

Analysis of E-Commerce Product Sales Prediction Based on Ga-Bp Neural Network

Jianwei Zhai*, Qing Li, Yufei Xia

Department of Software Engineering, HeiLongJiang University Harbin, China, 150080

* Corresponding author: fgowuming@gmail.com

Abstract. In recent years, the market for the sale of various goods has been expanding and with it comes huge amounts of data that needs to be specifically processed and analyzed. For a company, the adoption of artificial intelligence algorithms with better performance in handling big data is an important step in influencing the future decisions and future growth of the company. In this paper, an optimized neural network based on GA-BP is constructed to analyze samples with different characteristics to predict the future sales in the short term. After testing, the prediction results are found to be good. First, the degree of influence between the indicators is examined using BP neural network. By comparing whether the inventory indicator at the beginning of the month is used as the model input, the error of the end-of-month inventory forecast content is obtained, which leads to the conclusion that the inventory at the beginning of the month has a greater impact on the inventory at the end of the month. Similarly, it can be concluded that the commodity type has a greater impact on the commodity pool. Second, to further analyze the relationship between commodity type and commodity inventory, the Eta test was used. It was found that item type A2 had the greatest impact on merchandise inventory, followed by type A1 and finally type A3. This finding is consistent with the characteristics of the dataset. Considering that the optimized GA-BP neural network model is mainly used to predict merchandise sales in the next 3 months; secondly, considering the specificity of some data in the sample, it is not suitable to use the neural network for analysis and prediction. For all data, the data that could be analyzed by GA-BP were found by gradient descent method, and it was found that the data with target value below 5500 had the best prediction effect. Finally, according to the given accuracy evaluation index, the score of the model is calculated as 0.678 5, and the model works well.

Keywords: GA-BP optimization neural network, accuracy evaluation index, electronic goods.

1. Introduction

The development of e-commerce is the product of the Internet social era. The development of e-commerce has become the booster of the new wave of society and is a great progress in the development of the Internet. At present, the major e-commerce brands have a wide range of goods, and the sales volume of goods has been increasing. It is not difficult to find that the historical sales data of goods for enterprises to plan their marketing programs, for a more in-depth analysis of the market, for the rationality of logistics planning. It is of great significance. Therefore, the use of excellent statistical and predictive modeling algorithms is particularly important for e-commerce companies to maintain product inventories, reduce inventory costs, shorten delivery cycles, improve resilience to various risks, and ultimately increase profits [1].

2. The fundamental of data analysis

In this paper, we analyze the data characteristics of the test set for the sample data to obtain the relationship between the inventory at the beginning of the month and the inventory at the end of the month. Since it is difficult to directly analyze the amount of data in the test set in months as a time unit, a BP neural network is used to fit the inventory at the end of the month. By keeping the inventory at the beginning of the month as the input to the neural network and eliminating its effect, the prediction effects of the two month-end inventories are obtained separately. The error of the month-end inventory obtained by using the inventory at the beginning of the month as the input to the neural

network is shown in Figure 1 below. By excluding the inventory at the beginning of the month as the input to the neural network, the error in the month-end inventory obtained is shown in Figure 2 below. It is concluded that the inventory at the beginning of the month has a significant effect on the inventory at the end of the month, and the prediction effect based on BP neural networks can vary significantly depending on whether the inventory at the beginning of the month is used as input. BP neural networks are also used to detect whether there is a significant relationship between commodity type and commodity inventory [2-3]. Second, the relationship between item type and item inventory is further quantified by the Eta test. The commodity type was used as the input to the neural network as shown in Figure 3 below.

