Applied Mathematics, Modeling and Computer Simulation
C.-H. Chen et al. (Eds.)

doi:10.3233/ATDE221088

Prediction of Agricultural Products Logistics Demand in Five Provinces of North China Based on BP Neural Network

Wei DONG, Kexin ZHANG, Min ZUO[1], Wenjing YAN, Qingchuan ZHANG
National Engineering Research Centre for Agri-product Quality Traceability, Beijing Technology and Business University, Beijing, 100048, China.

Abstract. Agricultural product logistics is the key to ensuring people's livelihood. The prediction of agricultural products' logistics demand is an important guarantee for the rational planning of agricultural products logistics. However, the demand for agricultural product logistics is affected by many factors, which increases the complexity of its prediction. Therefore, taking the logistics demand of agricultural products as the research object, this paper constructs an index system from five aspects: the level of economic development, the level of industrial structure, the level of logistics development, the supply factors of agricultural products and human factors. Using the nonlinear mapping ability of the BP neural network, this paper constructs a BP neural network model to predict the logistics demand of agricultural products and takes the five provinces in North China as an example to predict the logistics demand of agricultural products. The results show that the established model has a strong ability to describe the nonlinear relationship between agricultural products logistics demand and its influencing factors, and can provide a basis for rational planning and policy- making of agricultural products logistics.

Keywords. Agricultural products logistics, Demand prediction, BP neural network

1. Introduction

The prediction of agricultural products' logistics demand is an important guarantee for the rational planning of agricultural products logistics. However, the demand for agricultural product logistics is affected by many factors, which increases the complexity of forecasting the demand for agricultural product logistics. North China includes two municipalities directly under the central government, Beijing and Tianjin, and three provinces, Hebei, Shanxi, and Inner Mongolia. It has a certain influence on both economic development and logistics scale. And since the five provinces of North China are geographically adjacent and supply food to each other, they share some common factors in agricultural products logistics demand. Therefore, it is very necessary to predict the logistics demand of agricultural products in the five provinces of North China.

1 Corresponding Author, Min ZUO, National Engineering Research Centre for Agri-product Quality Traceability, Beijing Technology and Business University, Beijing, 100048, China; E-mail: zuomin2022@126.com.

In the existing research on demand forecasting, most methods are proposed based on linear fitting concepts, such as multiple linear regression[1]. BP neural network has a strong nonlinear mapping ability and flexible network structure, which exhibits great advantages in dealing with nonlinear problems. Therefore, this paper applies BP neural network model to predict the logistics demand of agricultural products in the five provinces of North China, to provide a basis for the rational allocation of agricultural products logistics resources, future development planning, and policy formulation in the five provinces of North China.

2. Datasets and Index

2.1. Datasets

When forecasting the demand for agricultural product logistics in the five provinces of North China, the primary work is data collection. The basic data used in this paper are obtained from the Beijing Statistical Yearbook, Tianjin Statistical Yearbook, Hebei Statistical Yearbook, Shanxi Statistical Yearbook, Inner Mongolia Statistical Yearbook from 2000 to 2021, and Statistical bulletin of national economic and social development of five provinces in 2021.

2.2. Index

The logistics demand forecast of agricultural products needs to comprehensively consider a variety of influencing factors. Based on the analysis of the actual data from 2000 to 2021, the main influencing factors of agricultural products' logistics demand in the five provinces of North China are screened according to the principles of comprehensiveness, applicability, and availability. Finally, the level of economic development, the level of industrial structure, the level of logistics development, the supply of agricultural products, and human factors are taken as the first level indicators[2].

