function with the highest classification accuracy is found, and the c and g parameters affecting the prediction accuracy of support vector machine are optimized by ant colony algorithm to establish improved the prediction model of support vector machine. Compared with other prediction methods, ACO-SVM prediction method is more accurate to 97.28%. It can not only consider the influence of other factors on the prediction object, but also has the advantage of improving the prediction accuracy degree, in the prediction result, people can get more reliable and accurate results that support Qingdao's logistics development in data aspect.

II. INDEX SYSTEM CONSTRUCTION OF URBAN LOGISTICS DEMAND FORECASTING

The accurate selection of urban logistics demand forecasting index is the key step to successful implementation of forecasting. In order to ensure that the forecast results of logistics demand are closer to the actual outcome, this paper analyzes the complex and diverse factors affecting urban logistics demand. Based on the investigation of relevant texts and the actual situation, this paper analyzes the influencing factors of urban logistics demand from the perspectives of economic influence and non-economic influence.

Economic influence factor:

(1) The scale of urban economic development. Since the ''13th Five-Year Plan'', with the development of urban economy, the growth rate of logistics demand has accelerated, and the development speed of urban economy is closely related to the development of logistics industry.

(2) Urban industrial structure. Adjusting the urban industrial structure can have a major impact on the size and structure of logistics demand. As a regional center, the city usually has the same industrial structure as the local area. To stabilize a city's industry, a city must ensure that the service industry can achieve stable, positive, and long-term development.

(3) Commercial trade. Commercial trade has a major impact on the logistics needs of the city, and the development of commercial trade also requires high standards of urban logistics support. For the import and export volume of goods, the trend of foreign trade volume must be taken into account when forecasting urban logistics demand. The overall retail sales of urban social consumption products can appropriately present the activity of their business and trade flows, reflecting the scale of logistics demand.

(4) Residents consumption level. With the development of urbanization, progress has been made in e-commerce distribution services with the increase of both the consumption and purchasing power of local people. Online stores and third-party logistics companies have cooperated to obtain more business opportunities. And urban goods distribution services will directly affect the development of urban logistics demand. The development trend of logistics demand.

(5) Capital investment. The capital investment affects the logistics demand to a large extent. The city needs to optimize resource allocation and policy guidance, and the capital investment environment, and improve the logistics service level.

TABLE 1. Urban logistics demand forecasting indicator set.

Influence Factors	Detailed Index	Unit	Symbol
Scale of economic development	GDP	100million	X_1
industrial structure	Added value of primary industry	100million	X_2
Commercial trade	total retail sales of consumer goods	100million	X_3
income level	urban per capita disposable income	yuan	X_4
capital input	Total fixed asset investment of the whole society	100million	X_5
consumer market	Total population	Ten thousand	X_6

(6) Consumer market. The consumer market refers to the material exchange and economic activities that people carry out in the social environment. The size of the consumer market is limited by the population density. In areas with high population densities, the increase in commercial circulation has led to the occurrence of logistics activities.

Due to the factors affecting urban logistics demand such as non-economic factors, macroeconomic policies and external environment, the demand for urban logistics in the short term has little impact and thus can be ignored. Therefore, this paper only focuses on the relationship between economic indicators and urban logistics demand, and uses freight volume to measure the scale of urban logistics demand. In the actual forecasting study, according to the availability of statistical data and the different statistical data of each region, the influencing factor index system shown in Table 1 below was constructed.

III. PREDICTIVE MODEL CONSTRUCTION

A. SUPPORT VECTOR MACHINE PRINCIPLE

Support Vector Machines (SVM) is a newly advanced technology, which, based on statistical learning theory, uses the principle of structural risk minimization to avoid local minimums to effectively solve over-learning problem. It can obtain meaningful law information, ensure people acquire generalization ability, and possess better prediction accuracy degree. The SVM method realizes the principle of minimizing structural risk. Aiming to find the unique solution by using the quadratic programming method to construct the optimal separation hyperplane in the high-dimensional feature space, thereby reducing the boundaries of classification errors. One of the core ideas of Support Vector Machines is the control of generalization ability. Statistical learning theory indicates that the optimal hyperplane has the best generalization performance, and the problem of solving the optimal hyperplane is transformed into Optimization problem. The optimal hyperplane can be represented by the classification function $f(x) = w^T x + b$ in the case of linear separability. You can use $y(w^T x + b)$ to indicate the correctness and faithfulness of the classification. The function interval is as follows:

$$\hat{\gamma}_i = y_i(w^T x_i + b) \tag{1}$$

Regarding the maximum geometric spacing of feature spaces, the constraint optimization problem is as follows:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \tag{2}$$

$$y_i = (w^T x_i + b) - 1 \geq 0, \quad i = 1, 2, \ldots, n \tag{3}$$

x_i represents the i-th point on the linear separable hyperplane, and b determines the position of the point on the hyperplane relative to the origin, y indicates the classification of the linear sample data points, and y_i indicates the linear classification category corresponding to the i-th data point. Where w is the normal of the hyperplane, when $f(x)$ is equal to 0, and $f(x)$ is greater than 0 corresponds to the data point of $y = 1$, $f(x)$ is less than 0 The point corresponds to the point of $y = -1$.

Considering that most of the classification problems in practical applications are linear (or nonlinear) and inseparable, it is difficult to find a hyperplane that can distinguish singular points without errors. By allowing a certain degree of error classification, the soft margin method introduces a slack variable ε that measures the degree of classification of the error, and a cost constant c whose important effect of weighted classification error on edge width [11]. For the nonlinear case, the input data is mapped to the high-dimensional