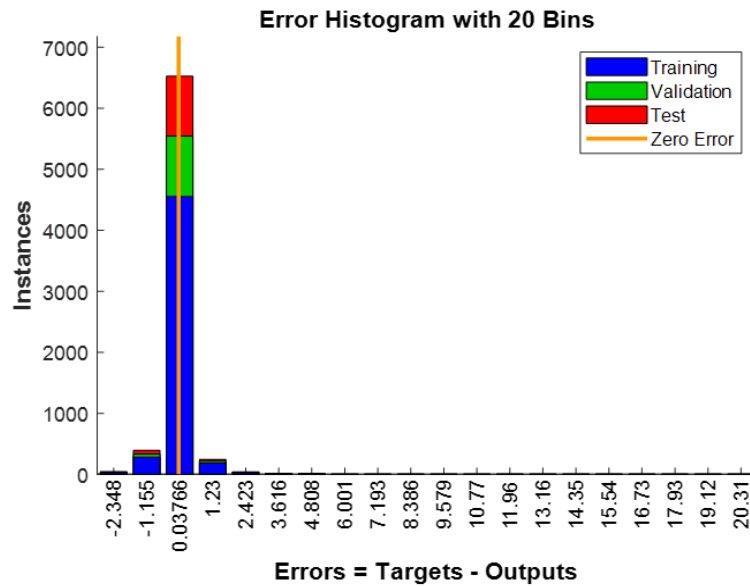


Figure 1. Forecast month-end inventory considering the beginning of the month inventory as input

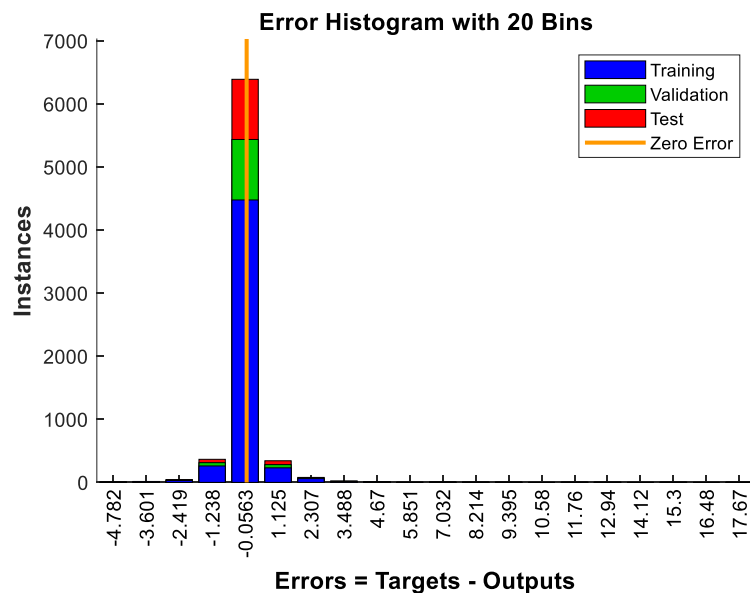


Figure 2. Inputs without considering the inventory at the beginning of the month as an effect of the forecasted month-end inventory

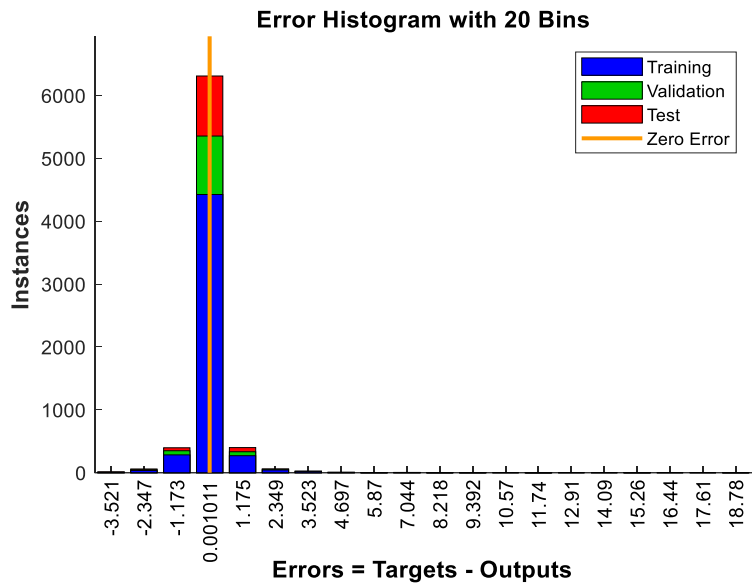


Figure 3. Effect of considering commodity type as an input to forecast month-end inventory

After extensive data analysis and processing, it can be concluded that item type has a strong influence on merchandise inventory, and the prediction effect based on BP neural network can vary greatly depending on whether the item type is used as input [4].

To quantify the relationship between item type and merchandise inventory, the effect of time is first removed from the data in the table. Here, the commodity inventories for each data set are summed by month.