The economic development level includes: per capita GDP (100 million yuan) X1, average disposable income (yuan) X2; The level of industrial structure includes: the added value of the primary industry (RMB 100 million) X3, the added value of the secondary industry (RMB 100 million) X4, and the added value of the tertiary industry (RMB 100 million) X5; Logistics development level includes: freight volume (10000 tons) X6, freight turnover volume (100 million tons km) X7, highway mileage (km) X8; The supply factors of agricultural products include: annual output of agricultural products (10000 tons) X9, planting and breeding area (10000 hectares) X10, the total power of agricultural and animal husbandry machinery (10000 kW) X11; Human factors include: population (10000 people) X12, per capita consumption expenditure (yuan) X13. The above 13 influencing factors and provinces are used as secondary indicators to measure the logistics demand for agricultural products in the five provinces of North China.

The agricultural products in this paper include eight categories: grain, oil, vegetables, fruits, meat, eggs, milk and aquatic products. Among these indicators, the annual output of agricultural products X9 is obtained by adding the annual output of eight types of agricultural products, and the planting and breeding area X10 is obtained by adding the planting and breeding areas of eight types of agricultural products.

This paper uses the number of permanent residents in each province multiplied by the per capita consumption of agricultural products in each province as the dependent variable Y, which is the logistics demand of agricultural products[3]. The per capita consumption of agricultural products is the sum of the per capita consumption of eight types of agricultural products. The calculation formula is as follows:

$$Y = M \times N \tag{1}$$

M——number of permanent residents in each province

N——per capita consumption of agricultural products in each province

Using 14 influencing factors highly related to the logistics demand of agricultural products as independent variables, a prediction model for the logistics demand of agricultural products in the five provinces of North China is established. The datasets example is presented in table 1.

Table 1. Sample datasets of indicators affecting agricultural products logistics demand in five provinces of North China

Year	Province	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
2019	Beijing	800.04	161776	67756	114.4	5667.4	29663.4	27338.22	901	22365.94	233.619	12.58	122.8	2190.1	43038
2019	Tianjin	574.08	101557	42404	185.41	4947.18	8922.87	56941	2244	16132	625.59	43.42	359.84	1385	31854
2019	Hebei	2656.56	47036	25665	3518.4	13597.3	17988.8	374501	14179.53	196983	11308.37	813.27	7830.73	7446.56	17987
2019	Shanxi	1189.04	48468.67	23828	825.34	7466.3	8669.97	219312	4690	144283	3311.25	352.44	1517.57	3496.88	15863
2019	Inner Mongolia	861.19	71170	30555	1863.26	6763.14	8586.13	182706.62	4586.84	206089	5940.48	888.5	3866.42	2415.3	20743
2020	Beijing	840.79	164889	69434	107.6	5716.4	30278.6	26345.86	843	22264	251.4	13.57	120.2	2189	38903
2020	Tianjin	573.36	101614	43854	210.18	4804.08	9069.47	53566	2371	16411.02	655.67	44.29	365.08	1386.6	28461
2020	Hebei	2838.72	48564	27136	3880.1	13597.2	18729.6	368246	13735.52	204737	11542.50	808.9	7965.74	7463.84	18037
2020	Shanxi	1345.80	50527.92	25214	946.68	7675.44	9029.81	190238	5712	144323	3491.73	354.15	1595.26	3490.5	15733
2020	Inner Mongolia	935.48	72062	31497	2025.12	6868.03	8466.66	170550.26	4431.47	210217	5961.32	888.3	4057.14	2402.8	19794
2021	Beijing	868.42	183980	75002	111.3	7268.6	32889.6	28132	881	22290	265.81	13.04	111.55	2188.6	43640
2021	Tianjin	591.95	113732	47449	225.41	5854.27	9615.37	57568	2682	16895.45	638.55	56.07	340.4	1373	33188
2021	Hebei	2831.07	54172	29383	4030.3	16364.2	19996.7	388868	14774.6	206989	11844.35	816.8	8096.81	7448	19954
2021	Shanxi	1339.15	64821	27426	1286.87	11213.13	10090.16	217637	6444.7	147069.57	3358.14	351.51	1487.86	3480.48	17191
2021	Inner Mongolia	923.37	85422	34108	2225.2	9374.2	8914.8	211904	4891.7	212603	6048.89	902	4239.4	2400	22658

2.3. Data Normalization

Due to the different dimensions of each input data of the model, the variation range of the input data varies greatly. Therefore, it is necessary to normalize the data to prevent small value information from being drowned. The normalization operation is carried out according to the following formula:

$$x'=(x-x_{min})/(x_{max}-x_{min}) \tag{2}$$

Where, x' is the normalized data; x is the original data, x_{min} and x_{max} are the minimum and maximum values in the original data, respectively.

