Indian Journal of Data Analysis and Information Systems
VOLUME 11 • NUMBER 2 • DECEMBER 2019

Research on Regional Logistics Demand Forecast Based on Support Vector Machine

Yu Kaili, Xu Wei
Department of Logistics Engineering, College of Transportation, Shandong University of Science and Technology, Qingdao, China

ABSTRACT

This paper introduces the support vector machine algorithm based on ant colony algorithm optimization to predict the logistics demand of Qingdao. Since the penalty coefficient c of the support vector machine and the parameter g of the kernel function greatly affect the accuracy of the prediction model, the ant colony is utilized. The algorithm optimizes the parameter effect and obtains the optimized support vector machine prediction model,using the support vector machine to predict the superior performance of small samples and nonlinear data, the demand for logistics in Qingdao is predicted. The experimental results show that the prediction accuracy of the ACO-SVM prediction model established in this paper is higher than that of the traditional SVM prediction model, and the error is smaller, which provides data guarantee for Qingdao logistics demand forecasting.

Key words: Support vector machine; ant colony algorithm; regional logistics demand; prediction model

INTRODUCTION

In the past 30 years, with the logistics theory and logistics technology entering China, the logistics industry has been concerned by local governments and core enterprises. The logistics development policies across the country have sprung up. Regional logistics demand forecasting is an important basis for regional logistics planning. It is an important prerequisite for comprehensive planning of regional logistics infrastructure construction scale, network space layout and logistics enterprise development direction. Scientific and rational regional logistics demand forecasting can ensure regional logistics services. The relative balance between supply and demand effectively improves the efficiency and efficiency of regional logistics.

Carbonneau R, Laframboise K, Vahidov R. (2008)[1]in the article, the application of machine learning algorithm in the supply chain demand forecasting scenario is introduced. The machine learning model is constructed to predict the supply demand.

In recent years, machine learning algorithms have been widely used in logistics demand forecasting. Support vector machine (SVM) is a new machine learning method proposed by Vanpik[2-4]. It is based on statistical theory and can effectively solve small samples, nonlinearities and high dimensional issues.

Jontahon T. Fite, John R.(2002)[5]proposed to select a set of economic indicators related to freight volume, and use freight volume to measure the scale of logistics demand more effectively.

Zhongsheng Hua, Bin Zhangi.(2006)[6]study the traditional single model in the text, which can only reflect part of the change law, and can not fully reflect the overall change law of demand, resulting in poor prediction effect. Therefore, a hybrid support vector machine prediction model is proposed to improve the prediction accuracy.

J.Kim.S.Won,(2002)[7]use the combination of fuzzy reasoning theory and support vector machine. According to the commonality of fuzzy reasoning and support vector machine, the support vector machine and fuzzy reasoning theory are combined, and the fuzzy reasoning model based on support vector machine is proposed and constructed.

C. C Chung.(2001)[8]predicted the small and medium-sized samples for the regional logistics demand, and minimized the dimension of the original data. The LIBSVM support vector machine regression model was used for comparison, and the prediction accuracy was significantly improved.

Cao Lijuan Tay, Francis E. H.(2007)[9]based on the characteristics of regional logistics demand forecasting, combined with the method of financial time series and support vector machine, he not only considers the fitting error of the support vector machine regression function, but also applies new data to the development of time, so that the model can be more Precise predictions.

In order to improve the accuracy of time series prediction, Thao-Tsen Chen, Shie-Jue Lee.(2014)[10] proposes a LSSVM prediction model. The combination method not only improves the prediction accuracy of a single LSSVM, but also accelerates the convergence speed of model prediction.

Dorigo M, Blum C.(2005)[11]pointed out that the ant colony algorithm is a swarm intelligence bionic heuristic algorithm, which has been widely used in the optimization problems in different fields. The basic principle of ant colony algorithm and its path optimization and production scheduling are introduced. , image processing and other aspects of the application.

In order to avoid the blindness of parameter selection in support vector machine prediction process, ant colony algorithm (ACO) is used to optimize the parameters of support vector machine under the premise of summing up a large number of scholars' research on support vector machine. The experimental results show that, The prediction model based on ACO-SVM has achieved good results.