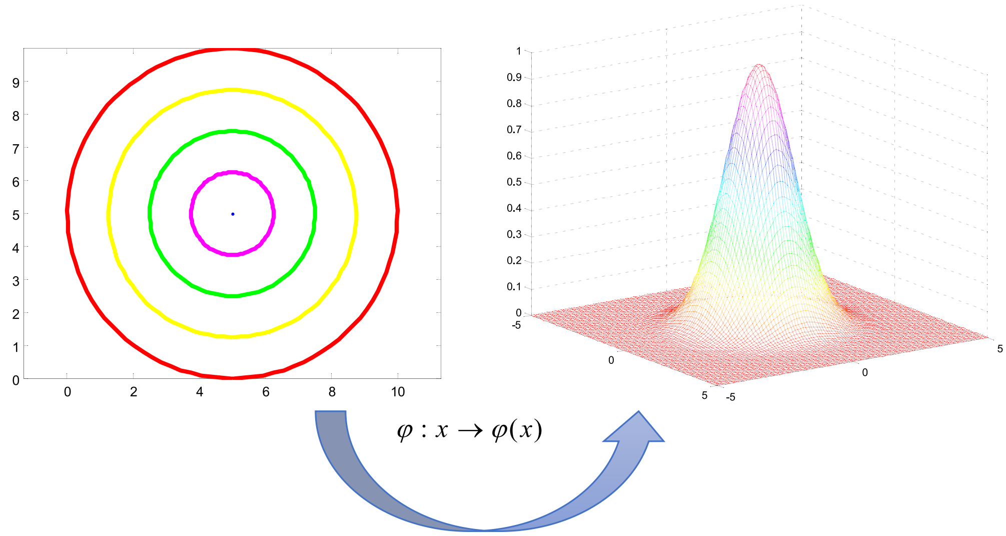

FIGURE 1. Support vector machine nonlinear map.

feature space by using the nonlinear map a + b = c, and finally the optimal hyperplane is constructed in the high-dimensional feature space, so that the nonlinear data on the plane is separated, as shown in Fig. 1 Shown.

Being generality, a slack variable ε_i is introduced for each sample point to mark x_i, at the same time the corresponding objective function becomes:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + c \sum_{i=1}^{R} \varepsilon_i \tag{4}$$

$$s.t.\ y_i(w^T x_i + b) \geq 1 - \varepsilon_i \tag{5}$$

For the convex quadratic optimization problem, the objective function and constraints are integrated into the Lagrangian function by introducing the Lagrangian multiplier, which is convenient for solving the maximum or minimum value Introducing the Lagrangian multiplier α for each inequality constraint, the Lagrangian function is obtained as follows:

$$\begin{aligned} &\min_{\alpha} \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n} \alpha_i \\ &\text{s.t. } c \geq \alpha_i \geq 0, \quad i = 1, 2, \ldots, n \\ &\sum_{i=1}^{n} \alpha_i y_i = 0 \end{aligned} \tag{6}$$

Obviously, the paper completes the design of linear decision surface, and classifies the training set in high-dimensional feature space. The high-dimensional feature space can be the input space itself or its nonlinear transformation space, which form linear and nonlinear SVM respectively.

B. PRINCIPLE OF NUCLEAR FUNCTION

Kernel function is the key technology of support vector machine, which can effectively solve the "dimension mutation problem" in the high-dimensional feature space operation of traditional methods in dealing with nonlinear problems. Given a kernel function, a feature space associated with it can be constructed so that the inner product in the space can be represented by it. Let X be a tight subset of the n-dimensional space F^n, $K(x, y)$ is a kernel function defined on $X \times X$, and there is a nonlinear mapping $\varphi : x \to \varphi(x)$ from X to the high-dimensional feature space H, such that:

$$K(x, y) = \varphi(x)^T \varphi(y) \tag{7}$$

1) CHOICE OF KERNEL FUNCTION

In practical applications, depending on the difference of problems and the sample datas, different kernel functions need to be selected, so that different support vector machine algorithms are obtained. Kernel functions that are used frequently are as follows [12]:

(1) Linear kernel function: $K(x, y) = x^t y$;

(2) Radial Basis Function (RBF) kernel function: $K(x, y) = \exp(-\frac{\|x-y\|^2}{2\sigma^2})$;

(3)Polynomial kernel function: $K(x, y) = (sx^t y + 1)^d$;

(4)Sigmoid kernel function: $K(x, y) = \tanh(ax^t y + b)$.

Since the kernel function affects the computational accuracy and classification speed of the support vector machine algorithm, the kernel function corresponding to the optimal precision can be selected through multiple trainings, when solving different problems. Therefore, the article uses four common kernel functions to nonlinearly classify the training sample data. The accuracy results obtained by the test are as follows:

The first set of data: $x = [1, 0, -1, 2, 0, 1]$; $y = [1, 1, -1, 2, 0, 1]$; $z = [-1, 1, 1, -1, 1, 1]$
where (x, y) represents a two-dimensional data point and z is a type attribute of the corresponding point.

Accuracy rate: linear kernel function 0.6667, RBF kernel function 0.8333, polynomial kernel function 0.6667,

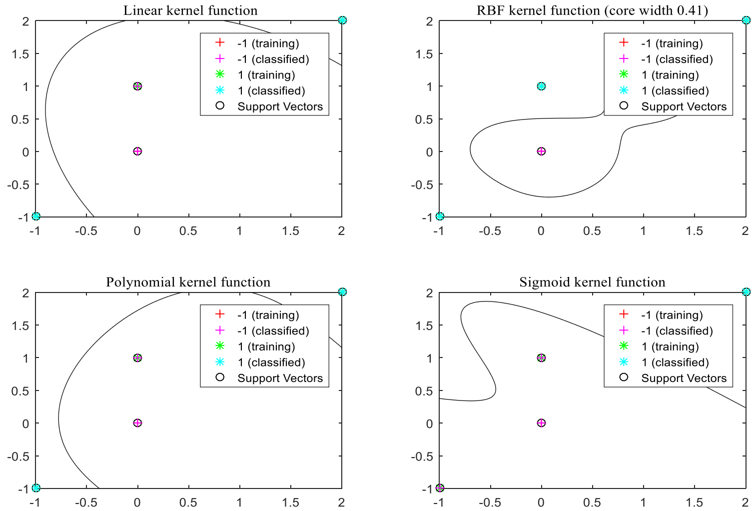

FIGURE 2. Classification of the first set of data by four common kernel functions.

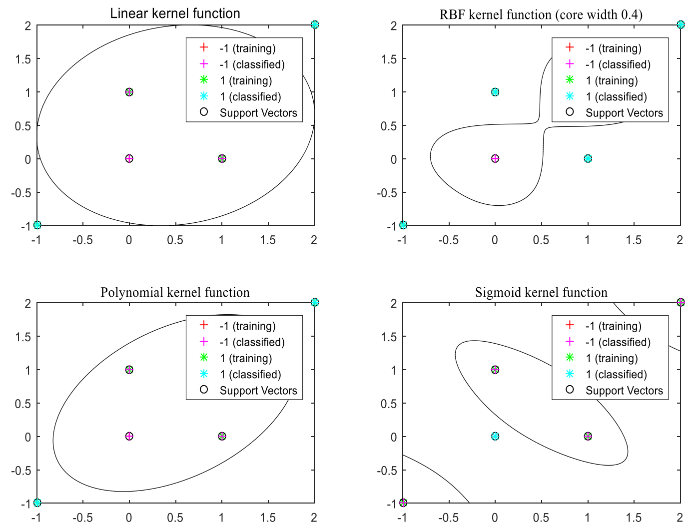

FIGURE 3. Classification of the second of data by four common kernel functions.