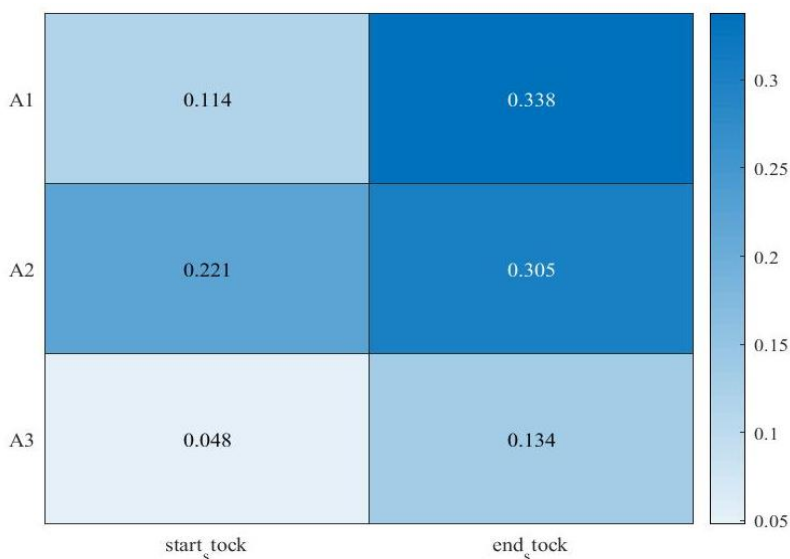


Figure 4. Confusion Matrix

An Eta test was used to analyze the commodity inventory of various commodity types and all training data sets as inputs to obtain a quantitative relationship between the two. The data with item IDs 1144 and 1145 were excluded due to the missing amount of data for item type IDs 1144 and 1145, which avoided outliers and had less impact on the evaluation results.

Each item type and item inventory that passed the Eta test is shown in Figure 4 below. Based on the results, the degree of influence of various commodity types on commodity inventory was obtained [5].

It is concluded that there is a significant relationship between item type and item inventory. The three merchandise types are ranked in order of their influence on merchandise inventory. a2, a1, a3. This result coincides with the characteristics of the dataset: by analyzing the dataset, most of the data with more monthly sales belong to the A2 category; the A1 and A3 categories tend to be difficult to

sell or even have no sales for three years. Quantity. It can be assumed that the type of goods further affects the inventory of goods through the sales volume of goods, which is in line with the perception.

3. Results

3.1. The establishment of ga-bp model

By removing the outlier data, the remaining data are used for the neural network model prediction task. Because ordinary BP neural networks fall into local optima during gradient descent, resulting in poor model training. In order to get better prediction results, this problem completes the downstream prediction task based on the optimized GA-BP neural network and uses the average relative error to test the model prediction results [6-7]. The model structure of the GA-BP neural network is shown in Figure 5 below.

