MLP neural network-based regional logistics demand prediction

Hongpeng Guo[1] · Cheng Guo[1] · Beichun Xu[1,2] · Yujie Xia[1] · Fanhui Sun[1]

Abstract

The introduction of logistics theory and logistics technology has made the government and enterprises gradually realize that the development of logistics has an important strategic role, which can effectively solve the changing needs of users, optimize resource allocation, improve the investment environment, and enhance the overall strength and overall competitiveness of the regional economy. This paper carries out matrix–vector multiplication operations and weight update operations, designs a perceptron neural network model, and realizes a simulation platform based on MLP neural network. Moreover, on the basis of the standard MLP neural network, this paper proposes to use the deep learning training mechanism to improve the MLP neural network, which provides effective technical support for the improvement of the prediction model. In addition, through the fusion of deep learning and MLP neural network, an MLP neural network with three hidden layers is determined. Finally, this paper builds a model based on the MLP neural network algorithm, selects the RBF kernel function as the kernel function of the model by referring to the relevant literature, and uses PSO to optimize the combination of parameters. It can be seen from the result of the evaluation index that each evaluation index is relatively small. The result shows that the prediction is accurate, and the empirical result shows the feasibility of the model to predict the demand for industrial logistics in Shanxi Province.

Keywords MPL neural network · Regional logistics · Demand prediction · Prediction model

1 Introduction

Urban logistics is the foundation to ensure the normal operation of various functions of the city and is a new engine to promote the healthy and stable development of the urban economy. With the continuous deepening of economic transformation and the optimization and upgrading of industrial structure, the process of professionalization and socialization of logistics has accelerated significantly, which has stimulated the continuous growth of logistics demand in various industries and improved the level of logistics demand. In particular, the development of e-commerce has increased the demand for urban logistics services [1].

In recent years, with the continuous strengthening of regional integration and regional coordinated development, the division of labor and cooperation in industries has gradually matured, resulting in the continuous rise in the level of socialization and specialization of production. Moreover, the modern logistics industry has gradually developed into an important basic service industry in China's national economic system and has gradually become another key breakthrough in profit acquisition after reducing material consumption and increasing labor productivity. As an important part of China's modern service industry, modern logistics industry has an all-round and multi-angle role in promoting economic and social development. The role is mainly reflected in accelerating the development of the national economy to high-quality, optimizing, and improving economic business processes, adjusting the economic structure, expanding domestic demand, and enhancing social welfare expenditures. From a time perspective, the modern logistics industry can speed up circulation and reduce delays, thereby reducing the possibility of inventory backlogs and out of stocks in various industries and helping to increase production and

✉ Yujie Xia
    shendan@jlau.edu.cn

1   College of Biological and Agricultural Engineering, Jilin University, Changchun 130025, China

2   Changchun Normal University, Changchun 130000, China

circulation speeds, and optimize industrial economic operation processes. From a spatial perspective, the modern logistics industry can realize the effective transfer and docking of products between the place of production and the place of consumption to reduce ineffective production, thereby realizing the optimization and upgrading of resource allocation and industrial structure, and ultimately promote the efficient and coordinated development of related industries and the continuous improvement of economic operation quality. At present, although there are many types and forms of modern logistics, from the perspective of the main scope, modern logistics can be roughly divided into enterprise logistics, urban logistics, national and international logistics. Although their service levels and objects are different, they do support each other in function. In particular, enterprise logistics and urban logistics are complementary and dependent on each other. Enterprise logistics contributes to the development of urban logistics, and urban logistics development in turn provides opportunities for enterprise logistics. The urban logistics mentioned here refers to the logistics that exists in the urban area and serves the city and belongs to the category of middle logistics. In some mega-consumer cities, the logistics are mainly import and urban distribution. With the further development of the urban economy, the scale and level of logistics demand are also quietly changing with the adjustment of the city's functional structure, spatial form, and industrial structure under the background that the consumption growth rate of modern urban life is accelerating year by year. Various new types of logistics, such as cold chain logistics, e-commerce logistics, community logistics, and express logistics, are developing rapidly. Moreover, consumers' demand for personalized and specialized terminal logistics services is also increasing, and their requirements for urban logistics in terms of service capabilities and service forms are also increasing. With the rapid growth of logistics demand, some shortcomings of urban logistics system planning are gradually exposed. The traditional basic theory of logistics facility layout and network node planning and construction lack systemicity, coordination and standardization, and are far behind the diversified needs of urban logistics development. Moreover, the construction of the original logistics service system failed to fully consider the city's future logistics demand situation, the construction of logistics infrastructure was repeated, and there was an imbalance between logistics supply and logistics demand. The reason lies in the lack of strong scientific basis and scientific theory as the basis for logistics demand and analysis, and it is impossible to effectively evaluate logistics demand. Therefore, it cannot follow the development trend of logistics demand and reasonably respond to and guide the

demand space of various new types of logistics business [2].

Urban logistics demand prediction is to find relevant factors affecting logistics demand on the basis of analyzing the relationship between logistics supply and demand and to predict the indicators that can reflect logistics demand with the aid of quantitative and qualitative analysis methods and objective data. Logistics demand prediction is an important prerequisite for the implementation of spatial partition guidance of logistics facilities to maximize the efficiency of logistics resource allocation. The prediction results are an important basis for ensuring the relative balance of regional logistics demand and supply. In addition, as the basis of urban logistics system planning, accurate logistics demand prediction has important theoretical and practical significance. (1) Theoretical significance. Quantitative prediction of the overall urban logistics demand in the future can not only effectively reveal the inherent relationship between relevant influencing factors and urban logistics demand, but also further analyze the theoretical feasibility and practical operability of the logistics demand prediction ideas through the establishment of a combined prediction model, which broadens the theoretical horizon of urban logistics demand prediction research. At the same time, it has also enriched the relevant theories of logistics management and urban planning to a certain extent. (2) Practical significance. Logistics demand prediction can help the government have a more comprehensive understanding of the factors that affect the healthy development of the logistics industry and the extent of the impact, so as to accurately grasp the overall situation of changes in urban logistics demand. Moreover, it provides a more accurate quantitative basis for the overall spatial layout planning of urban logistics and the effective allocation of logistics resources and ensures the relative balance between urban logistics service supply and demand [3].

2 Related work

Foreign research on demand prediction began in the 1870s, and there have been many related results, some of which have been used in enterprises and governments. The literature [4] proposed a combination method for the first time when studying time series problems. Through comparison, it was found that the more combination methods, the more accurate the prediction. The literature [5] analyzed and extracted the data in the US aviation report, and then used the gray theory model to predict the US aviation demand and proved the nonlinear characteristics of the US aviation service. After studying the prediction results of three different prediction methods of BP neural network, linear

regression, and sliding regression, the literature [6] found that the prediction accuracy of neural network is much higher than the other two prediction methods. The literature [7] used support vector machine (SVM) and neural network two learning machine prediction methods to predict the distorted demand in the end supply chain and compared the two prediction results with the prediction results of traditional prediction methods. Finally, it found that the prediction accuracy of the learning machine is higher than that of the traditional model. The literature [8] used a combined prediction method when predicting the sales volume of Chilean supermarkets, combining neural network algorithms, and moving average algorithms into a combined prediction system. The empirical study found that the method of this literature greatly improved the accuracy of prediction and played a role in reducing inventory levels. The literature [9] proposed the important role of prediction for operational performance, and based on the prediction, it reduced the total cost by simulating the data on the supply chain.