Sigmoid kernel function 0.5000. The second set of data: $x = [1, 0, -1, 2, 0, 1]$; y = [0,1,-1,2,0,1]; z = [-1,1,1,-1,1,1]

Accuracy rate: linear kernel function 0.6667, RBF kernel function 1, polynomial kernel function 0.6667, Sigmoid kernel function 0.

Since the RBF kernel function has high classification accuracy for the test set data, it has a high degree of flexibility. Therefore, the article selects the RBF kernel function to handle the nonlinear separability of the support vector machine.

Since the calculation of equation (6) only needs to calculate the dot product $x_i^T x_j$ in the high-dimensional feature space, so that $x_i^T x_j = \varphi^T(x_i)\varphi(x_j) = K(x_i, x_j)$, by using the appropriate kernel function, not only can the computation amount be greatly reduced, but also calculate the value of the kernel function on the sample set in the high-dimensional feature space, the convex quadratic optimization problem for α can be expressed as:

$$\begin{aligned} &\min_{\alpha} \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i\alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^{n}\alpha_i \\ &\text{s.t.} \sum_{i=1}^{n}\alpha_i y_i = 0, \quad c \geq \alpha_i \geq 0,\ i = 1, 2, \ldots, n \end{aligned} \tag{8}$$

C. CONSTRUCTION OF ACO-SVM MODEL

1) BASIC PRINCIPLES OF ANT COLONY ALGORITHM

Ant Colony Optimization (ACO) is a new type of heuristic evolutionary algorithm. It is a group of non-intelligent or lightly intelligent individuals that cooperate to express intelligent behaviors, thus it provides a new algorithm to solve complex problems [13]. If the number of all ants

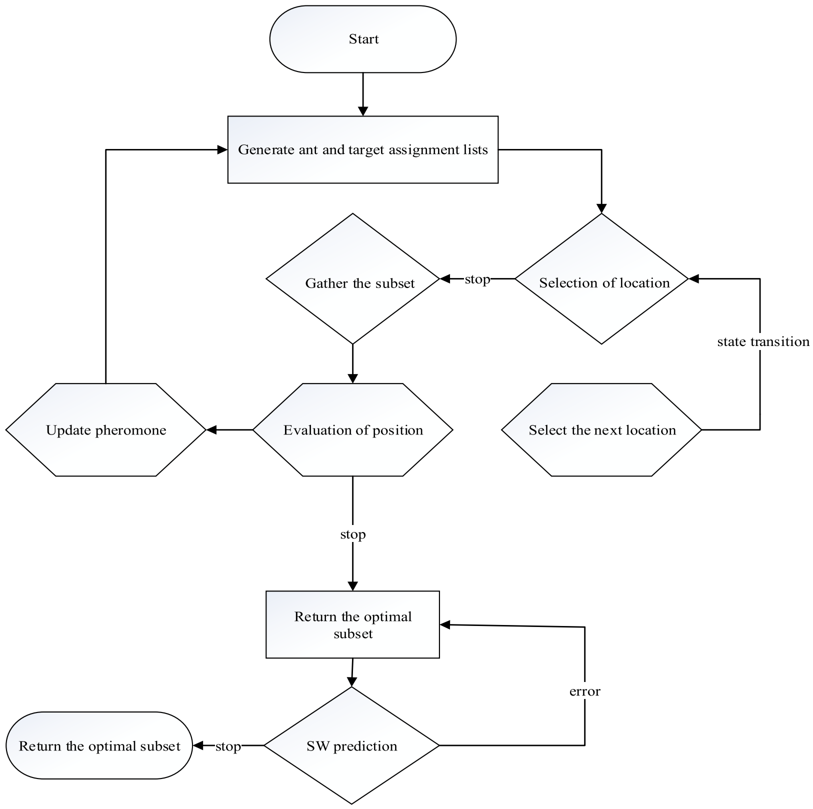

FIGURE 4. ACO-SVM prediction flow chart.

in the ant colony is m, the pheromone volatilizes factor ρ, and each ant k defines a one-dimensional array $path_k$ with n elements. The ordinates of the n nodes through which the kth ant passes are sequentially stored in the $path_k$, and can be used to represent the crawling path of the kth ant, where n is the total effective bit of the optimized parameter. Let the time counter $t = 0$, the number of cycles $N = 0$, and set the maximum number of cycles $N_{\max}$.

2) ANT COLONY ALGORITHM OPTIMIZATION SUPPORT VECTOR MACHINE

The parameters of the support vector machine determine its learning ability and generalization ability. It can be understood from the front that the support vector machine algorithm has two very important parameters c and g, c is the penalty coefficient, that is, the tolerance to the error, to determine the trade-off cost between minimizing training error and SVM model complexity. If c is too large, the model learning and its complexity will increase, and the model will easily fall into the phenomenon of "over-fitting"; if it is too small, the complexity of the model will be too low, and the model will fall into the "under-fitting" problem. g is the RBF function as a parameter after the kernel, g defines the nonlinear mapping from the input space to the high-dimensional feature space, which affects the shape of the RBF function [14], the g value is inversely proportional to the support vector, the number of support vectors affects the training and prediction speed of support vectors machine algorithm. In this paper, the ant colony algorithm is used to find the optimal parameters of SVM in a certain range, which avoids the blindness of manual selection of support vector machine parameters. What's more, it also improves the accuracy and adaptability of support vector machine prediction model. The specific process of ACO optimization SVM is shown in the figure 4 is shown.

In order to verify the effect of ant colony algorithm optimization, the simulation experiments were first carried out on two sets of sample data sets. On the right-angle plane xoy, the points in the circle $x^2+y^2 = 0.16$ are defined as one class, the points in the circle $0.16 < x^2 + y^2 < 0.64$ are defined as another class, and 314 samples of two types of samples are randomly generated for classification accuracy test.