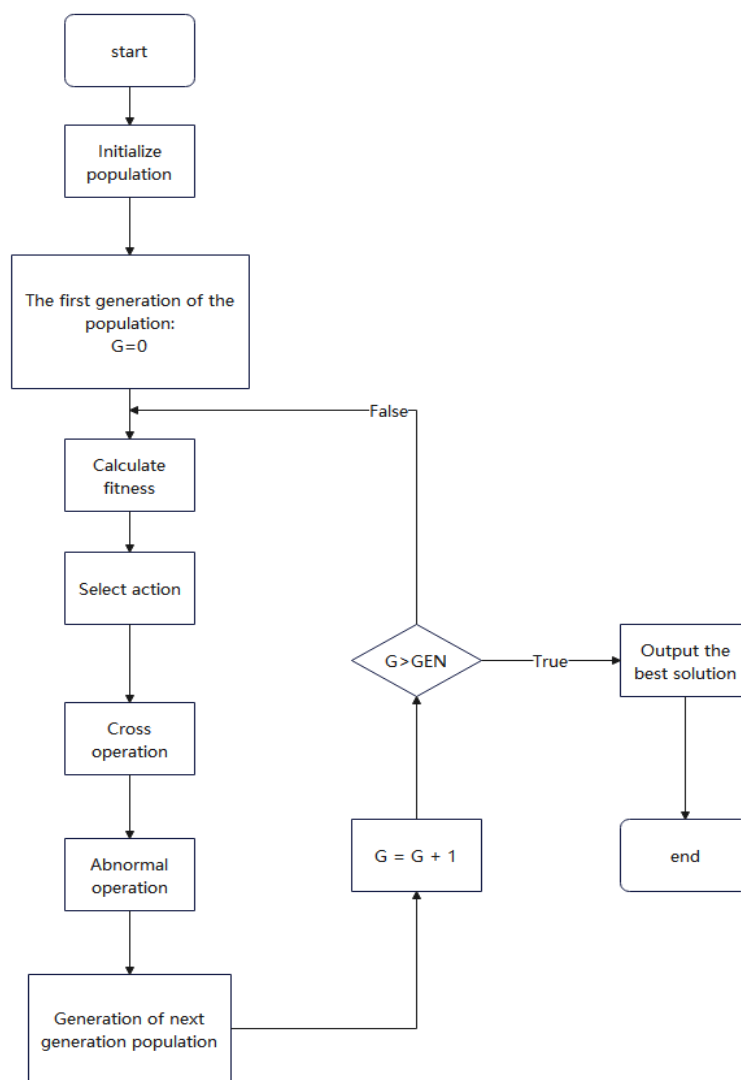


Figure 5. GABP neural network implementation process

Since the data range of the target forecast sales data is too large and the variance is too large, the effectiveness of the model will be greatly reduced. The calculation results are shown in Figure 6. In order to reduce the variance of the sales data and improve the accuracy of the model prediction, the sales data are logarithmically processed here.

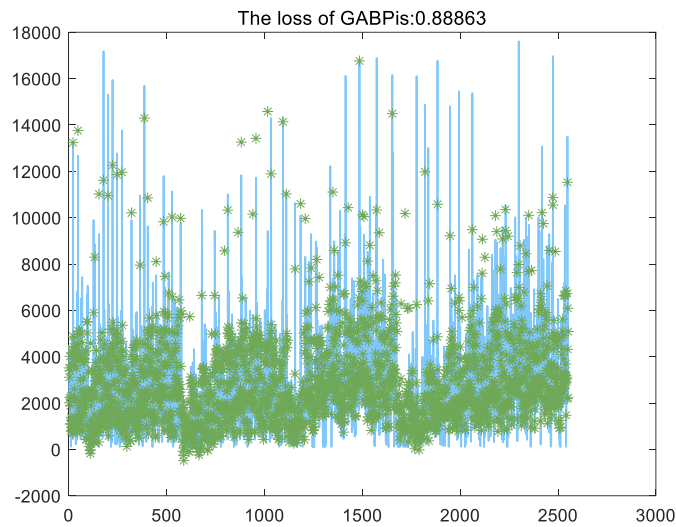


Figure 6. Prediction performance without data manipulation

After normalizing the input data, the prediction effect using GA-BP neural network is shown in Figure 7 below.

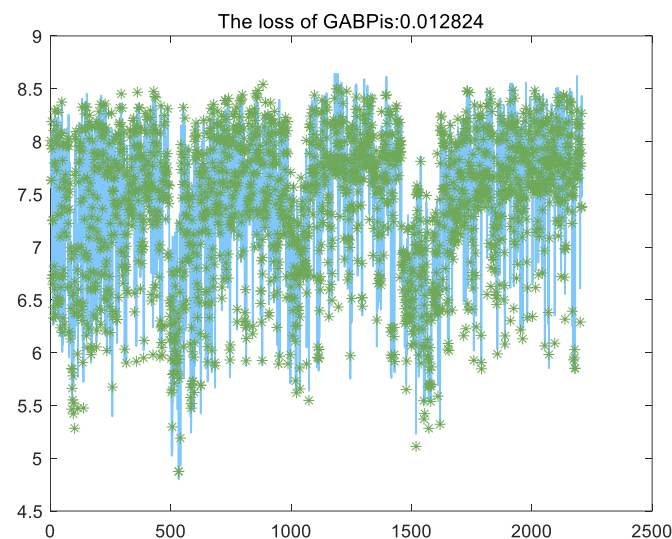


Figure 7. Predictive performance with data manipulation

3.2. Analysis of experimental results

Predict through the above steps to obtain the predicted sales value of all training sets, and calculate with the known accurate value to obtain an Accuracy [8]. The specific calculation formula is as follows:

$$Accuracy_1 = \sum_{i=1}^n \left[\left(1 - \frac{|predict_i - t \arg et_i|}{t \arg et_i} \right) * \frac{t \arg et_i}{\sum_{j=1}^n t \arg et_j} \right] \quad (1)$$