When conducting regional logistics research, the literature [10] combined DWT and ANN to reduce the increased inventory cost due to the bullwhip effect, and empirically proved the accuracy of the model. The literature [11] emphasized the role of prediction in alleviating traffic problems when studying regional logistics and established a 0-D matrix based on three types of considerations and proved the accuracy of the prediction with empirical evidence.

The literature [12] first established a logistics demand prediction index system, and then used neural network method, gray system, and regression analysis to predict on this basis and formed a combined model on this basis. The literature [13] first analyzed the relationship between regional economy and logistics and proposed a "regional economy-logistics demand" prediction model. The empirical results verify that the model is more accurate. The literature [14] discussed the characteristics and steps of demand prediction in detail and used two different methods to form a combined model. In the empirical research stage, actual data are used for prediction. The results show that the model works well and contributes to the government's formulation of relevant policies and infrastructure planning. When studying the development of regional logistics, the literature [15] first studied the factors that affect logistics demand and selected appropriate factors as the research indicators in the article. Moreover, it predicted the logistics demand level of Hubei Province through three methods: GM (1, 1), linear regression, and combined model. The literature [16] used the time series method to predict the logistics demand and used examples to explain the whole calculation process in detail, and how to use this method correctly in the prediction research. The empirical

calculation results show that the time series method is very effective in the medium- and short-term logistics demand prediction and can achieve the expected accuracy. The literature [17] mainly studied regional logistics from three levels: what is the content of the prediction, how to select the influencing factors, and what are the methods. In empirical research, the results of different methods are weighted. The literature [18] introduced the principal component method to deal with the influencing factors, thus forming several comprehensive factors, thus establishing a new prediction model. Under this premise, it compared the prediction results with the other two commonly used methods and concluded that the model established in this literature is better. The literature [19] studied regional logistics, and first established a predictive index system, expressed demand in terms of cargo turnover, and proposed a combination method based on support vector regression and neural network. At the same time, it uses methods of optimizing thresholds and weights to improve prediction accuracy. Through comparison, it is found that the effect of the combined model of this literature is much higher than that of a single method, and it enhances the generalization ability and application scope of model prediction. The literature [20] elaborated on the characteristics and principles of the four-stage method of transportation planning, and at the same time explained in detail how to apply the staged logistics demand in urban planning. Finally, the empirical part uses the four-stage method to predict.

The literature [21] used the BP neural network method to construct a demand prediction model, continuously optimizes the artificial neural network with the trainbfg function, and continuously adjusted the weights to make the results of the prediction method in this paper closer to the expected results. Finally, simulation research proved the effectiveness of the prediction model in this paper. The literature [22] used the kernel principal component analysis method to process the influencing factors and formed several comprehensive variables as the input variables of the least square support vector machine and established the model. Finally, it was proved by empirical evidence that the model established through nuclear principal component analysis is better than the model without nuclear principal component analysis. The literature [23] used the GM (1, 1) method to predict the freight volume of highway, railway, and waterway in the next three years in an article studying the prediction of industrial logistics demand. The literature [24] used gray relational analysis to select high-relevance influencing factors from many influencing factors, established a logistics demand index system, and used it as an input variable of LSSVM to establish a predictive model. In view of the relatively strong subjectivity of parameter selection in the application of least squares support vector

machine (LSSVM), the literature [24] proposed to optimize the parameters of LSSVM by using the optimization ability of the dynamic acceleration coefficient particle swarm optimization (DACPSO) algorithm. Then, it used the LSSVM prediction method to predict the logistics demand and empirically proved that the model with DACPSO parameter optimization is more accurate.

3 Neural network foundation of logistics model

The McCulloch and Pitts (MP) model had a profound impact on the later perceptron model. The MP model is basically composed of a summing amplifier and a variable resistor, and the corresponding weight is adjusted by changing the value of the resistor. If the sum of the signals exceeds the threshold T, the threshold device composed of voltage comparators outputs 1, or if the signal is less than the threshold T, the threshold device outputs 0. The neuron model is a model that includes input, output, and calculation functions. The input can be analogous to the dendrite of a neuron, the output is analogous to the axon of a neuron, and the calculation is analogous to the nucleus, and different synapses have different weights. As shown in Fig. 1, it is a typical neuron model that contains three inputs, one output, and two calculation processes. First, the model performs weighted summation on all inputs and then passes the weighted summation value as an independent variable to the nonlinear activation function to calculate the output result. In the M-P model, the activation function f is the sgn function, which is used to determine the sign of the independent variable. The formula can be described as [25]:
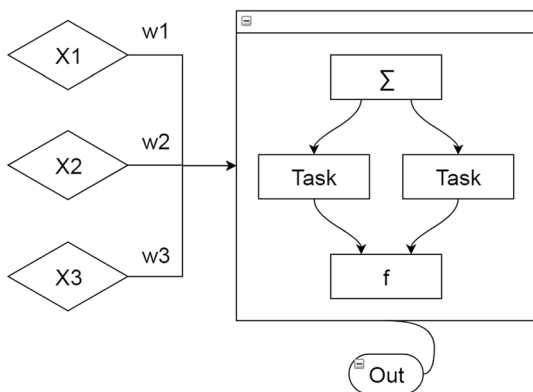
$$y = f\left(\sum_{i=0}^{n} w_i x_i\right) \qquad (1)$$

The network structure of the perceptron model is shown in Fig. 2. The perceptron is composed of two layers of neural networks. The input layer receives external input signals and then passes them to the output layer. The output layer is M-P neurons, also called "threshold logic unit." The activation function f of the perceptron is the sgn function, so the perceptron function can be written as:

$$y = \text{sgn}(\vec{w} \cdot \vec{x})$$
$$\text{sgn}(y) = \begin{cases} 1 & if\, y > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

We can regard the perceptron as a hyperplane in the n-dimensional instance space. For an instance on one side of the hyperplane, the perceptron outputs 1, and for an instance on the other side, the perceptron outputs 0. This decision hyperplane equation is $\vec{w} \cdot \vec{x} = 0$. The set of positive and negative examples that can be divided by a certain hyperplane is called the set of linearly separable examples, and they can be represented by perceptron [26].

