Z.B. Hu et al. (Eds.)

doi:10.3233/ATDE231338

Cold Chain Logistics Demand Forecasting Based on Improved BP Neural Network Model

Xudong DENG, Boheng ZHANG[1], Yong WANG
School of Management, Wuhan University of Science and Technology, Wuhan, China

Abstract. According to the characteristics of continuous value prediction and the imitation of the sample data size of cold chain logistics demand, an improved BP neural network model based on gating mechanism and GELU is proposed. Through gating mechanism, neural network pays more attention to important features in input, and GELU is used as the activation function of neural network, which can provide better nonlinear ability. In the empirical process, the model is used to forecast the demand of cold chain logistics of agricultural products in Hubei Province. The improved BP neural network achieves 96.07% similarity on the test dataset, which represents the state of the art in cold chain logistics demand forecast.

Keywords. Cold chain logistics demand forecast; BP neural network; Gating mechanism; GELU; Adam optimizer; Sub-models.

1. Introduction

Forecasting and analyzing the cold chain logistics demand of fresh agricultural products is an important basis for measuring the current level of cold chain logistics, which can promote the process of restoring economic development to ensure that cold chain logistics services achieve a relative balance between supply and demand, and then improve the operational efficiency of cold chain logistics.

With the development of emerging technologies such as the Internet of Things, big data, and blockchain, machine learning methods are also more used in the field of prediction. For example, in the industrial sector, Awad et al. constructed artificial neural networks using different types of optimization algorithms, successfully predicting the water demand in the Jenin city of Palestine [1]. In the aviation field, Mamdouh et al. utilized machine learning to build a model for predicting ground service resource demand by constructing future flight schedule resource demand curves, which has been proven to have good accuracy [2]. In the medical field, Howlader et al. compared naive Bayes, decision trees, random forests, and logistic regression using data mining techniques and accurately predicted heart diseases [3].

This paper introduces state-of-the-art deep learning techniques into cold chain logistics demand forecasting. The bp neural network is improved through GELU and the gating mechanism, the Adam optimizer is used in the training process, and the sub-model prediction input is used when performing the forecast, which not only improves the

[1] Corresponding Author: ZHANG Boheng, E-mail: boxdada@qq.com.

accuracy of the prediction results, but also improves the stability of the model. This study can not only provide reference and suggestions for cold chain logistics demand forecasting, but also provide new perspectives and ideas for research related to continuous value forecasting. Effective demand management based on the prediction and analysis results, guiding funds from all aspects of society to reasonably enter the field of agricultural product cold chain logistics services, is conducive to scientific and effective logistics system planning development of cold chain logistics, and effectively avoiding resource waste and overcapacity.

2. Literature Review

2.1. Research Status

Currently, traditional logistics demand forecasting methods include time series analysis, causal analysis, and random analysis [4]. Due to the advantages of machine learning and various algorithm optimizations, neural networks often have higher accuracy compared to traditional forecasting methods [5]. In recent years, many scholars have used neural networks to predict logistics demand. For example, Eksoz C used a neural network model and grey model to comprehensively analyze factors affecting demand forecasting and scientifically predict short-term cold chain logistics demand [6]. Huang L et al. used the GM (1,1) model and BP neural network model to simulate and forecast logistics demand. The results showed that the BP neural network model had smaller prediction errors and more stable results [7]. Munkhdalai L et al. compared and analyzed multiple linear regression models, MLP models, LSTM models, UBER-LSTM models, and MLP-SUL models, and concluded that the MLP-SUL model was the most effective for predicting logistics demand in Korea [8]. Ma H et al. combined the logistic regression algorithm to construct a neural network algorithm model and predicted logistics demand through examples [9]. Yu N et al. combined ant colony algorithm with support vector machine (SVM) to predict urban logistics demand. The experimental results showed that the improved SVM had higher prediction accuracy, stronger stability, lower error rate, and more realistic prediction results [10]. Zhang G et al. proposed a combination forecasting model that combines the ARIMA time series forecasting model with BP neural network. Through simulation experiments, it was found that the combination forecasting model could more comprehensively reflect the changing patterns of logistics demand [11].