BASIC PRINCIPLE OF SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) originated from the early work of Vapnik and Chervonenkis[12]on statistical learning in 1971. The "A training algorithm for optimal margin classifiers" published by Boser, Guyon and Vapnik[13]in 1992 involved convex analysis algorithms, kernel functions, and nerves. High-level fields such as the Internet. Generally speaking, the support vector machine is a two-class classification model. The basic model is defined as the linear classifier with the largest interval in the feature space. That is, the learning strategy of the support vector machine is to maximize the interval and finally transform into a convex. Solving the quadratic programming problem.

One of the core ideas of SVMs is to control the generalization ability, which is to maximize the classification interval. The statistical learning theory points out that the optimal hyperplane has the best generalization performance[14], so that the problem of finding the optimal hyperplane is transformed into the optimization problem. In the case of linear separability, the optimal hyperplane can be represented by the classification function $f(x) = w^T x + b$. You can use $y(w^T x + b)$ to represent the correctness and certainty of the classification. This is the function. interval:

$$\widehat{\gamma_i} = y_i(w^T x_i + b) \qquad (1)$$

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

$$\min_{w,b} \frac{1}{2} \|w\|^2 \qquad (2)$$

s. t.

$$y_i(w^T x_i + b) - 1 \geq 0, i = 1,2,\cdots,n$$

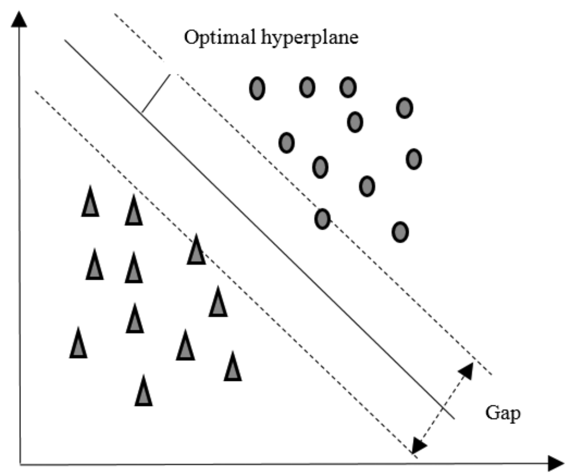

Figure 1. Optimal hyperplane diagram in the case of linear separability

Considering the real data in real life, there are some singular points that do not satisfy the constraints derived above[15]. To solve this problem, a slack variable is introduced for each sample point, and the corresponding target variable becomes :

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{R} \varepsilon_i \qquad (3)$$

s. t.

$$y_i(w^T x_i + b) \geq 1 - \varepsilon_i$$
$$\varepsilon_i \geq 0$$
$$i = 1,2,\cdots,n$$

Where C is the penalty parameter[16], which is used to control the weight between the target function to find the largest hyperplane of the margin and to ensure the minimum deviation of the data points. For the convex quadratic optimization problem, the objective function and constraints are integrated into the Lagrangian function by introducing Lagrangian multipliers, which is convenient for solving the most value problem. Then, the Lagrange multiplier α is introduced for each inequality constraint, and the Lagrangian function is obtained as follows:

$$\min_{\alpha} \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n} \alpha_i \tag{4}$$

s. t.

$$C \geq \alpha_i \geq 0, i = 1,2,\cdots,n$$
$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

The classic SVM algorithm only supports two classifications. For multi-classification problems, the model needs to be changed. Class-type data is not supported. Class-type data needs to be converted into discrete data in the pre-processing stage. In the face of linear inseparable data samples, the support vector machine first completes the calculation in low-dimensional space, and then maps the data into high-dimensional space through the kernel function, thus solving the optimal separation hypersurface in the high-dimensional feature space[17], as shown in Figure 2 shown.

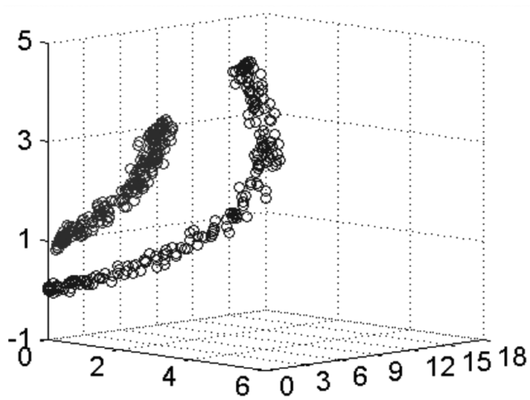

Figure 2. RBF kernel function mapping diagram

The kernel function can simplify the inner product operation of the mapping space[18]. The commonly used kernel function of the support vector machine is the RBF kernel, and the RBF kernel can map the low-dimensional space to the infinite dimension[19]. After selecting the RBF[20] kernel function, the objective function and constraints of the optimization problem become:

$$\min_{\alpha} \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{n} \alpha_i \tag{5}$$

s. t.