If one or more layers of neurons (called hidden layers) are added between the input layer and the output layer, a multi-layer forward network can be formed. When every neuron in one layer is connected to every neuron in the next layer, the network is said to be fully connected. It can handle relatively complex tasks and linearly inseparable classification. A typical example of this problem is the well-known exclusive OR (XOR). As shown in Fig. 3, it is a three-layer perceptron model. The input has three variables $[x_1, x_2, x_3]$, the hidden layer has four neurons, and the output layer has two neurons. Next, we analyze the process of forward propagation. First, we define some variable information. For the l-th layer, we use L_l to denote all neurons in this layer, and the output is y_l. Among them, the output of the j-th node is y_l^j, the input of the node is u_l^j, the weight matrix connecting the l-th layer and the $(l-1)$-th layer is w_l, and the weight of the i-th node in the $(l-1)$-th layer to the j-th node in the l-th layer is w_l^{ji} [27].
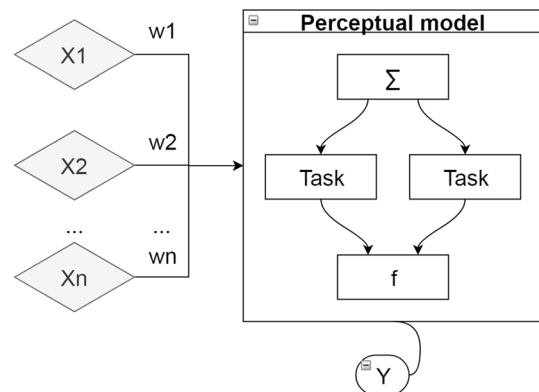

Fig. 1 Basic model of neuron

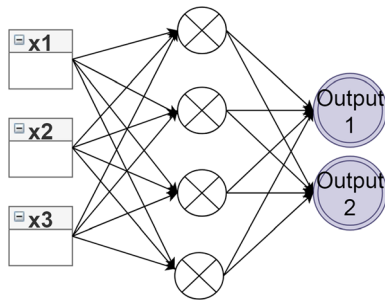
Fig. 2 Basic model of perceptron

Fig. 3 Three-layer perceptron model

Combining the above definitions, we can know that the outputs of the four neurons in the second layer are:

$$y_2^1 = f(u_2^1) = f\left(\sum_{i=1}^n w_2^{ji} x_i\right) =$$
$$f(w_2^{11} x_1 + w_2^{12} x_2 + w_2^{13} x_3) \tag{3}$$

$$y_2^2 = f(u_2^2) = f\left(\sum_{i=1}^n w_2^{ji} x_i\right) =$$
$$f(w_2^{21} x_1 + w_2^{22} x_2 + w_2^{23} x_3) \tag{4}$$

$$y_2^3 = f(u_2^3) = f\left(\sum_{i=1}^n w_2^{ji} x_i\right) =$$
$$f(w_2^{31} x_1 + w_2^{32} x_2 + w_2^{33} x_3) \tag{5}$$

$$y_2^4 = f(u_2^4) = f\left(\sum_{i=1}^n w_2^{ji} x_i\right) =$$
$$f(w_2^{41} x_1 + w_2^{42} x_2 + w_2^{43} x_3) \tag{6}$$

The above formula is transformed into a matrix expression:

$$y_2 = \begin{bmatrix} y_2^1 \\ y_2^2 \\ y_2^3 \\ y_2^4 \end{bmatrix} = f \begin{bmatrix} w_2^{11} & w_2^{12} & w_2^{13} \\ w_2^{21} & w_2^{22} & w_2^{23} \\ w_2^{31} & w_2^{32} & w_2^{33} \\ w_2^{41} & w_2^{42} & w_2^{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = f(W_2 \cdot X) \tag{7}$$

If the forward propagation calculation process of the second layer is extended to any layer in the network, then:

$$\begin{cases} y_l^j = f(u_l^j) \\ u_l^j = \sum_{i \in L_{l-1}} w_l^{ji} y_{l-1}^i \\ y_l = f(u_l) = f(W_l y_{l-1}) \end{cases} \tag{8}$$

4 BP neural network

BP neural network is a multi-layer feedforward network trained by error backpropagation algorithm, and it is one of the most widely used neural network models. The BP network can learn and store a large number of input–output pattern mapping relationships, and this mapping relationship has no clear mathematical equation to express. Its learning rule is to use the steepest descent method to continuously adjust the weights and thresholds of the network through backpropagation to minimize the sum of squared errors of the network.

As shown in Fig. 4, it is a commonly used nonlinear activation function. The activation function can perform a nonlinear combination of network inputs. Figure (a) shows the sigmoid function, which can be expressed as [28]:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

Its characteristic is that it can transform the continuous real number space into output between 0 and 1. If the value is large, the output is 1, and if it is a very large negative number, the output is 0.

Figure (b) shows the tanh function, and the analytical formula is:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{10}$$

The tanh function compresses the input value to the range of [− 1, 1]. Therefore, it is zero mean, so the non-zero-centered problem of the sigmoid function is solved. However, it also has the problems of vanishing gradient and exponentiation.

Figure (c) is the image of the ReLU function. It sets all the negative values of the independent variables to 0, and the other values remain unchanged. Its advantages are fast convergence speed and low computational complexity.

$$\text{ReLU}(x) = \begin{cases} x & if \; x > 0 \\ 0 & if \; x \le 0 \end{cases} \tag{11}$$

Figure (d) is the Leaky ReLU function graph. The curve looks very similar to the ReLU function. The difference is that Leaky ReLU assigns a nonzero slope to all negative values, as shown in the following formula [29]:

$$\text{LeakyReLU}(x) = \begin{cases} x & x > 0 \\ \dfrac{x}{a} & x \le 0 \end{cases} \tag{12}$$

In a neural network problem, we will define a loss function in advance and then use an optimization algorithm to minimize the loss function. In optimization, such a loss function is usually called the objective function of the optimization problem. By convention, optimization algorithms usually only consider minimizing the objective function.

The commonly used objective functions have the following forms:

(1) Mean square error function (MSE):

Fig. 4 Several common nonlinear activation functions

Sigmoid
(a)

Tanh
(b)

Relu
(c)

Leaky Relu
(d)
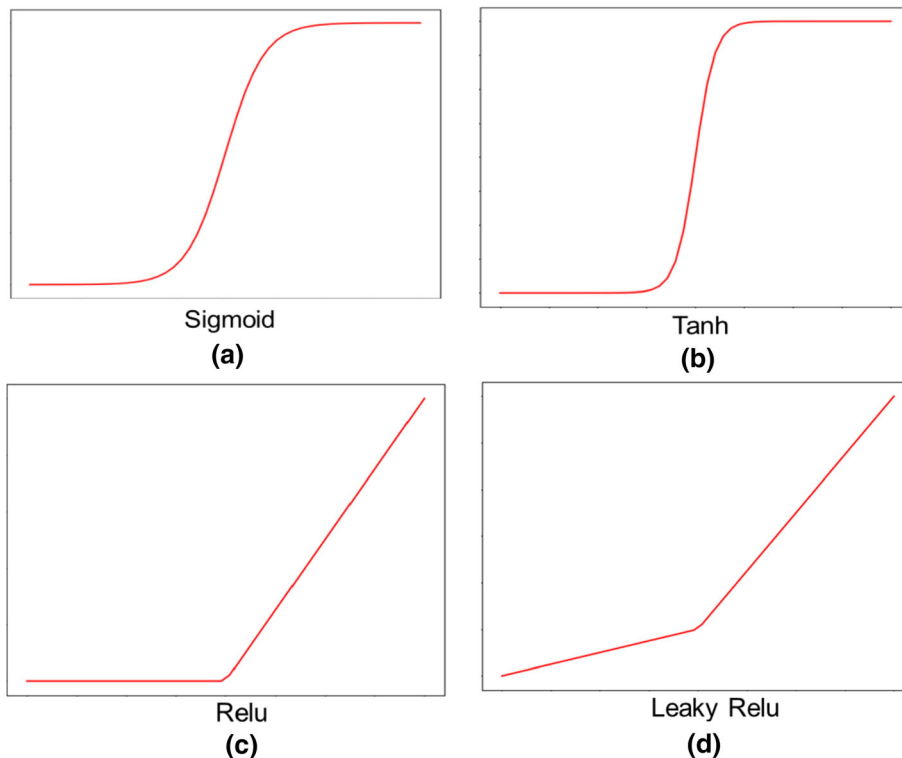
$$\text{mes}(y, y') = \frac{1}{2} \sum_{i=1}^{n} \left(y_i - y'_i \right)^2 \tag{13}$$

The mean square error function is to sum the Euclidean distance between the output value and the label value.