When selecting logistics demand influencing factors, scholars have established diverse logistics demand forecasting indicator systems based on different situations. For example, Nguyen TY proposed from the perspective of Southeast Asian logistics development that GDP, regional logistics volume growth rate, logistics regional attractiveness, regional logistics distribution, and regional distance had an inseparable impact on logistics demand [12]. Du B et al. proposed that national freight volume represented logistics demand and selected four indicators as factors affecting logistics demand [13].

Many scholars have established diverse logistics demand forecasting indicator systems based on different situations. Feng Y selected social logistics volume as a substitute variable to predict fresh agricultural products cold chain logistics demand. He predicted social logistics volume by constructing a prediction model that combined BP neural network and principal component regression analysis [14]. At the level of fresh agricultural product consumption, Wang S selected 14 indicators affecting fresh agricultural product consumption from five aspects: regional development level, market supply and demand factors, industry structure level, location advantage factors, and logistics industry factors, and then constructed a combined prediction model based on SVR [15]. Wang X predicted fresh agricultural product consumption from five aspects: agricultural product supply, cold chain level, socio-economic indicators, logistics demand scale, and human development angle, and optimized neural network using genetic algorithm [16].

2.2. Innovation

Different from the above research, this paper is more focusing on introducing the state of the art deep learning techniques into cold chain logistics demand forecast. The major contributions are: 1. Utilizing GELU and gating mechanism in the neural network. 2. Leveraging Adam optimizer during training. 3. Using sub-models to predict input when doing forecasting.

Current research only utilizes machine learning at a primary level, and majorly focuses on feature engineering. However, with more advance techniques, the most learn the relationship between the input and the label. For example, most of the research mentioned above use RELU as activation function, but GELU as an update on RELU has been proved to be much more effective in a wide range of applications. Although RELU function has the advantages of simple calculation, fast speed and outstanding performance in solving parallel data, RELU function has only two output situations, and completely ignores the negative part of input, which may lead to the activation function not making full use of input information, thus affecting the forecasting effect. Therefore, this paper uses GELU to retain more useful information in the data, so that it can bring better nonlinear ability in the process of forecasting agricultural logistics demand.

At the same time, in the process of model training, facing the pain points such as limited data volume of cold chain logistics prediction, most of the existing researches use Attention mechanism to train the prediction model. Although Attention mechanism is widely used in AI and other scenarios as the mainstream, it usually needs a large amount of data to learn reasonable weight parameters, and cannot learn the sequence relationship in the sequence [17]. Therefore, the Attention mechanism may not be effective in predicting continuous values with limited data.

The reason why this paper did not choose the SGD optimizer used by most scholars is that SGD is unstable and prone to local optimality. Therefore, this paper chooses to use the more advanced Adam optimizer. Adam adds second-order momentum based on SGD. Adam adds second-order momentum based on momentum SGD and controls the step size through adaptive learning. When the gradient is small, the overall learning rate will increase, otherwise it will shrink. Therefore, in general, Adam has a faster and more stable convergence speed than SGD. Finally, the existing literature on cold chain logistics demand forecasting describes the input very simply when forecasting. This method can easily lead to the accumulation of errors. Therefore, we use sub-models to eliminate errors and improve the final forecasting results. accuracy.

3. Research Method

3.1. Model Introduction

In essence, the prediction of the total output value of agricultural products belongs to the continuous value prediction problem.

Suppose we have k eigenvalues in a single sample represented by x, then the j-th eigenvalue for each sample is denoted as , x_j, 0 < = j < = k. Suppose we have n data, then x^i $o_{i=1..n}$ ur data set is denoted , $D = \{x^i\}_{i=1..n}$. Our objective function is:

$$\max_{w} \sum_{i=1}^{n} I\left(y^i,\, t^i\right) \tag{1}$$

$y = f\,(x)$ *is* the predicted value of the model, t is the true value, and the function return $I(\,)s$ 1 if equals, y^iotherwise t^i it is 0.

We use neural network to learn the mapping relationship between input space and output space, and introduce gating mechanism to help control the inflow of information.

3.2. Neural Network Model Structure

Because the forecast data of agricultural products can only come from annual statistics, the amount of data is greatly limited, and various data augmentation technologies are difficult to use. Therefore, the model structure should be flatter and avoid over-fitting, so we only use one hidden layer.