$$C \geq \alpha_i \geq 0, i = 1,2,\cdots,n$$
$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

$K(x_i, x_j) = exp\left(-\frac{d(x_i,x_j)^2}{2\sigma^2}\right)$ÿ $gamma = \frac{1}{2\sigma^2}$, the SVM model has two very important parameters C and $gamma$, where is the penalty factor, which is the tolerance for the error. The higher , the more the error can not be tolerated, the easy over-fitting, the smaller , the easy under-fitting, too large or too small, the generalization ability is worse, is a parameter that comes with the function after selecting the RBF function as the kernel. Implicitly determines the distribution of data after mapping to a new feature space[21]. The larger the , the smaller the support vector, the smaller the value, and the more support vectors, the number of support vectors affects the speed of training and prediction.

If the is set too large, will be small, and the small Gaussian distribution of will be tall and thin, which will only affect the vicinity of the support vector samples[22]. The classification effect for unknown samples is very poor, and the training accuracy rate can be very high. High, (if is infinitely small, then theoretically, the Gaussian kernel SVM can fit any nonlinear data, but it is easy to overfit) and the test accuracy is not high, which is usually said to be training; Too small will result in too much smoothing effect, can not get a particularly high accuracy on the training set, and will affect the accuracy of the test set.

CONSTRUCTION OF MATERIAL FLOW FORECASTING MODEL BASED ON ACO-SVM

3.1 Ant Colony Optimization Algorithm for SVM Parameters

Ant colony optimization (ACO), also known as ant algorithm, is a probabilistic algorithm used to find optimized paths in the graph. It was proposed by Marco Dorigo[23] in his doctoral thesis in 1992, inspired by the behavior of ants finding paths in the search for food. During the movement, ants will leave something called pheromone, and the pheromone will be less and less scattered with the distance of movement. Therefore, the concentration of pheromone is often at home or around food. Strong, and the ants themselves will choose the direction according to the pheromone. Of course, the stronger the pheromone, the greater the probability of being selected, and the pheromone itself has a certain volatilization effect. Ant colony algorithm[24] has the effectiveness and application value of a new simulated evolution optimization method, and it has been used in all aspects of our lives.

The optimization of the support vector machine parameters is to solve the problem of continuous domain. In this paper, the K-fold cross-validation error is selected as the target value, and the kernel parameter g and the penalty coefficient C are optimized by the ACO algorithm[25]. The key of the ACO algorithm lies in the movement rules and pheromone update. The ant colony performs the mobile search through the positive feedback of the pheromone volatilization accumulation, searches the optimal combination of the parameters of the support vector machine, and obtains the SVM prediction model with the best modeling accuracy.

Influencing factors	Indicator name	Indicator unit	Symbol
Scale of economic development	GDP	Billion	X_1
Industrial structure	Primary industry added value	Billion	X_2
Commercial trade	The total retail sales of social consumer goods	Billion	X_3
Income level	Per capita disposable income of urban residents	RMB	X_4
Capital investment	Total investment in fixed assets of the whole society	Billion	X_5
Consumer market	Total population people	Ten thousand	X_6

The process of algorithm design is as follows:

Step 1: Initialize the relevant parameters, set the ant number m, the pheromone volatilization factor ρ, and define a one-dimensional array $path_k$ with n elements for each ant k. [26]

The ordinates of the n nodes through which the kth ant passes are sequentially stored in the , which can be used to represent the crawling path of the th ant, where is the total effective bit of the optimized parameter. Let the time counter , the number of cycles , and set the maximum number of cycles .

Step 2: Randomly place the ants at different starting points and calculate the next visiting city for each ant until there are ants accessing all the cities.

Step 3: Calculate the path length of each ant, record the optimal solution of the current iteration number, and update the pheromone concentration on the path.

Step 4: Calculate the and corresponding to the path according to the path that the ant has traveled, that is, the .