(2) Cross-entropy function:

$$c(y, a) = \frac{1}{n} \sum_{i=1}^{n} [y \ln a + (1 - y) \ln(1 - a)] \tag{14}$$

Cross-entropy function is mostly used in multi-classification problems.

(3) Log likelihood function:

$$c(y, a) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{i} y_k \log a_k \tag{15}$$
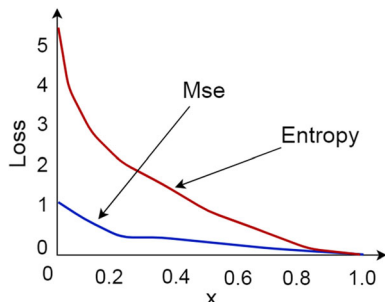

Fig. 5 Schematic diagram of the curve of MSE and cross-entropy function

As shown in Fig. 5, when the expected output value is 1, different loss function formulas give loss variation curves. It can be seen intuitively from the figure that the relationship between the MSE function and the predicted value is a quadratic function, and the cross-entropy curve is an exponential trend. When the predicted value is 0, the loss value given by the cross-entropy function is close to infinity, so it is impossible for the model to make such a wrong prediction during training. The penalty value given by the mean square error is relatively limited. However, the MSE function is relatively simple and performs well for handwritten digit recognition problems. Therefore, the program in this article uses the MSE function as the loss function between the predicted value and the expected value [30].

For complex neural network models, the goal is to minimize the objective function. However, under normal circumstances, the objective function does not have an explicit solution, so it is necessary to use an optimization algorithm based on numerical methods to find an approximate solution. This type of optimization algorithm generally finds an approximate solution by constantly iteratively updating the value of the solution, and there are two challenges in the solution process: local minimum and saddle point.

The gradient descent method is the most commonly used optimization algorithm. First, we take one-dimensional gradient descent as an example to explain why the gradient

descent algorithm can reduce the value of the objective function. One-dimensional gradient descent is a scalar, that is, derivative. We assume that the input and output of the function $f : \mathbb{R} -> \mathbb{R}$ are both scalars. Then, according to the Taylor expansion formula, we get the following formula:

$$f(x + \varepsilon) \approx f(x) + f'(x)\varepsilon \qquad (16)$$

If we assume that η is a constant, we get the following result after we replace ε with $-\eta f'(x)$:

$$f(x - \eta f'(x)) \approx f(x) - \eta f'(x)^2 \qquad (17)$$

If η is a small positive number, then $f(x - \eta f'(x)) \leq f(x)$. That is to say, if the current derivative is $f'(x) \neq 0$, updating the value of x according to $x : x - \eta f'(x)$ may reduce the value of $f(x)$.

The η (taking a positive number) in the above gradient algorithm is called the learning rate. If the learning rate is too large, x will skip the optimal solution, or even continue to diverge and fail to converge. If the learning rate is too small, the convergence speed of the optimization algorithm is very slow. Therefore, setting a reasonable learning rate is very important to optimize the algorithm. As shown in Fig. 6, it is a schematic diagram of finding the minimum value of the function along the gradient descent analysis.

For multi-dimensional gradient descent, we first consider a more generalized situation: The input of the objective function is a vector, and the output is also a vector. We assume that the input of the objective function $f : \mathbb{R}^d -> \mathbb{R}$ is a multi-dimensional vector $x = [x_1, x_2, x_3, \cdots, x_d]^T$. The gradient of the objective function $f(x)$ with respect to x is a vector composed of partial derivatives:

$$\nabla_x f(x) = \left[\frac{\partial f(x)}{\partial(x_1)}, \frac{\partial f(x)}{\partial(x_2)}, \frac{\partial f(x)}{\partial(x_3)}, \cdots, \frac{\partial f(x)}{\partial(x_d)} \right] \qquad (18)$$

For simplicity, $\nabla f(x)$ is sometimes used instead of $\nabla_x f(x)$. Each partial derivative element $\frac{\partial f(x)}{\partial(x_i)}$ in the gradient represents the rate of change of f in x with respect to input x_i. In order to measure the rate of change of f along the
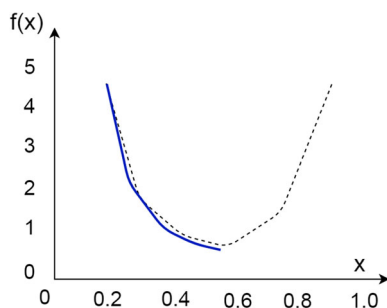

Fig. 6 Schematic diagram of one-dimensional gradient descent

direction of the unit vector u, in multivariate calculus, we define the directional derivative of f along the direction of x as:

$$D_u f(x) = \lim_{h \to 0} \frac{f(x + hu)}{h} \qquad (19)$$

According to the chain derivation rule, the directional derivative can be rewritten as:

$$D_u f(x) = \nabla f(x) \cdot u \qquad (20)$$

The directional derivative $D_u f(x)$ gives the rate of change of f along all possible directions in x. In order to minimize f, we hope to find the direction in which f can decrease the fastest. Therefore, we can use u to minimize the directional derivative $D_u f(x)$.

$D_u f(x) = \|\nabla f(x)\| \cdot \|u\| \cdot \cos(\theta)$. Among them, θ is the angle between $\nabla f(x)$ and u, and when $\theta = \pi$, $\cos(\theta)$ gets the minimum value. Therefore, when u is in the opposite direction of the gradient direction $\nabla f(x)$, the directional derivative $D_u f(x)$ is minimized. Therefore, we can reduce the value of the objective function through the following gradient algorithm:

$$x := x - \eta \nabla f(x) \qquad (21)$$

Similarly, η, taking a positive number, is called the learning rate or step size.