GELU function and its Derivative showed in Fig. 1. For the activation function of hidden layer, GELU is a better function to introduce nonlinear relationship. Compared with the traditional RELU, GELU will change it to 0 according to Bernoulli distribution based on its numerical value. If a value is smaller, it has a greater chance of being converted to 0, and vice versa. Therefore, GELU can retain more useful information than RELU.

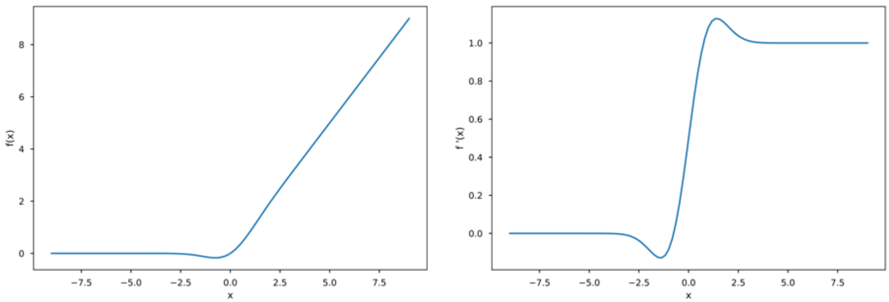

Figure 1. GELU function and it's derivative

3.3. Gating Mechanism

For the input of forecasting the total output value of agricultural products, some features have some similar characteristics. For example, per capita GDP and added value of primary industry can be regarded as data to measure the level of economic development. When we divide the input features into multiple feature groups according to their physical meanings, the importance of the features in each group relative to the information that the feature group wants to represent will be different from each other. The gating mechanism can adaptively control the information amount of each feature group involved in neural network calculation.

There are two main uses of gating mechanism, one is to act on the feature embedding layer, and the other is to act on the hidden layer. Because there is only one hidden layer in the model and the amount of training data is limited, if the gate control is applied to the hidden layer, it will be counterproductive, which will be mentioned in the introduction of the experiment. Therefore, the gating mechanism is applied to every feature group in the input layer, and the bit-wise mode is adopted. Compared with vector-wise, bit-wise can control the amount of information on each input value, which means that the processing accuracy of information will be higher than vector-wise. The excellent performance of bit-wise is also verified in the follow-up experiments.

For each input sample x, suppose there are m feature groups, where $m <= k$. Then the input after gate control is:

$$g_i = \sigma(W_i \cdot x_i) \quad for\ i\ in\ [0, m] \tag{2}$$

$$x_{gated} = concat(g_i \odot x_i) \quad for\ i\ in\ [0, m] \tag{3}$$

Where $\sigma(\,)$ is signmoid function to control the inflow size of information, and W_i is used as a learnable parameter. The model structure is showed as Fig. 2.

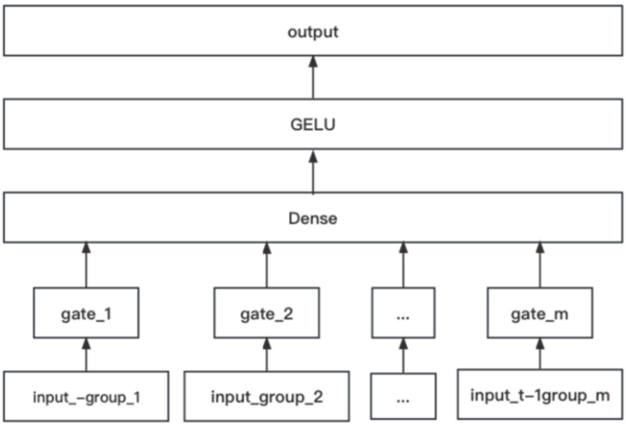

Figure 2. Improved BP neural network model structure based on gating mechanism

Generally, the neural network has a strong ability to fit various mapping relationships. The gating mechanism makes the neural network pay more attention to the important features in the input, while GELU can bring better nonlinear ability to the neural network and improve the model capacity.

4. Establish a Prediction Model

4.1. Data Sources

Based on the existing research results in the field of cold chain logistics demand and consumption of fresh agricultural products, and following the principles of availability, practicality and comprehensiveness, this paper selects five first-level indicators: economic development level, supply and demand level, transportation development level, cold chain logistics support level and cold chain logistics sustainable development level, and systematically summarizes and screens out 21 second-level indicators, as shown in Table 1.