Step 5: Train the SVM based on the calculated to calculate the k”fold cross validation error.

Step 6: Record the optimal path of the current cycle with the k”fold cross-validation error as the fitness value. Let , , according to the amount of information on each node of the update formula, and clear all elements in . [27]

Step 7: Determine whether the maximum number of iterations is reached. If not, return to step 2; Yes, end the program.

Step 8: Output the results and output the optimal combination in the optimization process as needed.

The flow of ACO algorithm to optimize SVM parameters is shown in Figure 3.

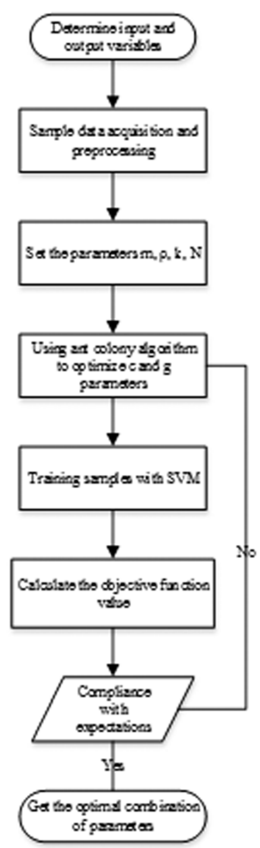

Figure 3. SVM parameter optimization flow chart based on ant colony algorithm

3.2 Logistics demand forecast impact indicator

The logistics industry, which is known as the "third profit source", has become more and more important in the process of global economic integration and has become one of the pillar industries of the major developed countries. The logistics industry has become an important part of the national economy and a strategic driving force for China's economic growth. The rapid development of the regional economy is not only an important indicator of the prosperity of China's market economy, but also puts forward higher requirements for the development of the logistics industry.

The regional economy is an important gathering area for resource elements in the process of market economy development. The development of the logistics industry can promote the exchange frequency of regional funds, information and the entire economic system, and improve the efficiency of economic operations as a whole. Therefore, various regions have successively issued a series of planning and policy measures to promote the development of the logistics industry in the region, thereby enhancing core competitiveness.

The logistics demand in Qingdao is affected by many factors. Generally speaking, it is divided into economic factors and non-economic factors. The economic factors are the main influencing factors of regional logistics demand forecasting[28]. Economic factors include:

(1) The level of regional economic development. The higher the level of regional economic development, the greater the demand for production materials, semi-finished products and finished products. The greater the circulation of materials, the greater the logistics demand. Fundamentally speaking, the development of regional economy is the internal driving force of regional logistics demand, which plays a decisive role in the amount of logistics demand.

(2) Industrial structure. As the social division of labor becomes more and more detailed, and more and more production departments are produced, China has divided its production parts into three industries, namely, agriculture, industry and service industries, namely, the first, second and third industries, and different industries. The level and scale of demand for logistics are different.

(3) Regional and foreign trade. Trade in the region promotes cooperation between enterprises and departments, and has greatly promoted large-scale chain companies and major supermarkets, effectively promoting the prosperity of logistics and distribution business.

(4) Income level and consumption level. The improvement of the social and economic level has led to an increase in per capita disposable income and household consumption levels. Different income classes have different ideas about consumption, different willingness to purchase goods, and then different consumption behaviors.

(5) Capital investment. Capital includes physical capital and human capital. The capital investment of the state and region affects the logistics demand of the region to a large extent. In particular, the material investment in logistics parks and transportation facilities can improve the regional logistics service capacity.

(6) Consumer market. People's material exchange or economic activities form a consumer market, so the size of the consumer market is limited by population density.

Based on the systematic review of domestic and foreign research results and the theory of logistics demand forecasting, this paper constructs the Qingdao logistics demand forecasting index system based on the correlation between regional economy and logistics industry.This paper uses the freight volume to measure the scale of regional logistics demand. In the actual forecasting research, according to the availability of statistical data and the statistical data of each region, the following influencing factor index system is constructed:

Table 1:Logistics demand forecasting impact indicator system

EMPIRICAL STUDY ON FORECAST OF LOGISTICS DEMAND IN QINGDAO CITY

4.1 Indicator data acquisition and preprocessing

This paper summarizes and summarizes the current situation and characteristics of Qingdao logistics demand, and selects cargo freight volume indicators to measure the scale of logistics demand in Qingdao. The logistics demand data of Qingdao from 1999 to 2017 was selected as the raw data table. The logistics demand data from 1999 to 2014 was used as the training sample, and the logistics demand data from 2015 to 2017 was used as the forecast sample, as shown in Table 2 shown.