In order to visually explain the influence of parameters on the classification results, we respectively test and determine the classification accuracy rate of c values and different g values, as shown in table 2. Table about determining the classification accuracy rate of g values and different c values is shown.

It can be seen from Table 2 and Table 3 that under the same test sample data set, the classification accuracy of different c and g parameters are very different. Therefore, the article constructs the ACO-SVM optimization mathematical model. For a given sample data set of the two category, the support vector machine classification accuracy rate can be regarded

TABLE 2. When $c = 1$, the effect of g on classification accuracy.

g	Classification Accuracy
0.00005	48.95%
0.005	65.02%
0.5	91.65%
50	59.07%

TABLE 3. When $g = 1$, the effect of c on classification accuracy.

c	Classification accuracy
0.1	92.33%
1	89.86%
10	76.94%
100	75.03%

as a binary function about c and g, denoted as $L(c, g)$, then the ACO-SVM optimization mathematical model is :

$$\begin{aligned} &\max L(c, g) \\ &s.t.\ c \in (0, a), \quad g \in (0, b) \end{aligned} \tag{9}$$

Assuming $a = 5000$ and $b = 0.0000125$, the parameters in the ant colony algorithm include the information heuristic factor γ and the expected heuristic factor β, where the pheromone constant $Q = 2$, the pheromone volatilization factor $\rho = 0.2$, the number of ant colonies $m = 20$, the maximum number of cycles $N_{\max} = 30$. After setting the above parameters, according to the given 314 sample data sets, the ACO-SVM optimization model can be solved, and the highest classification accuracy (c, g) combination can be obtained. The parameter optimization results and classification accuracy rate are as follows Table 4.

Comparing the results of the above table classification, it can be seen that the classification accuracy of the ACO-SVM model parameters in Table 4 is better than other combinations. In general, for the sample data sets that is inseparable and very difficult the ACO-SVM model improves the classification accuracy more obviously.

IV. EMPIRICAL ANALYSIS

A. OVERVIEW OF QINGDAO LOGISTICS DEVELOPMENT

Qingdao's geographical location is superior and it is actively building into an international metropolis. It is a development zone for the development of China's modern marine industry and a major node city along the New Belt Eurasia Bridge Economic Corridor of the "Belt and Road". Qingdao is an important economic and trade gateway of Shandong Province. Logistics development has provided an important impetus for accelerating the implementation of the "two districts, one circle and one belt" regional development strategy put forward by the Shandong Provincial Government.

B. DATA ACQUISITION AND PREPROCESSING

According to the the index system of Qingdao logistics demand forecasting established in the second chapter of the article, this paper obtains the explanatory variables of the logistics demand refinement index set of Qingdao from 1999 to 2017 by consulting the literature and the national statistical yearbook and the Qingdao statistical information website. X6, the explanatory variable (freight volume) y, as shown in Table 5.

In order to eliminate the influence of the dimension between each forecasting indicator datum on the accuracy of Qingdao's logistics demand forecasting results, the raw data needs to be normalized before the forecasting of Qingdao's logistics demand. Here, the *map* min max function [15] is used. It normalizes and converts the sample data into data between [−1,1] as follows:

$$y_{ij} = \frac{(y_{\max} - y_{\min}) * (x_{ij} - x_{\min})}{(x_{\max} - x_{\min})} + y_{\min} \tag{10}$$

TABLE 4. ACO-SVM model parameter optimization results.

Sample Data Set	(c, g)	Accuracy
Classification	(0.12,0.48)	97.28%

TABLE 5. Statistics on Qingdao's logistics demand indicators from 1999 to 2017.

	Y	X_1	X_2	X_3	X_4	X_5	X_6
2003	30553	1780.42	148.92	636.42	10075	547.55	720.68
2004	35570	2163.80	163.49	756.50	11089	984.56	731.12
2005	37636	2695.82	178.33	965.32	12920	1403.3	819.55
2006	39294	3206.58	183.95	1006.70	15328	1485.69	829.42
2007	40758	3786.52	203.59	1199.20	17856	1635.36	838.67
2008	42484	4436.18	223.4	1464.80	20464	2019.01	845.61
2009	24408	4854.00	230.25	1730.20	22368	2458.89	850.03
2010	26971	5666.20	276.99	1961.10	24998	3022.48	871.51
2011	29184	6615.60	306.38	2302.40	28567	3502.54	879.51
2012	29246	7302.10	324.41	2635.60	32145	4153.91	886.85
2013	31318	8006.60	340.5	2986.80	35227	5027.86	896.41
2014	26061	8692.10	349.62	3361.70	37346	5766.03	904.62
2015	26965	9300.07	363.98	3713.70	40370	6555.67	909.7
2016	27955	10011.29	371.01	4104.90	43598	7454.7	920.4
2017	28418	11037.28	380.97	4541.00	47176	7777.1	929.05

where x_{ij} represents the original sample data, y represents the normalized data of the sample data, $y_{\max} = 1$ and $y_{\min} = -1$ are the maximum and minimum values of the set parameters, and the normalized data is shown in Table 6.

C. IMPLEMENTATION OF ACO-SVM PREDICTION MODEL

1) PREDICTION MODEL PARAMETER SETTING AND MODEL TRAINING

The prediction model of improved Support Vector Machine is a short time series prediction model, considering that the recent data in the time series has a relatively profound and lasting impact on the future. In order to verify the accuracy of the support vector machine's logistics demand forecast in Qingdao,and reduce the effect of the error, this paper takes the data of Qingdao City through 1999 to 2014 as the training set, and those through 2015 to 2017 as the test set, and forecast the study of the logistics demand of Qingdao City.

On the basis of the previous analysis, RBF is selected as the kernel function of the support vector machine, and the training set is trained by Matlab R2014a programming. The range of the penalty parameter c and the RBF kernel function parameter g is set to [0.1, 1000], the cross-validation

TABLE 6. Normalized data.