Relevant data come from the statistical data of China Statistical Yearbook, China Cold Chain Logistics Development Report, China Logistics Yearbook and related websites from 2001 to 2022, which are directly quoted or indirectly calculated and collated.

Table 1. Index system of influencing factors of logistics demand

First-class index	Secondary index	Variable
Level of economic development	Per capita GDP	X1
	Added value of tertiary industry	X2
	Total retail sales of social consumer goods	X3
	Added value of primary industry	X4
	Residents' consumption expenditure	X5
Supply and demand level	Resident population	X6
	Fruit yield	X7
	Meat production	X8
	Output of aquatic products	X9
	Commercial price quantity of agricultural products	X10
	Freight volume	X11
Development level of transportation	Goods turnover	X12
	Added value of transportation, warehousing and postal services	X13
	Railway operating mileage	X14
	Total mileage of highway	X15
Support level	Number of outlets	X16
	Road mileage	X17
	Quantity of express delivery	X18
	Warehouse holdings	X19
Sustainable Development Level of Cold Chain Logistics	Number of college graduates	X20
	Fixed investment in scientific and technological research	X21

Therefore, we selected 22 pieces of data from 2001 to 2022 from the data source. The input of each data consists of 5 characteristic groups and 21 characteristics, including economic development level (5 characteristics), supply and demand level (6 characteristics), transportation development level (4 characteristics), cold chain logistics support level (4 characteristics) and cold chain logistics sustainable development level (2 characteristics). The label is the total output value of agricultural products, which is a continuous value.

Among the 22 pieces of data, we selected 18 pieces of data from 2001 to 2018 as training sets and 4 pieces of data from 2019 to 2022 as test sets.

4.2. Evaluation Criteria

We use the average error method to evaluate the performance of the model. The formula is as follows:

$$similarity = \frac{1}{n}\sum_{i=1}^{n} 1 - \frac{|y^i - t^i|}{t^i} \tag{4}$$

where similarity has a value range of $(-\infty, 1]$. When the similarity is higher, we think the model performance is better, and vice versa, the worse the model performance.

4.3. Trainer

4.3.1 Loss Function

Because this task is to predict continuous values, we choose MSE as the loss function. The advantage of MSE is that it can fully learn the mapping relationship from feature space to output space when the amount of training set data is small and the noise is small. The formula is as follows:

$$Loss = \frac{1}{n}\sum_{i=1}^{n}\left(y^i - t^i\right)^2 \tag{5}$$

4.3.2 Optimizer

We chose Adam as the optimizer. Adam absorbs the advantages of both Momentum and RMSProp, and is more stable in the weight update direction and update step size, thus finding the global optimum. The formula is as follows:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{6}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{7}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{8}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{9}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \tag{10}$$

where g_t is the gradient of the parameter, β_1 and β_2 are the attenuation coefficients of the two exponentially weighted averages, $\hat{m}_t$ and $\hat{v}t$ are the moving averages of the gradient after deviation correction, $\theta t + 1$ is the updated parameter, η is the learning rate, and ϵ is a small constant to avoid dividing by 0.

Among them, $[\beta_1, \beta_2]$ we choose $[0.98, 0.98]$ and the initial learning rate is set to 5e-6. From Table 2, we tested the epoch number of 50, 100, 200, 500, 1000, and found that when the epoch is 100, the model performs best in the test set. The training convergence state curve are showed as Fig. 3.