However, when the training set is large, the gradient descent algorithm may be difficult to use. To explain this problem, we consider using the objective function:

$$f(x) = \frac{1}{n} \sum_{i=1}^{n} f_i(x) \qquad (22)$$

Among them, $f_i(x)$ is the loss function of the training data point with index i. The computational cost of each iteration of gradient descent increases linearly with n. Therefore, when n is large, the computational cost of each iteration is high. In order to solve this problem, in each iteration, the algorithm uniformly samples i randomly and calculates $\nabla f_i(x)$. This gradient descent algorithm is called stochastic gradient descent algorithm (SGD). In fact, the stochastic gradient $\nabla f_i(x)$ is an unbiased estimate of the gradient $\nabla f(x)$:

$$E_i \nabla f_i(x) = \frac{1}{n} \sum_{i=1}^{n} \nabla f_i(x) = \nabla f(x) \qquad (23)$$

In a broad sense, each iteration can randomly and uniformly sample a mini-batch composed of training data point indexes. Similarly, we can use

$$\nabla f_B(x) = \frac{1}{|B|} \sum_{i \in B} \nabla f_i(x) \qquad (24)$$

to update x:

$$x := x - \eta \nabla f_B(x) \tag{25}$$

where $|B|$ represents the number of indexes in the batch. Similarly, the batch stochastic gradient is also an unbiased estimate of the gradient:

$$E_B \nabla f_B(x) = \nabla f(x) \tag{26}$$

This algorithm is called batch stochastic gradient descent. The computational cost of each iteration of the algorithm is $O(|B|)$. Therefore, when the batch size is small, the computational overhead of each iteration is also small.

5 Error backpropagation algorithms

The core of the error backpropagation algorithm is to adjust the weight value of each layer according to the error between the output of the forward propagation and the expected value and use the stochastic gradient descent algorithm to iteratively update the weight value. In the end, the error between the output value of the network and the expected value is minimized. The following uses a three-layer network without bias to derive the weight update process of the output layer and the hidden layer.

Figure 7 shows a schematic diagram of a three-layer network. The input vector is: $x = (x_1, x_2, \cdots, x_i, \cdots, x_n)^T$, the hidden layer output vector is: $y = (y_1, y_2, \cdots, y_i, \cdots, y_n)^T$, the output layer output vector is: $O = (o_1, o_2, \cdots, o_i, \cdots, o_n)^T$, the output layer output vector is: $D = (d_1, d_2, \cdots, d_i, \cdots, d_n)^T$, the weight vector from the output layer to the hidden layer is: $V = (v_1, v_2, \cdots, v_i, \cdots, v_n)^T$, the weight vector from the hidden layer to the output layer is: $W = (w_1, w_2, \cdots, w_i, \cdots, w_n)^T$.

$$o_k = f(net_k), k = 1, 2, \cdots, l \tag{27}$$

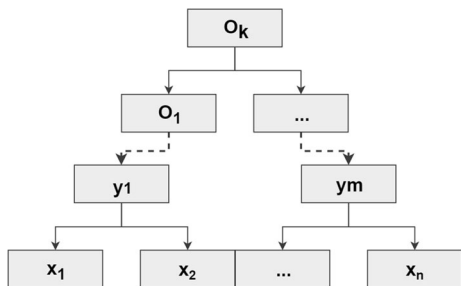$$net_k = \sum_{j=0}^{m} w_{jk} y_j, k = 1, 2, \cdots, l \tag{28}$$

For the hidden layer:

$$y_j = f(net_j), \quad j = 1, 2, \cdots, m \tag{29}$$

$$net_j = \sum_{i=0}^{m} v_{ij} x_i, \quad j = 1, 2, \cdots, m \tag{30}$$

The output error E is:

$$E = \frac{1}{2}(d - o)^2 = \frac{1}{2} \sum_{k=1}^{l} (d_k - o_k)^2 \tag{31}$$

The above error definition is expanded to the hidden layer:

$$E = \frac{1}{2} \sum_{k=1}^{l} [d_k - f(net_k)]^2$$
$$= \frac{1}{2} \sum_{k=1}^{l} \left[d_k - f\left(\sum_{j=0}^{m} w_{jk} y_j \right) \right]^2 \tag{32}$$

The error definition is further expanded to the input layer:

$$E = \frac{1}{2} \sum_{k=1}^{l} \left\{ d_k - f\left[\sum_{j=0}^{m} w_{jk} f\left(\sum v_{ij} x_i \right) \right] \right\}^2 \tag{33}$$

According to the gradient descent method, it can be seen that the two-layer weights are updated as follows:

$$w_{jk} = w_{jk} + \Delta w_{jk} = w_{jk} - \eta \frac{\partial E}{\partial w_{jk}} \tag{34}$$

$$v_{ij} = v_{ij} + \Delta v_{ij} = v_{ij} - \eta \frac{\partial E}{\partial v_{ij}} \tag{35}$$

In the above formula, the negative sign indicates the direction of gradient descent, and the constant $\eta \in (0, 1)$ indicates the learning rate.

6 Logistics demand model construction and performance analysis

The factors affecting logistics demand are shown in Fig. 8.

The GM-ARIMA. BPNN combination model constructed in this paper uses the ARIMA model to analyze the characteristics of non-stationary time data series and improves the gray GM (1, 1) model, so that the gray GM (1, 1) model can study the data samples with complex changes. Moreover, it uses the excellent nonlinear approximation ability of BP neural network to improve the deficiencies of gray GM (1, 1) model in approximating complex nonlinear functions. The improved combination model is suitable for non-stationary time data series with less sample data and complex nonlinear relationships and changing laws. The workflow of the GM-ARIMA. BPNN
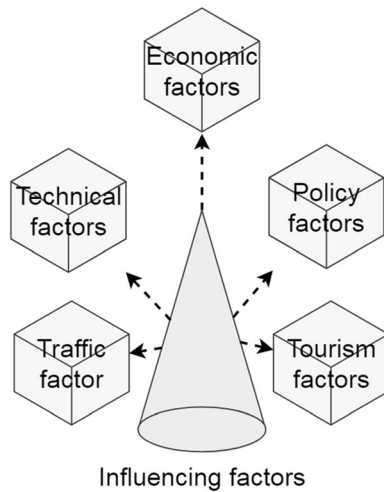

Fig. 7 Schematic diagram of three-layer BP network for the output layer:

Fig. 8 Analysis diagram of influencing factors

combination model constructed in this paper is shown in Fig. 9.

The principal component matrix is shown in Table 1 and Fig. 10.

SPSS17.0 will automatically use these values as new variables, denoted as FAC1-1, FAC2-1, FAC3-1. Moreover, they are stored as new indicators in the original data file, as shown in Table 2. The corresponding statistical graph is shown in Fig. 11.

On the basis of the above analysis, the model performance research is carried out. This article selects the data of Region A from 2005 to 2019 for analysis. Among them, the principal component data and freight volume data from 2005 to 2016 are used as the training sample data to build the model, and the principal component data and freight

Fig. 9 Predictioning process of the combined model

volume data from 2017 to 2019 are used as the prediction sample to test the prediction performance of the model. The model input vector is shown in Table 3.

In order to determine the accuracy of the model prediction results, it is necessary to determine the evaluation index of the model, that is, the evaluation of the accuracy of the model prediction value. In this paper, the commonly used mean square error (MSE) and average absolute percentage error (MAPE) are selected as the basis for comparison and judgment of prediction effects. The smaller the value of the above indicators, the better the prediction effect and the stronger the generalization ability of the model. When the value of MAPE is less than 10, it indicates that the prediction effect is very good. The evaluation index table is shown in Table 4 and Fig. 12.

We use 2015–2019 data to predict the data of FAC1-1, FAC2-1, and FAC3-1 in 2020–2024, as shown in Table 5 and Fig. 13.