	Y	X_1	X_2	X_3	X_4	X_5	X_6
1999	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
2000	-0.626	-0.967	-0.978	-0.946	-0.963	-0.994	-0.968
2001	-0.055	-0.934	-0.949	-0.906	-0.927	-0.981	-0.934
2002	-0.177	-0.894	-0.926	-0.830	-0.928	-0.961	-0.888
2003	0.062	-0.842	-0.912	-0.761	-0.860	-0.914	-0.843
2004	0.457	-0.766	-0.792	-0.706	-0.809	-0.798	-0.751
2005	0.619	-0.660	-0.669	-0.612	-0.717	-0.687	0.031
2006	0.749	-0.558	-0.623	-0.594	-0.597	-0.665	0.119
2007	0.864	-0.443	-0.461	-0.507	-0.470	-0.626	0.200
2008	1.000	-0.313	-0.298	-0.387	-0.339	-0.524	0.262
2009	-0.421	-0.230	-0.242	-0.267	-0.244	-0.408	0.301
2010	-0.219	-0.069	0.143	-0.163	-0.112	-0.258	0.491
2011	-0.045	0.120	0.386	-0.009	0.067	-0.131	0.562
2012	-0.040	0.257	0.534	0.141	0.247	0.041	0.627
2013	0.122	0.397	0.667	0.299	0.401	0.272	0.711
2014	-0.291	0.533	0.742	0.468	0.507	0.468	0.784
2015	-0.220	0.654	0.860	0.627	0.659	0.677	0.829
2016	-0.142	0.796	0.918	0.803	0.821	0.915	0.923
2017	-0.106	1.000	1.000	1.000	1.000	1.000	1.000

is $t = 6$, the number of ant colonies is 50, and the iteration is 100. The convergence path of the ACO algorithm in the optimization process is shown in Figure 5:

The ACO-SVM prediction model is trained by using the penalty factor c after optimization and the parameter g of the kernel function. The fitting effect of the training results is shown in Figure 7.

It can be seen from the comparison chart that there is a small fluctuation between the predicted and actual value of Qingdao's freight volume in Qingdao from 1999 to 2014 predicted by the ACO-SVM prediction model, and the average percentage between the predicted value and the true value is calculated. The error MAPE is 0.0343, and the goodness of fit R^2 is 0.9747. The closer the value of R^2 is to 1, the better the fit of the regression line to the observed value. From the calculation process, the parameter $c = 271.8591$, $g = 3.5748$, The parameters are optimized and the training results are highly accurate.

2) MODEL TESTING AND COMPARATIVE ANALYSIS

In order to verify the accuracy of the model, this paper takes the corresponding index data of Qingdao City's

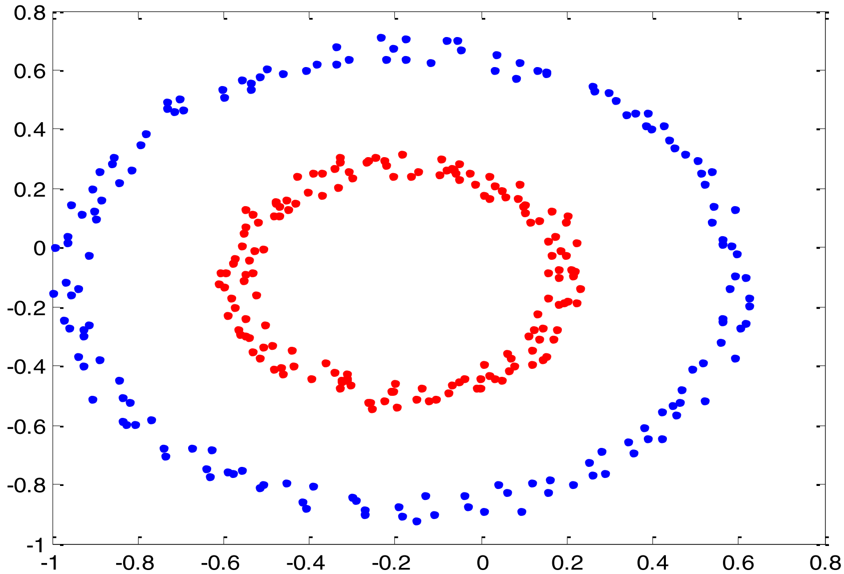

FIGURE 5. Classification problem training sample.

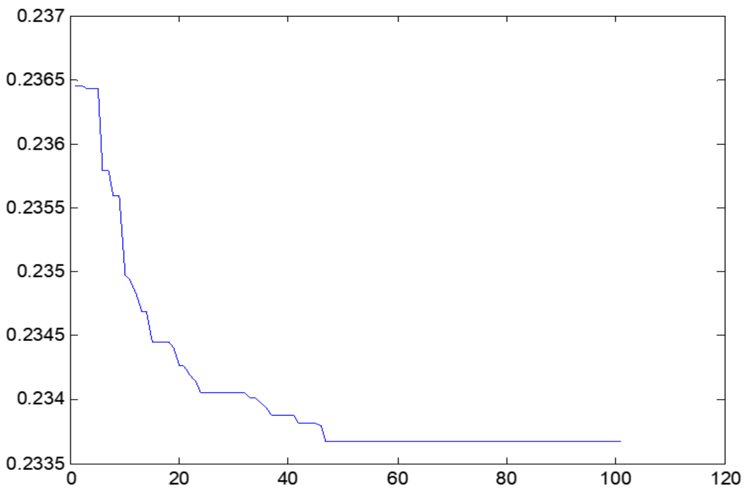

FIGURE 6. ACO algorithm convergence trajectory.

2015-2017 freight volume as a test set, and conducts an empirical study on the logistics demand of Qingdao City from 2015 to 2017. The model obtained by predicting the freight volume predicted by the model and the actual value is shown in Figure 8. In the figure, it can be seen that the difference between the predicted value and the true value is small, and the prediction effect is more ideal.

In order to illustrate the advantages and feasibility of the regression ACO-SVM model, we compare the prediction method of this paper with the prediction methods of other papers (RBF neural network prediction model, triple exponential smoothing-Logistic growth curve-BP neural network) [16]–[19] logistics demand of Qingdao, for RBF neural network, the systematic error is 0.01, the width of the kernel function σ^2 is 2, After normalizing all the sample data, the data is input into the ACO-SVM model and the RBF neural network for learning. As shown in Table 7, the root means square error RMSE of the ACO-SVM is 1084.86215, which is smaller than that obtained by the RBF neural network. The square root error indicates that the ant colony algorithm is feasible to optimize the parameters of the support vector machine.To explain why this addition makes sense, the paper sets other comparative prediction methods to test in the same environment. The comparative prediction results are shown in Table 7 below. Comparing the prediction results, the average relative error of the ACO-SVM prediction model is 0.038, which is better than the prediction results of other prediction methods. It is more accurate and the prediction results are more stable and reliable. It shows that the prediction model of improved support vector machine is meaningful.