Table 2. similarity of different epoch

epoch	Similarity
50	92.77%
75	95.03%
100	96.07%
125	95.12%
150	94.36%

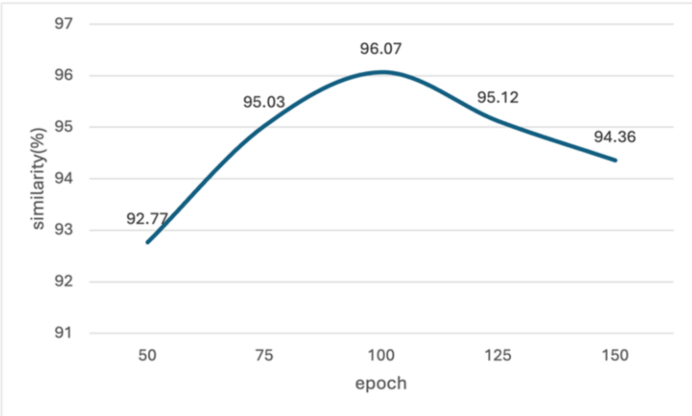

Figure 3. Training convergence state curve for BP neural networks

4.4. Comparison of Model Structures

We mainly compare three ways: no gating mechanism, gating mechanism in hidden layer and gating mechanism in input layer. Among them, the method of using gating mechanism in hidden layer is as follows:

$$g = \sigma\left(W_{gate} \cdot h\right) \tag{11}$$

$$h_{gated} = g \odot h \tag{12}$$

Where h represents the output of GELU, $\sigma(\ )$ is the signmoid function to control the inflow of information, and W_{gate} is a learnable parameter. Intuitively speaking, we use gating mechanism on GELU output to control how much information will be used to predict the total output value of agricultural products.

In the gating mechanism, for the learnable parameter W_i or W_{gate}, it can either map x_i or h to 1 dimension, which we call vector-wise, or it can map to the same dimension as x_i or h, which we call bit-wise.

Therefore, we can get the similarity using different model structures, showed as Table. 3:

Table3. similarity using different model structures

Model structure	similarity
No Gating mechanism is used	94.96%
Gating mechanism (vector-wise) in the hidden layer	95.04%
Gating mechanism (bit-wise) in the hidden layer	94.63%
Gating mechanism in the input level (vector-wise)	95.24%
Gating mechanism in the input level (bit-wise)	96.07%
Hidden layer (vector-wise) and input layer (vector-wise)	95.21%
Hidden layer (vector-wise) and input layer (bit-wise)	95.83%
Hidden layer (bit-wise) and input layer (vector-wise)	94.42%
Hidden layer (bit-wise) and input layer (bit-wise)	95.22%

4.5. Forecast Results

When we make predictions, we train multiple auxiliary neural networks separately. The input of each auxiliary neural network is the feature value of the previous three years, and the output is the feature value of the next year. When predicting the output value of agricultural products in the next year, first use the auxiliary neural network to obtain all input values, and then use the main network to predict the output value of agricultural products. The flowchart of the model is shown in Fig. 4.

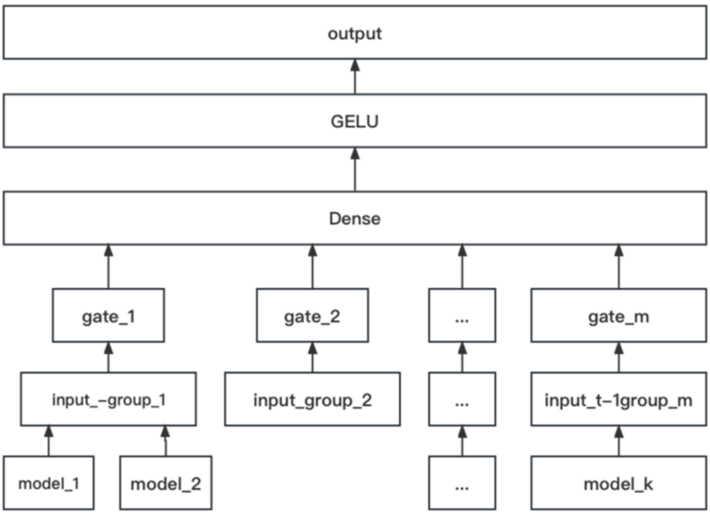

Figure 4. Improved BP neural network flowchart

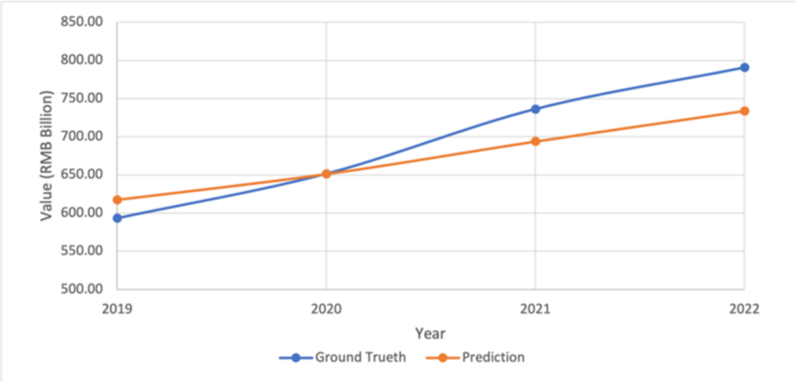

Figure 5. The contrast of ground truth and predicted value.