After bringing the above results into the established prediction model, the predicted value of freight volume in 2020–2024 can be obtained. The specific results are as follows:

The above results are plotted as a statistical diagram, as shown in Fig. 14.

Consumption, investment, and net exports are the troika of economic growth. Among them, consumption is an important driving force for economic growth, and consumption can have a direct impact on the economy. Moreover, the level of regional logistics development and local social and economic development is mutually reinforcing, and the balance of logistics supply and demand is affected by economic development. When the logistics supply capacity is strengthened, it can stimulate economic growth. Through the prediction and analysis of logistics demand, it is found that the continuous increase in logistics demand has increased national income and increased employment opportunities. However, logistics development in most regions of our country started late and the level of logistics development is low, facing the increasing demand for logistics. In order to better adapt to the changes in the market, only by further strengthening the logistics supply capacity can the balance of logistics supply and demand be achieved. In this way, economic development can be promoted, thereby accelerating the development of logistics. In the process of logistics development, we must pay attention to information construction. For all aspects of logistics, timely feedback and transmission of information can make the entire logistics circulation more flexible. The mature development of Internet technology has also brought new support for the development of logistics. Through the establishment of an information system to realize information sharing, producers can understand consumer needs in a timely manner, thereby producing
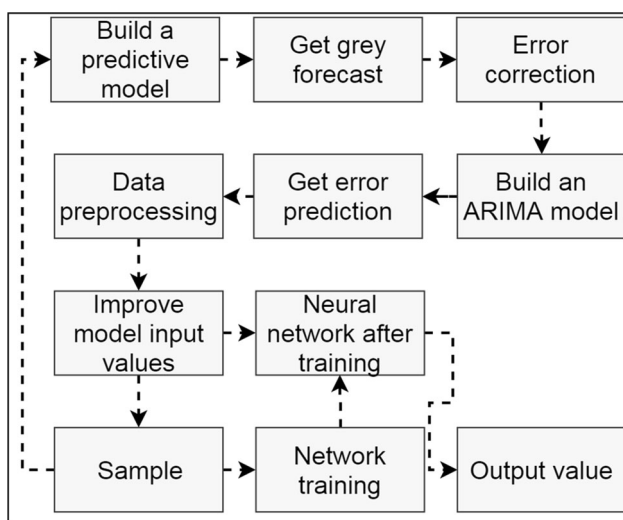
Table 1 Matrix of principal components

	Principal component 1	Principal component 2	Principal component 3
Zscore: Raw coal production	0.274	0.036	0.088
Zscore: Power generation	0.278	0.014	0.014
Zscore: Steel production	0.273	− 0.085	0.042
Zscore: Coke production	0.226	0.366	− 0.119
Zscore: Cement production	0.267	− 0.094	0.087
Zscore: Railway operating mileage	0.254	− 0.110	− 0.017
Zscore: Highway route mileage	0.261	0.159	− 0.149
Zscore: Truck	− 0.001	− 0.035	0.797
	0.196	0.460	− 0.140
Zscore: Practitioners	0.113	− 0.348	− 0.538
Zscore: Number of industrial companies	0.059	0.641	0.003
Zscore: Gross industrial production value	0.275	− 0.044	0.041
Zscore: Total industrial investment in fixed assets	0.268	− 0.154	0.056
Zscore: Industrial value added	0.272	− 0.057	0.048
Zscore: Industrial sales output value	0.275	− 0.069	0.040
Zscore: Total industrial assets	0.270	− 0.144	0.053
Zscore: Total industrial liabilities	0.268	− 0.158	0.050
Zscore: Total industrial profit	0.206	0.088	0.088

Fig. 10 Statistical diagram of the principal component matrix
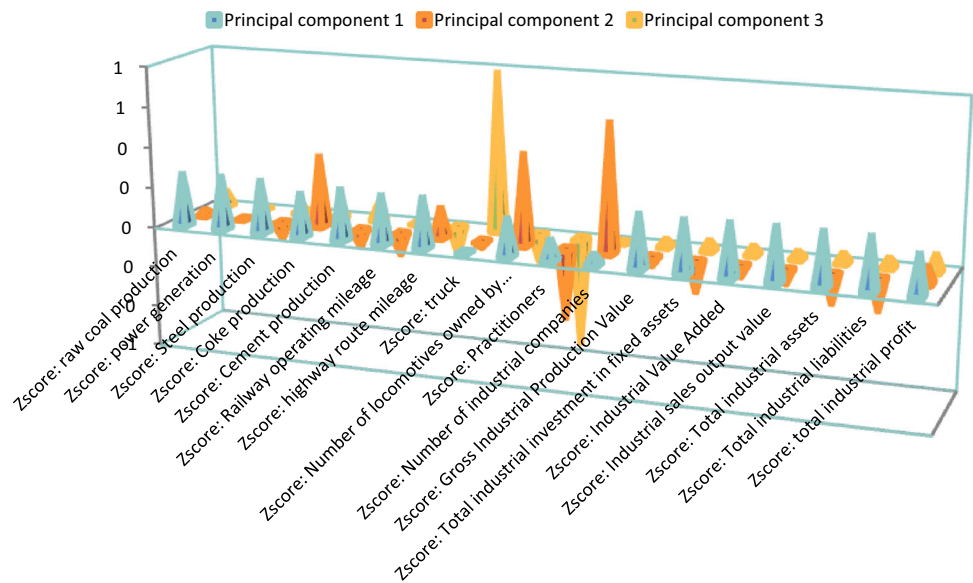
more suitable products, providing information services, facilitating the government's macro-control, and strengthening the government's supervision and management of logistics enterprises.

Since the logistics activities start from the upstream suppliers, continue to the hands of consumers, and even extend to feedback logistics, the entire process is a continuous process. In this process, every link involved needs the support of advanced technology and related equipment.

Therefore, the government should increase investment in the construction of logistics infrastructure, formulate corresponding policies according to different investment methods, and realize that logistics development drives the economic development of the entire region. The construction plan of the logistics center should be carried out in accordance with the logistics circulation links. When formulating a logistics center planning plan, the government should consider the size of the logistics center

Table 2 Principal component score

	FAC1-1	FAC2-1	FAC3-1
2005	− 1.3188984	− 1.8868517	− 2.1866096
2006	− 1.3615002	− 0.893042	0.7662769
2007	− 1.2438049	− 0.606202	0.8258265
2008	− 1.0219685	− 0.2054239	1.1383306
2009	− 0.7537731	1.11807	1.7818117
2010	− 0.4692864	1.059894	− 0.7839014
2011	− 0.1778812	1.2345432	− 1.024847
2012	0.1489851	1.5261403	− 1.0546723
2013	0.184729	0.7538943	− 0.6225741
2014	0.1726191	0.3052018	− 0.1654178
2015	0.6850224	0.5424508	0.2529545
2016	1.0739229	− 0.3977077	0.640138
2017	1.2708022	− 0.860924	0.1273711
2018	1.3727112	− 0.8000412	0.278457
2019	1.4383107	− 0.890012	0.026866