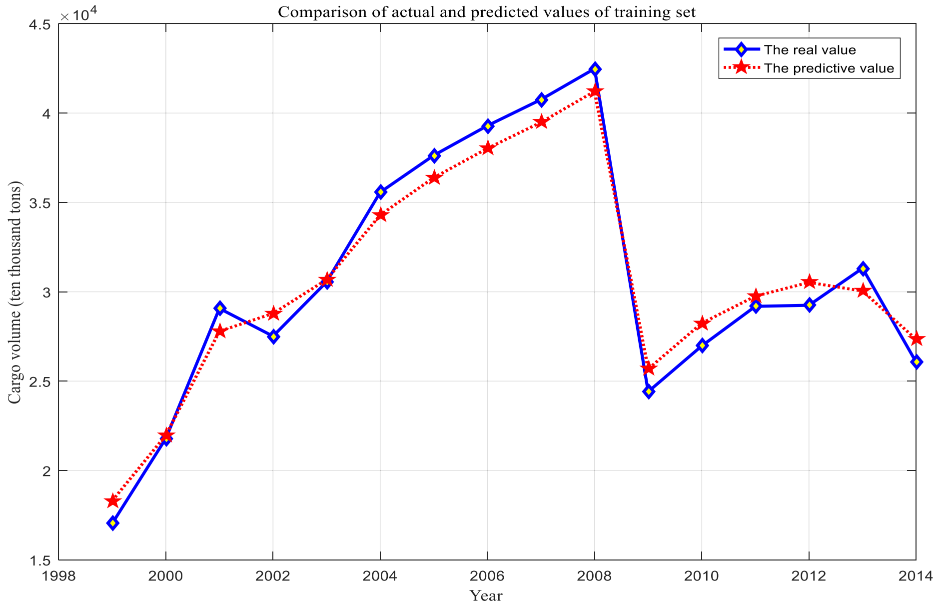

FIGURE 7. Fitting effect of freight volume training results of Qingdao City from 1999 to 2014.

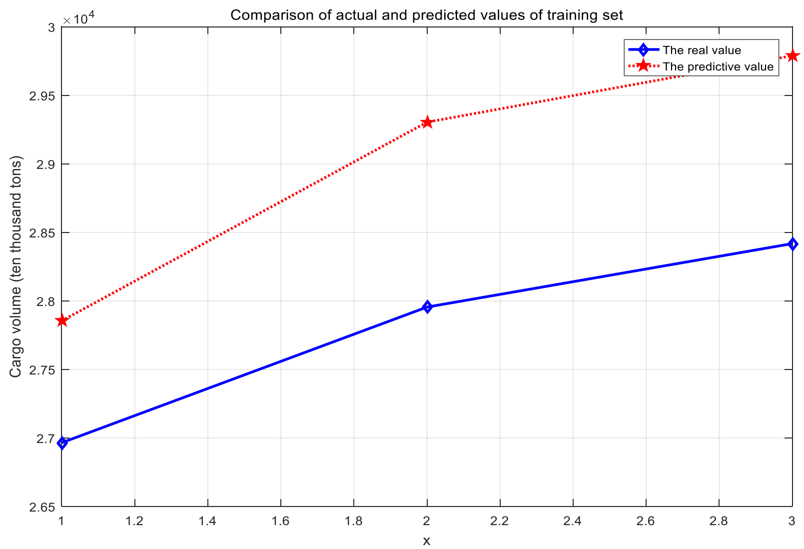

FIGURE 8. Fitting the predicted and actual value of the test sample freight volume.

In order to observe the prediction results more intuitively, the relative error comparison histogram of the ACO-SVM prediction model and the RBF neural network prediction model are plotted according to the above table, as shown in Figure 9.

The average relative error of the ACO-SVM prediction model is 0.038. The average relative error of the RBF neural network prediction model is 0.045. Triple exponential smoothing-Logistic the average relative error of the growth curve-BP neural network prediction model is 0.054. By comparing with other relative prediction methods, the results show that the ACO-SVM prediction model has better prediction effect, The operation time of prediction accuracy in acceptable range is short. the fitting effect and adaptability are better, and it has better applicability in regional logistics demand forecasting.

It can be seen from Figure 10 that the ACO-SVM prediction model is used to predict that the logistics demand deviation in Qingdao is small, and the prediction result is closer to the actual value. Therefore, the above-mentioned trained ACO-SVM prediction model is applied to the logistics demand forecast of Qingdao City in 2018-2022. The predicted freight volume of Qingdao can be predicted in the next five years as shown in Table 8.

As can be seen from Table 8, the logistics demand in Qingdao will continue to rise in the next five years.

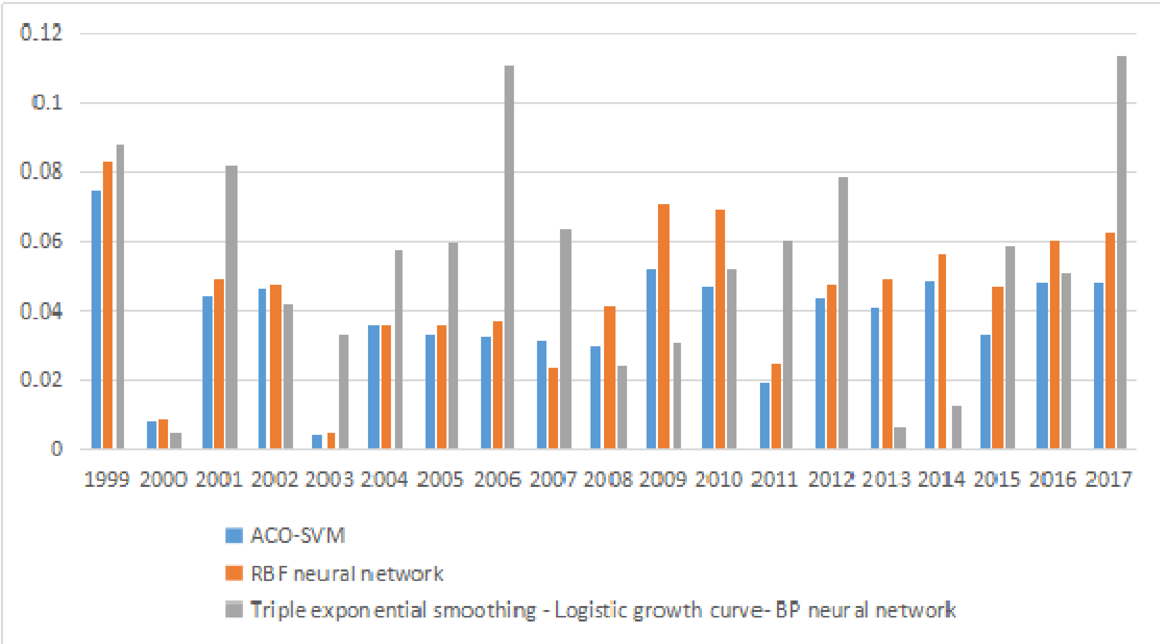

FIGURE 9. Comparison chart of prediction results error.