Our final prediction model of output value of agricultural products is a neural network model with bit-wise control mechanism in feature layer, a hidden layer with 42 neurons, activation function GELU, loss function MSE, optimizer Adam and 100 epoch training. The contrast of ground truth and predicted value is shown in Fig. 5.

Our auxiliary model consists of six neurons in multiple hidden layers, GELU in activation function, MSE in loss function and Adam in optimizer. Each model trains 25 epoch neural networks. Therefore, our forecast of the output value of agricultural products from 2023 to 2026 is shown in Table.4.

Table 4. Forecast of cold chain logistics demand in Hubei Province in the next 4 years

Year	Predicted value (RMB Billion)
2023	861.27
2024	920.91
2025	980.16
2026	1035.84

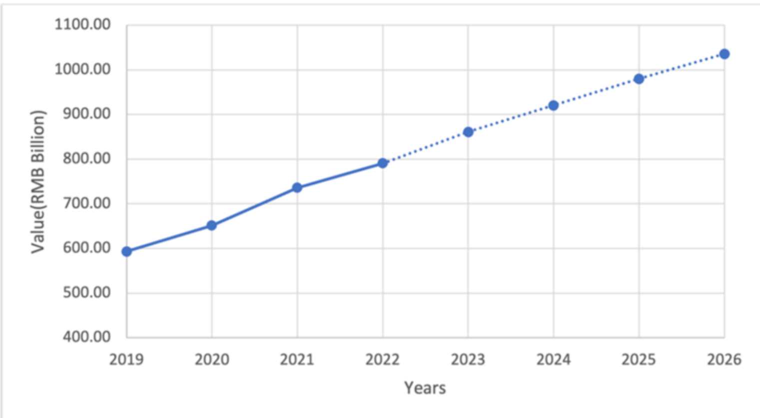

Figure 6. Hubei Province agricultural products cold chain logistics demand forecast curve.

5. Conclusion

Based on the above analysis, aiming at the characteristics and limited data volume of continuous value prediction in the relevant factors of agricultural product logistics demand forecasting, a gating mechanism is introduced to adaptively extract the information of each feature group. At the same time, in view of the slow convergence speed of common BP neural networks and the unstable defects in the prediction results, GELU is introduced to improve the training time of the model, improve the stability of the model, and build an improved BP neural network model for prediction. Experiments show that the improved BP neural network model achieves 96.07% similarity on the test dataset, with good accuracy and performance. From the forecast curve shown in Figure 6, the demand for cold chain logistics in Hubei Province will continue to grow in the future, and the cold chain logistics industry will continue to develop, but the development speed will slow down slightly.

5.1. Suggestion

The research results have certain reference value and can provide a certain reference basis for relevant departments in Hubei Province to formulate policies. In order to further promote the development of regional cold chain logistics industry in Hubei Province, in view of the influencing factors of cold chain logistics demand, combined with the development of agriculture and logistics industry in Hubei Province, the following suggestions are put forward: 1. Strengthen the cold chain logistics service system and build a cold chain logistics system that meets the characteristics of the local industrial structure and meets the needs of economic and social development. 2. Improve the low-temperature processing capacity of agricultural product production areas, expand the supply of high-quality fresh agricultural products, and support the development of cold chain logistics industry. 3. Improve the research and development of key technologies and advanced equipment in cold chain logistics, promote the establishment of a statistical evaluation system for cold chain logistics, and improve the standard system of cold chain logistics. 4. Increase investment in scientific research and education in cold chain logistics related fields, and cultivate more cold chain logistics professionals.