Fig. 11 Principal component score diagram

7 Conclusion
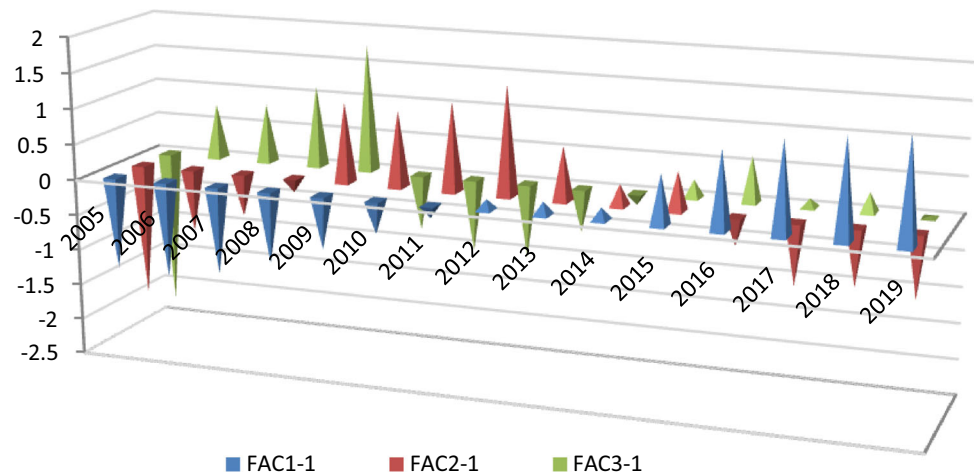
construction in light of the actual situation of the region, and the location of the logistics center should satisfy the optimal allocation of resources and ensure the timeliness of logistics distribution to the greatest extent. Through the construction of a logistics center, the efficiency of logistics circulation is improved, and the cost of logistics is reduced, which greatly promotes the development of logistics.

As there are few studies on logistics demand prediction, there is not much information that can be used for reference when selecting influencing factors. According to the characteristics of product structure and product transportation, and the principles of strong correlation, accessibility, comprehensiveness, and practicality, this paper selects the influencing factor indicators from three aspects: main product indicators, product transportation indicators, and enterprise scale indicators. Due to the relatively strong subjectivity in the selection of the influencing factors in this article, it is impossible to guarantee whether the selection of the indicators is appropriate. Therefore, this paper uses the MLP neural network analysis method to analyze the relevance of the indicators, and the results show that the indicators selected in this paper are all relatively high. Moreover, this article explains the accuracy of index selection from a quantitative aspect, thereby establishing a logistics demand forecast index system. The commonly used prediction methods are gray model, linear

Table 3 Model input vector

	FAC1-1	FAC2-1	FAC3-1	Freight volume
2005	− 1.3188984	− 1.8868517	− 2.1866096	− 1.6719136
2006	− 1.3615002	− 0.893042	0.7662769	− 1.3482187
2007	− 1.2438049	− 0.606202	0.8258265	− 1.0984053
2008	− 1.0219685	− 0.2054239	1.1383306	− 0.8346539
2009	− 0.7537731	1.11807	1.7818117	− 0.4393904
2010	− 0.4692864	1.059894	− 0.7839014	0.0481265
2011	− 0.1778812	1.2345432	− 1.024847	0.3444201
2012	0.1489851	1.5261403	− 1.0546723	0.7745286
2013	0.184729	0.7538943	− 0.6225741	0.1235533
2014	0.1726191	0.3052018	− 0.1654178	− 0.6393098
2015	0.6850224	0.5424508	0.2529545	0.0174932
2016	1.0739229	− 0.3977077	0.640138	0.6063131
2017	1.2708022	− 0.860924	0.1273711	0.9029703
2018	1.3727112	− 0.8000412	0.278457	1.4102428
2019	1.4383107	− 0.890012	0.026866	1.8042539

Table 4 Evaluation index

	2017	2018	2019
Actual value	0.903	1.410	1.804
Predictive value	0.945	1.432	1.761
Relative error	4.737%	1.545%	− 2.424%
Mean squared error, MSE	0.001		
Mean absolute percentage error, Mape	2.898		

regression analysis, Delphi method, etc. However, each method has its characteristics and applications. The gray model is the most commonly used time series method, which requires high smoothness and simple research problems. Regression analysis requires a linear relationship

between various variables. The Delphi method is applicable to situations where there is no historical data for the research content. However, the above methods are not applicable when the influencing factor data have a nonlinear relationship, high latitude, large scale, and nonnormal distribution. At this point, we need to use machine learning methods to make predictions. Commonly used machine learning methods include support vector machines (SVM) and BP neural network algorithms. The main principle of BP neural network algorithm is based on the
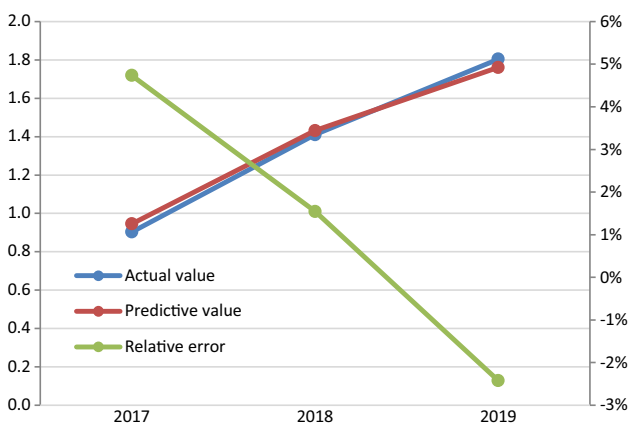

Fig. 12 Evaluation index diagram

Table 5 Raw data

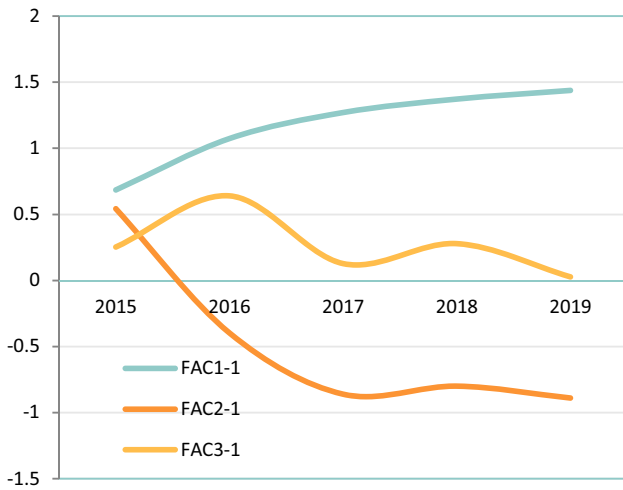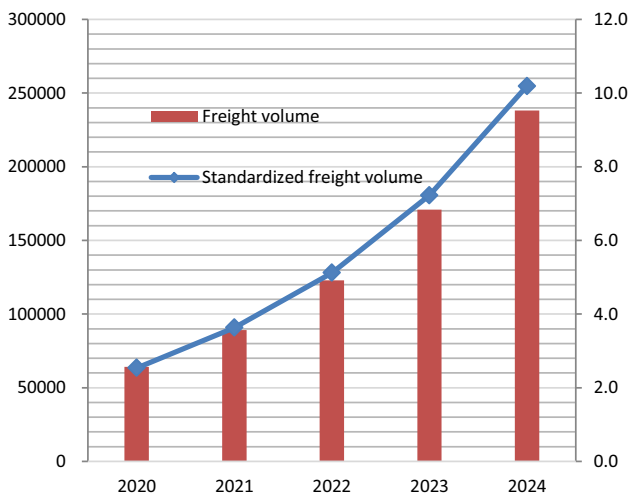|	FAC1-1	FAC2-1	FAC3-1
2015	0.6850224	0.5424508	0.2529545
2016	1.0739229	− 0.3977077	0.640138
2017	1.2708022	− 0.860924	0.1273711
2018	1.3727112	− 0.8000412	0.278457
2019	1.4383107	− 0.890012	0.026866