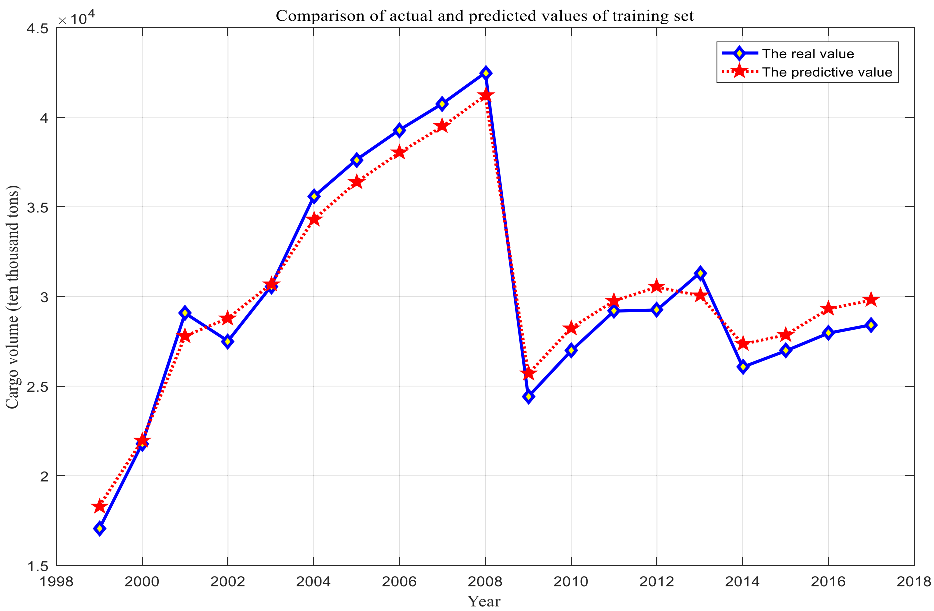

FIGURE 10. Comparison of actual and predicted values of Qingdao's freight volume training set for 1999-2017.

This phenomenon not only provides a major opportunity for the development of logistics in Qingdao, but also puts forward higher requirements for the operation of Qingdao Logistics. Therefore, Qingdao City should encourage the standardization development of the logistics industry, strengthen the standardization management according to the local economic level and the actual situation of logistics development and build logistics infrastructure. In the process of future development, we must not only implement the "Thirteenth Five-Year Plan", but also improve the cultivation degree of the key enterprises, encourage logistics enterprises to actively integrate the "Internet +" innovative logistics service platform model, and extensively attract investment and positively strengthen strategic cooperation with HNA Group to meet the growing logistics needs of Qingdao.

V. SUMMARY

Accurate forecasting of urban logistics demand is an important basis for logistics planning, and it is the premise and basis for comprehensive planning of regional logistics

TABLE 7. Comparison of forecast results.

Year	Real value	ACO-SVM		RBF neural network		Triple exponential smoothing - Logistic growth curve- BP neural network	
		Predicted value	Relative error	Predicted value	Relative error	Predicted value	Relative error
1999	17038	18309.48	0.0746	18449.29	0.0829	18535.73	0.0879
2000	21791	21970.71	0.0082	21984.36	0.0089	21683.37	0.0049
2001	29068	27790.20	0.0440	27625.38	0.0496	26689.42	0.0818
2002	27515	28786.81	0.0462	28829.55	0.0478	28673.15	0.0421
2003	30553	30681.79	0.0042	30701.54	0.0049	31564.62	0.0331
2004	35570	34296.85	0.0358	34286.48	0.0361	37622.45	0.0577
2005	37636	36388.39	0.0331	36287.36	0.0359	35379.42	0.0600
2006	39294	38022.51	0.0324	40762.82	0.0374	43642.58	0.1107
2007	40758	39483.18	0.0313	41725.02	0.0237	43360.37	0.0638
2008	42484	41214.99	0.0299	44246.97	0.0415	43526.56	0.0245
2009	24408	25679.48	0.0521	22682.72	0.0707	23652.64	0.0309
2010	26971	28242.84	0.0472	28834.49	0.0691	25568.82	0.0520
2011	29184	29753.40	0.0195	29920.77	0.0252	27420.62	0.0604
2012	29246	30524.39	0.0437	30638.29	0.0476	31547.84	0.0787
2013	31318	30041.54	0.0408	32864.26	0.0494	31527.75	0.0067
2014	26061	27338.72	0.0490	27527.67	0.0563	26385.43	0.0124
2015	26965	27854.34	0.0330	28227.31	0.0468	28536.31	0.0583
2016	27955	29304.08	0.0483	29634.28	0.0601	29378.28	0.0509
2017	28418	29789.16	0.0482	30202.86	0.0628	31648.64	0.1137

TABLE 8. Forecast value of Qingdao freight volume in 2018-2022.

Year	2018	2019	2020	2021	2022
Cargo Volume (10000tons)	29885.46	29960.25	29995.89	30068.07	30074.43

infrastructure construction scale, network space layout, logistics enterprise development direction and functional positioning. Taking as an example the actual data of economic factors affecting the logistics demand of Qingdao, for the complex logistics system of Qingdao, the ACO-SVM prediction model of Qingdao complex logistics system was first established for training and testing. The experimental results show the prediction result is closer to reality. In order to illustrate the feasibility of the selection method, the paper compares the ACO-SVM prediction model with the prediction results of the RBF neural network method. The experimental results show that the ACO-SVM prediction model has better prediction accuracy and the relative error is 0.0374. Therefore, the training is adopted. The well-prepared ACO-SVM forecasting model predicts the volume of freight from Qingdao to 2018-2022. According to the forecast results, further planning for Qingdao logistics development, Qingdao's logistics demand will continue to grow in the next five years, and it is necessary to further promote the healthy development of Qingdao logistics industry from the aspects of policy support, logistics service and logistics "Internet +" model.