3. BP Neural Network

3.1. Model Structure

BP neural network is a kind of multilayer feed-forward neural network. It trains the network according to the back propagation of the error, estimates the error of the leading layer of the output layer by using the error after the output, and calculates the error of the upper layer according to this error. It is calculated in this way until the error of each layer is obtained.

BP neural network is composed of three parts: input layer, hidden layer, and output layer. Layers are fully connected, and there is no interconnection between the same layer. The input layer directly connects the input variables, the hidden layer realizes the nonlinear mapping from the input layer to the hidden layer, and the output layer realizes the linear output of the output variables of the hidden layer[4].

The Activation function is a method of transforming input data in the neural network. In the hidden layer of the neural network, the activation function is responsible for summarizing the information entering the neuron and converting it into a new output signal, and then passing the new output signal to the next neuron. Its function is to increase the nonlinearity of the neural network model. Relu is a function that takes the maximum value, which is a commonly used activation function of BP neural network. Relu will make the output of some neurons 0, which can make the network more sparse. Also, it can reduce the interdependence of parameters, and alleviate the over fitting problem. Compared with other activation functions, Relu can effectively solve the gradient disappearance problem and improve the calculation speed. Its mathematical formula is as follows:

$$f(x) = \max(x, 0) \tag{3}$$

The Loss function is a way to measure the difference between the predicted value and the actual value of the output of the neural network. MSE is a commonly used loss function, which represents the expected value of the square of the difference between the predicted value and the actual value. It reflects the degree of change in the data. The smaller the value of MSE, the higher the accuracy of the experimental data of the prediction model. Its mathematical formula is as follows:

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(predicted_t - observed_t)^2 \quad (4)$$

Figure 1 shows the BP neural network structure of logistics demand prediction of agricultural products in five provinces of North China.

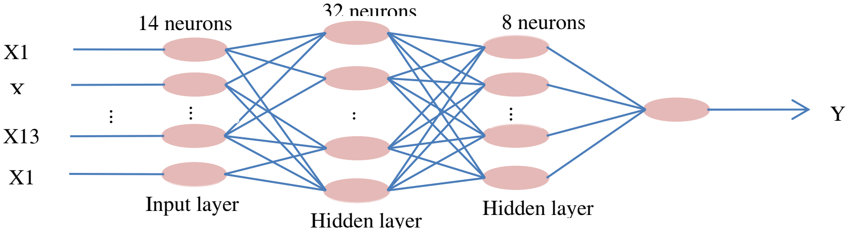

Figure 1. BP neural network structure for forecasting the logistics demand of agricultural products in the five provinces of North China

3.2. Model Training

This paper uses Python language programming to train BP neural network for forecasting the logistics demand of agricultural products in the five provinces of North China. We apply 5-fold cross-validation for the model training. More specifically, the original dataset is divided into 5 approximate average folds. At each training procedure, it selects 4 folds as the training set, and the remaining 1 fold as the testing set. The training is repeated 5 times, and the selected training sets and testing sets are different each time.

To avoid over fitting and increase the generalization ability of the model, dropout is added to the neural network model. The training parameters of the model are listed in table 2, and the final parameters are chosen based on the best performance model.

Table 2. Training parameter setting

Parameters	**Parameter Setting**
Dropout	[0.1,**0.2**,0.5]
Activation	Sigmoid, **Relu**, Tanh
loss	MAE, **MSE**, RMSD
Optimizer	AdaGrad, **Adam**, RMSProp, AdaDelta
Epochs	[100,**1000**]
Batch_size	4
Learning rate	[0.001, **0.01**, 0,1]

*The finally chosen parameters are labeled with BOLD.