Fig. 13 Raw data diagram

Fig. 14 Prediction result diagram of freight volume

Table 6 Prediction result of freight volume

	Standardized freight volume	Freight volume
2020	2.5	64,153.2
2021	3.6	89,042.6
2022	5.1	122,952.4
2023	7.2	170,779.9
2024	10.2	238,243.9

principle of empirical risk minimization, and its shortcomings are low generalization ability and difficult to determine the model structure. When the number of samples is small, the prediction process easily reaches a local minimum, and it is difficult to ensure the accuracy of the prediction. If there are many prediction samples, it is easy to lead to "dimensionality disaster." However, the MLP neural network method can solve all the above problems. This paper uses principal component data and freight volume data as training sample data to establish a model. It can be seen from the relative error in the prediction result comparison table that each (Table 6) evaluation index is relatively small, which shows that the prediction result is accurate. Finally, this article proves the feasibility of the model to predict logistics demand through empirical research.

Compliance with ethical standards

Conflict of interest The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. You SI, Chow JYJ, Ritchie SG (2016) Inverse vehicle routing for activity-based urban freight forecast modeling and city logistics[J]. Transportmetrica A: Transp Sci 12(7):650–673
2. Liu M, Zhang D (2016) A dynamic logistics model for medical resources allocation in an epidemic control with demand forecast updating[J]. J Operational Res Soc 67(6):841–852
3. Ya B (2016) Study of food cold chain logistics demand forecast based on multiple regression and AW-BP forecasting method on system order parameters[J]. J Comput Theor Nanosci 13(7):4019–4024
4. Li R, Zhang WY (2016) Energy demand forecast in the logistics sector based on RBF neural networks[J]. Resour Sci 38(3):450–460
5. He Y, Liu N (2015) Methodology of emergency medical logistics for public health emergencies[J]. Transp Res Part E: Logist Transp Rev 79:178–200
6. Li L, Liu Y (2014) Nonlinear regression prediction of the social logistics demand forecast in our country[J]. J Jiangnan Univ (Natural Science Edition) 3(13):375–377
7. Yu N, Xu W, Yu KL (2020) 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[J]. IEEE Access 8:9551–9564
8. Sheu JB, Kundu T (2018) Forecasting time-varying logistics distribution flows in the one belt-one road strategic context[J]. Transp Res Part E: Logist Transp Rev 117:5–22
9. van der Laan E, van Dalen J, Rohrmoser M et al (2016) Demand forecasting and order planning for humanitarian logistics: an empirical assessment[J]. J Operations Manag 45:114–122
10. Kosenko V, Gopejenko V, Persiyanova E (2019) Models and applied information technology for supply logistics in the context of demand swings[J]. Innovative technologies and scientific solutions for industries. 10.30837/2522-9818.2019.7.059
11. Thomé AMT, Hollmann RL, Carmo L (2014) Research synthesis in collaborative planning forecast and replenishment[J]. Ind Manag Data Syst 114(6):949–965

12. Zhu J, Zhang H, Zhou L (2015) Research on logistics demand forecasting and transportation structure of Beijing based on grey prediction model[J]. Sci J Appl Math Stat 3(3):144–146

13. Nuzzolo A, Comi A (2014a) A system of models to forecast the effects of demographic changes on urban shop restocking[J]. Res Transp Bus Manag 11:142–151

14. Wu J, Han X (2015) Forecasting and analysis of jiangsu province's logistics demand based on multiple line-ar regression model[J]. Acta Agriculturae Shanghai 31(4):62–68

15. Rostami-Tabar B, Babai MZ, Syntetos A et al (2014) A note on the forecast performance of temporal aggregation[J]. Naval Res Logistics (NRL) 61(7):489–500

16. Wang G, Gunasekaran A, Ngai EWT et al (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications[J]. Int J Prod Econ 176:98–110

17. Kwesi-Buor J, Menachof DA, Talas R (2019) Scenario analysis and disaster preparedness for port and maritime logistics risk management[J]. Accid Anal Prev 123:433–447

18. Sellitto MA (2018) Reverse logistics activities in three companies of the process industry[J]. J Clean Prod 187:923–931

19. Perboli G, Musso S, Rosano M (2018) Blockchain in logistics and supply chain: a lean approach for designing real-world use cases[J]. IEEE Access 6:62018–62028

20. Barreto L, Amaral A, Pereira T (2017) Industry 4.0 implications in logistics: an overview [J]. Procedia Manuf 13:1245–1252. https://doi.org/10.1016/j.promfg.2017.09.045

21. Choi TM, Wen X, Sun X et al (2019) The mean-variance approach for global supply chain risk analysis with air logistics in the blockchain technology era[J]. Transp Res Part E: Logist Transp Rev 127:178–191

22. Prakash C, Barua MK (2015) Integration of AHP-TOPSIS method for prioritizing the solutions of reverse logistics adoption to overcome its barriers under fuzzy environment[J]. J Manuf Syst 37:599–615

23. Wang DZ, Lang MX, Sun Y (2014) Evolutionary game analysis of co-opetition relationship between regional logistics nodes[J]. J appl res technol 12(2):251–260

24. Havenga JH, Simpson ZP, De Bod A et al (2014) South Africa's rising logistics costs: an uncertain future[J]. J Transp Supply Chain Manag 8(1):1–7

25. Winkelhaus S, Grosse EH (2020) Logistics a systematic review towards a new logistics system[J]. Int J Prod Res 58(1):18–43

26. Bhattacharya A, Kumar SA, Tiwari MK et al (2014) An intermodal freight transport system for optimal supply chain logistics [J]. Transp Res Part C: Emerging Technol 38:73–84

27. Bouzon M, Govindan K, Rodriguez CMT (2018) Evaluating barriers for reverse logistics implementation under a multiple stakeholders' perspective analysis using grey decision making approach [J]. Resour Conserv Recycl 128:315–335

28. Tunji-Olayeni PF, Afolabi AO, Ojelabi RA et al (2017) Impact of logistics factors on material procurement for construction projects [J]. Int J Civil Eng Technol IJCIET 8(12):1142–1148

29. Nuzzolo A, Comi A (2014b) City logistics planning: demand modelling requirements for direct effect forecasting [J]. Procedia-Social Behav Sci 125:239–250

30. Nowicka K (2014) Smart city logistics on cloud computing model [J]. Procedia-Social Behav Sci 151((Supplement C)):266–281