Table 2 :Qingdao logistics demand forecast indicators and data

	Y	X_1	X_2	X_3	X_4	X_5	X_6
1999	17038	985.20	138.19	105.30	7282	220.59	702.97
2000	21791	1151.61	140.85	225.60	8016	242.68	706.65
2001	29068	1316.08	144.35	312.85	8731	293.47	710.49
2002	27515	1518.17	147.21	482.64	8721	368.36	715.65
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

Because the index is different in magnitude and unit, this paper reduces the sample data to [-1,1], which eliminates the influence of dimension on the sample data. The input data of the sample set is normalized according to the function mapminmax.

Table 3 :Normalized data

	Y	X_1	X_2	X_3	X_4	X_5	X_6
1999	-1	-1	-1	-1	-1	-1	-1
2000	-0.62642	-0.96689	-0.9781	-0.94576	-0.9632	-0.9942	-0.96745
2001	-0.05447	-0.93417	-0.9493	-0.90642	-0.9274	-0.9807	-0.93347
2002	-0.17653	-0.89396	-0.9257	-0.82986	-0.9279	-0.9609	-0.88783
2003	0.06225	-0.84178	-0.9116	-0.76052	-0.8600	-0.9135	-0.84333
2004	0.45657	-0.76550	-0.7916	-0.70638	-0.8091	-0.7978	-0.75097
2005	0.61896	-0.65965	-0.6693	-0.61223	-0.7174	-0.6870	0.03132
2006	0.74927	-0.55803	-0.6230	-0.59357	-0.5966	-0.6652	0.11863
2007	0.86434	-0.44264	-0.4612	-0.50677	-0.4699	-0.6255	0.20046
2008	1	-0.31338	-0.2980	-0.38702	-0.3391	-0.5240	0.26185
2009	-0.42073	-0.23025	-0.2416	-0.26735	-0.2437	-0.4076	0.30096
2010	-0.21929	-0.06865	0.1434	-0.16324	-0.1118	-0.2584	0.49098
2011	-0.04535	0.12025	0.3855	-0.00936	0.0671	-0.1314	0.56175
2012	-0.04048	0.25683	0.5341	0.14088	0.2465	0.0410	0.62668
2013	0.12238	0.39700	0.6666	0.29923	0.4010	0.2724	0.71125
2014	-0.29081	0.53339	0.7417	0.46827	0.5072	0.4677	0.78388
2015	-0.21976	0.65436	0.8600	0.62698	0.6588	0.6767	0.82882
2016	-0.14195	0.79587	0.9180	0.80337	0.8206	0.9147	0.92348
2017	-0.10556	1	1	1	1	1	1

4.2 Predictive model solving

The non-dimensionalized data is substituted into the ACO-SVM object flow prediction model for training and testing models. In this paper, the RBF kernel function is selected, and the model parameter C, g is set to [0.1, 1000], cross-validation $t = 6$, modeling the freight demand series of Qingdao from 1999 to 2014. Under the operation of MATLAB software, the final result of the model training is shown in Figure 4.

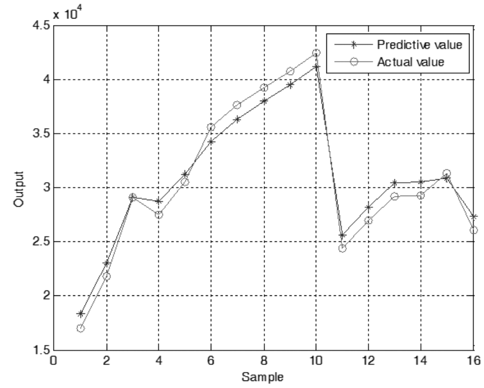

Figure 4. Comparison of actual and predicted values of training set

After obtaining the training model, the model needs to be tested to verify its accuracy. In this paper, the corresponding indicators of freight volume from 2015 to 2017 are test data, and the input vector is a set of these indicators. The predicted and actual values of the test year obtained by the model prediction are shown in Figure 5. It can be seen from the figure that the difference between the predicted value and the actual value is small, and the prediction effect is ideal.