5.2. Deficiencies and Prospects

This paper presents a novel improvement to the backpropagation neural network model based on gate mechanism and GELU activation function. It demonstrates a method for optimizing the neural network connection weights using the gate mechanism. The effectiveness of the algorithm is validated through experiments, which is determined by the high robustness and efficiency of GELU and Adam optimization algorithm. However, there are certain limitations in this study. When performing predictions, the inputs are based on the predictions of multiple sub-models. As these sub-models are trained separately and may have errors compared to the true input values, these errors accumulate further when the main model makes predictions. To overcome this limitation, future improvements could involve adopting cascade training, where the sub-models and the main model are trained together to reduce the accumulation of errors.

Acknowledgment

This project is supported by Major Project of Philosophy and Social Science Research in Hubei Province Higher Education Institutions (22ZD046); The Ph.D. Scientific Research and Innovation Foundation of Sanya Yazhou Bay Science and Technology City (HSPHDSRF-2022-03-032); The WUST & Int-plog Research Project (2022H20537); The CSL & CFLP research project plan (2023CSLKT3-306); The Department of Education of Hubei Province Young and Middle-Aged Talents Project (20211102).

References

[1] Awad M, Zaid-Alkelani M. Prediction of water demand using artificial neural networks models and statistical model[J]. International Journal of Intelligent Systems and Applications, 2019, 11(9): 40.

[2] Mamdouh M, Ezzat M, Hefny H. Optimized Planning of Resources Demand Curve in Ground Handling based on Machine Learning Prediction[J]. International Journal of Intelligent Systems and Applications, 2021, 13(1).

[3] Howlader S, Biswas T, Roy A, Mortuja G, Nandi D. A Comparative Analysis of Algorithms for Heart Disease Prediction Using Data Mining[J]. International Journal of Intelligent Systems and Applications, 2023, 15(5).

[4] Heikkurinen P, Forsman - Hugg S. Strategic corporate responsibility in the food chain[J]. Corporate Social Responsibility and Environmental Management, 2011, 18(5): 306-316.

[5] Benkachcha S, Benhra J, El Hassani H. Demand forecasting in supply chain: Comparing multiple linear regression and artificial neural networks approaches[J]. International Review on Modelling and Simulations, 2014, 7(2): 279-286.

[6] Eksoz C, Mansouri S A, Bourlakis M, et al. Judgmental adjustments through supply integration for strategic partnerships in food chains[J]. Omega, 2019, 87: 20-33.

[7] Huang L, Xie G, Zhao W, et al. Regional logistics demand forecasting: a BP neural network approach[J]. Complex & Intelligent Systems, 2021: 1-16.

[8] Munkhdalai L, Park K H, Batbaatar E, et al. Deep learning-based demand forecasting for Korean postal delivery service[J]. IEEE Access, 2020, 8: 188135-188145.

[9] Ma H, Luo X. Logistics demand forecasting model based on improved neural network algorithm[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(4): 6385-6395.

[10] Yu N, Xu W, Yu K L. 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, 2020, 8: 9551-9564.

[11] Zhang G, Xu X. A logistics demand forecasting model based on ARIMA-BPNN [J]. Control Engineering, 2017, 24(05): 958-962(in Chinese).

[12] Nguyen T Y. Research on logistics demand forecast in southeast Asia[J]. World Journal of Engineering and Technology, 2020, 8(03): 249.

[13] Du B, Chen A. Research on logistics demand forecast based on the combination of grey GM (1, 1) and BP neural network[C]//Journal of Physics: Conference Series. IOP Publishing, 2019, 1288(1): 012055.

[14] Feng Y. Research on demand forecast of agricultural products cold chain logistics in Beijing [D]. Beijing. North China Electric Power University (Beijing), 2016(in Chinese).

[15] Wang S. Research on cold chain logistics demand forecast of fresh agricultural products [D]. Xi'an: Xi'an Engineering University, 2017(in Chinese).

[16] Wang X, Yan F. Demand forecast of cold chain logistics of urban agricultural products in Beijing based on GA-BP model [J]. Practice and Understanding of Mathematics, 2019, 49 (21): 17-27(in Chinese).

[17] Wu P, Lei Z, Zhou Q, et al. Multiple premises entailment recognition based on attention and gate mechanism[J]. Expert Systems with Applications, 2020, 147: 113214.