REFERENCES

[1] A. Samvedi and V. Jain, "A grey approach for forecasting in a supply chain during intermittentdisruptions," *Eng. Appl. Artif. Intell.*, vol. 26, no. 3, pp. 1044–1051, Mar. 2013, doi: 10.1016/j.engappai.2012.12.002.

[2] S. Jaipuria and S. Mahapatra, "An improved demand forecasting method to reduce bullwhip effect in supply chains," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2395–2408, Apr. 2014, doi: 10.1016/j.eswa.2013.09.038.

[3] A. Zendehboudi, M. Baseer, and R. Saidur, "Application of support vector machine models for forecasting solar and wind energy resources: A review," *J. Cleaner Prod.*, vol. 199, no. 10, pp. 272–285, Oct. 2018, doi: 10.1016/j.jclepro.2018.07.164.

[4] F. Ge, Y. Ju, Z. Qi, and Y. Lin, "Parameter estimation of a Gaussian mixture model for wind power forecast error by Riemann L-BFGS optimization," *IEEE Access*, vol. 6, pp. 38892–38899, 2018, doi: 10.1109/access.2018.2852501.

[5] T. LIU, "Logistics demand forecasting method combination of Shandong province and its application," M.S. thesis, Dept. Manage. Sci. Eng., Univ. Econ. Bus. Univ., Beijing, China, 2016.

[6] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short–term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, pp. 42946–42955, Mar. 2019, doi: 10.1109/access.2019.2907739.

[7] J. Q. Sun, "Research on logistics demand forecasting of Beijng," M.S. thesis, Dept. School Transp., Beijing Jiaotong Univ., Beijing, China, 2016.

[8] B. Li, J. Zhang, Y. He, and Y. Wang, "Short–term load–forecasting method based on wavelet decomposition with second–order gray neural network model combined with ADF test," *IEEE Access*, vol. 5, pp. 16324–16331, Aug. 2017, doi: 10.1109/access.2017.2738029.

[9] Q. M. Wang, A. W. Fan, and H. S. Shi, "Network traffic prediction based on improved support vector machine," *Int. J. Syst. Assurance Eng. Manage.*, vol. 8, no. 11, pp. 1976–1980, 2017, doi: 10.1007/s13198-016-0412-8.

[10] T. Sujjaviriyasup, "A new class of MODWT-SVM-DE hybrid model emphasizing on simplification structure in data pre-processing: A case study of annual electricity consumptions," *Appl. Soft Comput.*, vol. 54, pp. 150–163, May 2017, doi: 10.1016/j.asoc.2017.01.022.

[11] D. Jiang, Z. Xu, and Z. Lv, "A multicast delivery approach with minimum energy consumption for wireless multi-hop networks," *Telecommun. Syst.*, vol. 62, no. 4, pp. 771–782, Aug. 2016, doi: 10.1007/s11235-015-0111-9.

[12] G. Jiang, *Kernel Function Theory and Signal Processing*. Beijing, China: Science Press, 2013, pp. 97–98.

[13] A. Jayaprakash and C. Keziselvavijila, "Feature selection using ant colony optimization (ACO) and road sign detection and recognition (RSDR) system," *Cogn. Syst. Res.*, vol. 58, pp. 123–133, Dec. 2019, doi: 10.1016/j.cogsys.2019.04.002.

[14] Jayadeva, S. Chandra, S. S. Batra, and S. Sabharwal, "Learning a hyperplane regressor through a tight bound on the VC dimension," *Neurocomputing*, vol. 171, no. 1, pp. 1610–1616, Jan. 2016, doi: 10.1016/j.neucom.2015.06.065.

[15] J. P. Yu, Z. J. Chen, and L. J. Wu, "Forecasting slope displacement based on support vector machine optimized by ant colony algorithm," *J. Yangtze River Sci. Res. Inst.*, vol. 4, no. 32, pp. 22–27, 2015.

[16] A. Rout and P. Dash, "Forecasting foreign exchange rates using hybrid functional link RBF neural network and Levenberg–Marquardt learning algorithm," *Intell. Decis. Technol.*, vol. 10, no. 3, pp. 299–313, Jun. 2016, doi: 10.3233/idt-160257.

[17] S. Liu and L. Tian, "The application of Grey–Markov combined model for port cargo throughput forecasting," in *Proc. 9th Int. Conf. Fuzzy Syst. Knowl. Discovery*, Chongqing, China, May 2012, pp. 973–976.

[18] P. Yan, L. Zhang, Z. Feng, and J. Zhang, "Research on logistics demand forecast of port based on combined model," *J. Phys., Conf. Ser.*, vol. 1168, Feb. 2019, Art. no. 032116, doi: 10.1088/1742-6596/1168/3/032116.

[19] Y. Huang, "Research on container throughput forecast of Qingdao port based on combined forecasting model," in *Proc. 4th Int. Conf. Ind. Bus. Eng. (ICIBE)*, New York, NY, USA, 2018, pp. 99–103.

NAN YU was born in 1995. She received the bachelor's degree in management from the Heze College and the master's degree from the Shandong University of Science and Technology, where she is currently a Graduate Student in logistics engineering. Her areas of interest are logistics system planning and design, port logistics, and supply chain management.

WEI XU was born in 1979. He received the bachelor's degree in management and the D.Eng. degree from the Shandong University of Science and Technology, in 2001 and 2015, respectively. He is currently with the Shandong University of Science and Technology, where he is currently an Associate Professor with the Department of Logistics Engineering, School of Transportation. His research interests include logistics system planning and design, port logistics, and supply chain management.

KAI-LI YU was born in 1993. She received the bachelor's degree in management from Qingdao Agricultural University, in 2016, and the master's degree from the Shandong University of Science and Technology, in 2019. She is currently with the Technology Vocational College of Dezhou, where she is currently a Teacher with the Department of Logistics Engineering. Her research interests include logistics system planning and design, port logistics, and supply chain management.