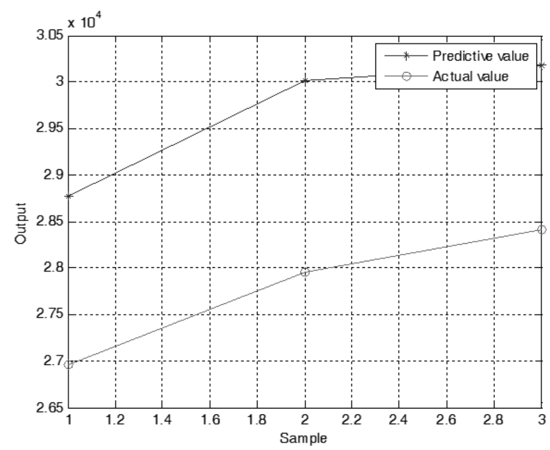

Figure 5. Test set actual value and predicted value comparison chart

From the running process, it can be concluded that the parameter $C = 271.8591$, $g = 3.5748$, the goodness of fit $R^2 = 0.9747$, the closer the value of R^2 is to 1, the better the fitting degree of the regression line to the observed value, the average percentage error MAPE=0.0343.

4.3 Predictive model comparison

As shown in Table 4, the prediction model of ACO-SVM is compared with the prediction model of traditional SVM, which is the fitted value from 1999 to 2014 and the predicted value from 2015 to 2017. Among them, the root mean square error of RMSE=1084.86215 of ACO-SVM is smaller than the root mean square error obtained by SVM, which indicates that the method of ant colony algorithm optimization of support vector machine parameters is feasible.

Table 4 :Comparison of ACO-SVM and SVM prediction models

	Actual value	ACO-SVM prediction model		SVM prediction model	
		Predictive value	Relative error	Predictive value	Relative error
1999	17038	18309.47876	0.0746	18449.28642	0.0829
2000	21791	21970.71477	0.0082	21984.35681	0.0089
2001	29068	27790.19778	0.0440	27625.37560	0.0496
2002	27515	28786.81243	0.0462	28829.54723	0.0478
2003	30553	30681.79045	0.0042	30701.53782	0.0049
2004	35570	34296.84539	0.0358	34286.48207	0.0361
2005	37636	36388.39069	0.0331	36287.35963	0.0359
2006	39294	38022.51413	0.0324	40762.82145	0.0374
2007	40758	39483.18196	0.0313	41725.01823	0.0237
2008	42484	41214.99398	0.0299	44246.96510	0.0415
2009	24408	25679.47893	0.0521	22682.72305	0.0707
2010	26971	28242.84163	0.0472	28834.48923	0.0691
2011	29184	29753.40007	0.0195	29920.76527	0.0252
2012	29246	30524.39197	0.0437	30638.28743	0.0476
2013	31318	30041.54089	0.0408	32864.25686	0.0494
2014	26061	27338.72138	0.0490	27527.66712	0.0563
2015	26965	27854.34177	0.0330	28227.30561	0.0468
2016	27955	29304.08331	0.0483	29634.27682	0.0601
2017	28418	29789.15670	0.0482	30202.86432	0.0628

It can be seen that the ACO-SVM prediction model shows better prediction results. The prediction accuracy of the prediction model is within the acceptable range, and the prediction model runs for a short time. This also proves that the support quantity machine can convert the optimization problem into Solve linear equations and solve the advantages of fast speed.

4.4 Forecast Analysis of Freight Volume in Qingdao City

Using the ACO-SVM prediction model established in this paper, the logistics demand data of Qingdao from 1999 to 2017 will be used as a training sample to predict the freight volume of Qingdao in the next three years. The actual value and forecast of the freight volume of Qingdao will be trained. The comparison of values is shown in Figure 6.