4. Result

This paper constructs a four-layer BP neural network with an input layer, two hidden layers, and an output layer. 13 indicators affecting the logistics demand of agricultural products and the province are selected as the input of the network, so the number of input nodes is 14. The logistics demand of agricultural products is taken as the output of the network, so the number of output nodes is 1. The nodes of the two hidden layers are set to 32 and 8. The 5 training results are shown in figure 2.

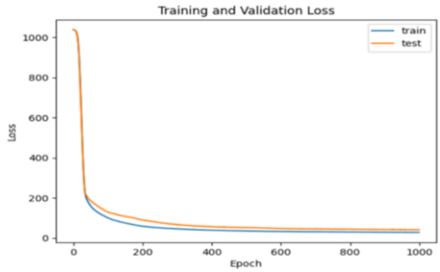

Figure 2. Training and Validation Loss

As can be seen from figure 2, the loss function value is large at the beginning of the iteration, rapidly decreases during the iteration, and then gradually flattens. When the number of iterations is 1000, the loss function value is close to 0.

Table 3 shows the absolute error and percent error of each training set and test set, as well as the average absolute error and average percent error of five-fold cross-validation.

Table 3. BP neural network training and test error

Fold Number	Training error		Test error	
	Absolute error	Percent error	Absolute error	Percent error
1-fold	31.14	3.21%	33.24	4.23%
2-fold	27.36	3.23%	47.03	4.23%
3-fold	33.35	3.43%	28.51	3.80%
4-fold	24.6	2.69%	37.9	5.33%
5-fold	25.98	3.28%	57.84	5.40%
Average	28.49	3.17%	40.91	4.60%

It can be seen from table 3 that the absolute error of each training set is 28.49 in the 5-fold training set and 40.91 in the 5-fold test set, indicating that there is little difference between the predicted value and the actual value. In the training set, the average error rate of 5-fold cross-validation is 3.17%, and in the test set, the average error rate of 5-fold cross-validation is 4.60%, and the percentage of training and testing has reached about 96%. This shows that the BP neural network prediction model has high prediction accuracy and practical value, and can be used to predict the logistics demand of agricultural products in the five provinces of North China.

5. Conclusion

This paper constructs a BP neural network model to predict the logistics demand of agricultural products in the five provinces of North China. Select the actual value and the predicted value of agricultural products logistics demand in Beijing in 2019 for comparison and analysis. According to the prediction results, the predicted value is 826.45, which is very close to the actual value of 800.04, and the absolute percentage error is within 1%. The result shows our model could extract nonlinear characteristics of indexes and achieve a good performance in agricultural products' logistics demand prediction. Therefore, the method used in this paper can provide a basis for the reasonable planning of agricultural product logistics in the five provinces of North China, and has practical value.

Acknowledgments

This study is supported by Beijing Natural Science Foundation (No.4202014), Natural Science Foundation of China (61873027), Humanity and Social Science Youth Foundation of Ministry of Education of China (No.20YJCZH229).

References

[1] Liu Jiong. Logistics demand forecasting analysis based on multiple linear regression:Taking Anhui province as an example.Sichuan University of Arts and Science Journal.2022,32(2):51-57
[2] Li Minjie, Wang Jian. Prediction of demand for cold-chain logistics of aquatic products based on RBF neural network. Chinese Journal of Agricultural Resources and Regional Planning.2020,41(6):100-108
[3] Wang Xiaoping, Yan Fei. Prediction of cold chain logistics demand for agricultural products inBeijing based on GA-BP model[J].Mathematics in Practice and Theory,2019,49 (21):18-27
[4] Chen Yongdang, Liu Shan, Qin Shujuan,et al.Prediction of maximum deformation of single nail riveting based on BP neural network[J].Journal of Xi'an Polytechnic University,2021,35(6):90-95