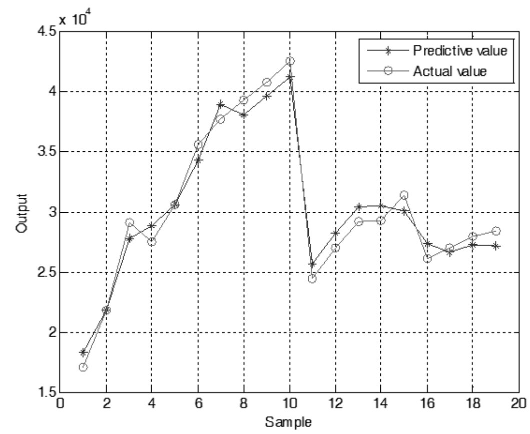

Figure 6. Comparison of actual and predicted values of training set

According to the relevant forecasting index data of Qingdao City, using the ACO-SVM forecasting model established in this paper, it is possible to predict the freight volume of Qingdao in the next three years, which are 30,630,565,620 tons, 306,925,573 tons, and 3,065,666,600 tons.

According to the above research results, the predicted and actual values of the support vector machine model constructed in this paper are close. It shows that the predicted results of the constructed logistics demand forecasting model can accurately reflect the trend of market demand and have good practical value[29]. Therefore, the model is a more effective and applicable forecasting method for logistics demand.

CONCLUSION

With the deepening of the national "Belt and Road" strategy, according to Qingdao's "13th Five-Year Plan" logistics development plan, Qingdao logistics industry is required to actively integrate into the "One Belt, One Road" strategy and play a leading role. Regional logistics demand forecasting is an important basis for regional logistics planning, and it is the premise and basis for comprehensive planning of regional logistics infrastructure construction scale, network space layout, logistics enterprise development direction, and functional positioning. If the demand is not adequately estimated, the logistics enterprises will lose a lot of profit opportunities; but if the demand estimation is too exaggerated, the regional logistics planning will cause the abuse and shortage of the logistics enterprise funds due to excessive investment, so the logistics demand science Forecasting has very important theoretical and practical implications.Based on the analysis of the impact forecast of logistics demand forecast in Qingdao, this paper proposes an improved model of freight volume forecast based on ant colony algorithm optimization support vector machine parameters, and applies the data from Qingdao to 1999 to verify the example.

Based on the comprehensive consideration of Qingdao logistics demand, the index system of regional logistics demand forecasting is constructed. Due to the large quantity, high value, time difference and reliability of Qingdao logistics and freight transportation, Qingdao logistics and freight transportation for society and economy Has a very big impact, the selection of freight volume to measure the scale of Qingdao logistics demand has practical significance for the scientific forecast of Qingdao logistics demand.

According to the statistical data of Qingdao from 1999 to 2017 as an empirical research object, predicting the trend of freight volume in Qingdao in the next three years is of great significance to the development of regional logistics and analysis of the logistics market situation. The support vector machine model parameters $\{C, g\}$ have a great influence on the prediction accuracy and the generalization ability. In order to avoid the blindness of artificial selection parameters, this paper proposes to use the improved ant colony algorithm to optimize SVM and search for the optimal parameters. The research shows that the ACO-SVM model is feasible for regional logistics demand prediction, and the support vector machine analysis method is implemented. The application in logistics demand forecasting has certain practical application value.

In the freight volume forecast of Qingdao from 1999 to 2017, the ACO-SVM model and the traditional SVM prediction are more accurate and more accurate, indicating the correctness of the ACO-SVM model.

Although this paper has achieved certain research results, due to the problem of time and the level of knowledge in this paper, the research in this paper is still in a certain deficiency, which is also the direction to be studied in this paper:

The regional logistics demand forecast includes the source of goods, transportation speed, transportation direction and composition of goods. However, this paper only predicts the total amount of freight, and does not involve the planning and distribution of logistics network. This is the direction of future research.

The predictive index system established in this paper is not authoritative. It is only established for the current economic and logistics situation in Qingdao. In the future research, it is necessary to consult more authoritative data and expand the index system of this paper to make it more rational and complete.

Due to the problem of time, this paper does not compare the prediction results with other prediction methods, but only compares the parameters before and after optimization. In the future research, try to compare more prediction algorithms to make the model of this paper more popular and effective.

References

[1] Carbonneau R, Laframboise K, Vahidov R. (2008). Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research, 184(3):1140-1154.

[2] VAPNIK V.N. (1995),The Nature of Statistical Learning Theory, New York: Springer-Verlag.

[3] Vapnik V, Golowich S, Smola A.(1996).Support vector method for function approximation, regression estimation, and signal processing.Advance in Neural Information Processing Systems Cambridge M IT Press,1996:281-287.

[4] VAPNIK V.N.(2012).Statistical Learning Theory, Encyclopedia of the Sciences of Learning,41(4):3185-3185.

[5] Jontahon T. Fite, John R. English, John N. Roberts, G. Dom Taylor, John S. Usher. Forecasting freight demand using economic indices [J]. International journal of physical distribution. 2002 (32): 3-4.

[6] Zhongsheng Hua, Bin Zhangi.(2006).A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. Applied Mathematics and Computation, (181):1035–1048.

[7] J.Kim.S.Won,(2002).New fuzzy inference system using a support vector machine, Proceeding of the IEEE Conference on Decision and Contro,12:1349-1354.

[8] C. C Chung.(2001). Libsvm: a libraly for support vector machines.

[9] Cao Lijuan Tay, Francis E. H. (2007). Application of support vector machines in financial time series forecasting. Journal of University of Electronic Science Technology of China, 48(1):847–861.

[10] Thao-Tsen Chen, Shie-Jue Lee.(2014).A weighted LS-SVM based learning system for time series forecasting, Information Sciences.

[11] Dorigo M, Blum C.(2005). Ant colony optimization theory: A survey. Theoretical Computer Science, 344(2/3): 243"278.

[12] Vapnik, V.N., Chervonenkis, A.Ya.(1971).On the uniform convergence of relative frequencies of events to their probabilities. Dokl. Akad. Nauk SSSR,181(4),781–783.

[13] Boser B.E, Guyon I.M, Vapnik V.N.(1992).A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. Pittsburgh, Pennsylvania, USA: ACM. pp. 144–152.

[14] FABIEN Lauer, GéRARD Bloch. (2008).Incorporating Prior Know ledge in Support Vector Machines for Classification: A Review, Neurocomputing, 71(7/9): 1578-1594.

[15] Wang, Jianguo, Zhang, Zhijie, Zhang, Wenxing.(2013).Support Vector Machine based on Double-population Particle Swarm Optimization, Journal of Convergence Information Technology, 8(8):898-905.

[16] CHAPPELLE O,VAPNIK V,BOUSQUET O.(2002),Choosing Multiple Parameters for Support Vector Machines, Machine Learning,46(1):131 ÿ160.

[17] ZHENG Chun-hong, JIAO Li-cheng. (2004). Automatic parameters selection for SVM based on GA//Proceedings of the the 5th World Congress on Intelligent Control and Automation. Piscataway. NJ: IEEE Press, 1869"1872.

[18] Burges J C.(1999).Geometry and invariance in kernel based methods. Advances in Kernel Methods-Support Vector Learning .Cambridge:M IT PPres,89-116.

[19] Cristianini N,Shawe-Taylor J.(2000).An introduction to support vector machines and other kernel-based learning methods.

[20] Cr Istian Nin, Shawe-Ta Ylor J.(2000).An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. The Syndicate of the Press of the University of Cambridge.

[21] SCHOLKOPF B,SMOLA J.(2001).Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond.

[22] SONG Rui, XUE Xingjian. (2017).Freight Volume Forecast based on Improved Radial Basis Function Neural Network, Boletin Tecnico/Technical Bulletin.

[23] Dorigo M.(1992).Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. IEEE Transactions on Evolutionary Computation, 1(1): 53–66.

[24] Sumim T,Ongsakul W.(2003).Ant colony search algorithmfor unit commitment. IEEE International Confer-ence on Industrial Technology.

[25] Dorigo M, Stutzle T. (2004). Ant colony optimization. Cambridge: MIT Press.

[26] WANG F XÿZHANG L S.(2009).Combination Gray Forecast Model Based on the Ant Colony Algorithm.Mathematics in Practice&Theory,39(14) :102-106.

[27] AFSHAR M H.(2010).A Parameter Free Continuous Ant Colony Optimization Algorithm for the Optimal Design of Storm Sew er Netw orks: Constrained and Unconstrained Approach, Advances in Engineering Softw are, 41(2): 188-195.

[28] Ling Zhou.(2017),Prediction of a service demand using combined forecasting approach. Journal of Physics: Conference Series.

[29] Theodore B. Trafalis Huseyin Ince.(2008).Short term forecasting with support vector machines and application to stock price prediction. International Journal of General Systems, 37(6):677–687.