For the training set, an accuracy evaluation result of 0.6785 was obtained. Finally, the error test was performed. The training data set is again divided into the test set and the training set, and the data from the training set is used to predict the data from the test set to check the prediction effect of the model. The error test of the GA-BP neural network prediction model is shown in Figure 8 below. The

test results are good, with slightly lower accuracy than the experimental results [9-10]. This may be due to the large overall variance of the test set data, the lack of distinctive features, and the absence of special data series for separate analysis. Noise treatment is added to the target values of the model. Here, a noise value of 5% is added and the data is trained to observe the accuracy of the model prediction. Since gray prediction is for small sample prediction and is suitable for short-term prediction, the method of adding noise treatment is not applicable, and here only noise treatment is applied to the GA-BP neural network model. the stabilization effect of GA-BP neural network is shown in Figure 9 below.

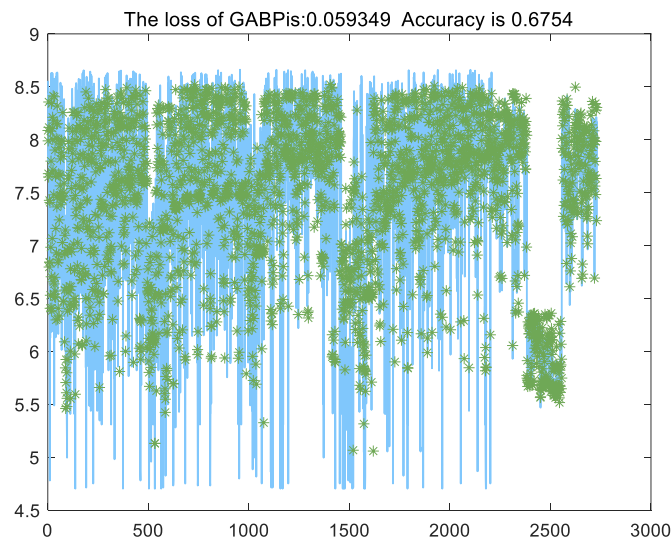


Figure 8. Error analysis of neural network

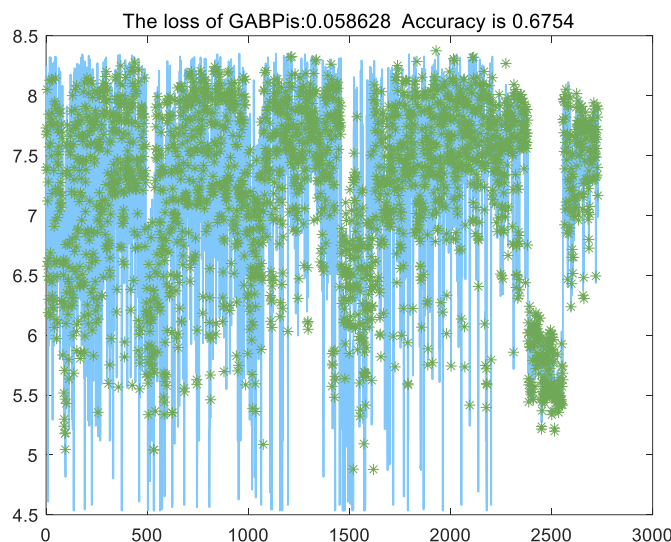


Figure 9. GA-BP neural network stability test

4. Conclusions

In this paper, we conducted a study related to the inventory at the beginning and the end of the month and found that the inventory at the beginning of the month has a greater impact on the inventory at the end of the month. By studying the impact of commodity type on commodity inventory, it is found that different commodity types have different degrees of impact on the monthly inventory of commodities. Among them, commodity type A2 has the greatest impact on commodity inventory,

commodity type A1 has a medium impact, and commodity type A3 has the least impact on commodity inventory. Based on the historical sales records of commodities in the past, the sales of various commodities in the next three months were predicted, and the prediction model index score was 0.6789, indicating that the historical data has some reference significance for future commodity sales prediction. Considering the comprehensive interests of the company, by mining other links between the data, it is concluded that there are large differences in merchandise sales in different time periods and the time of sales has different effects on merchandise sales.

References

- [1] Liao C.H., Chen C.Q., Wang C.L. Analysis of yield prediction of Guangdong Huai Shan based on GA-BP neural network [J]. *Agricultural Mechanization Research*, 2023, 45(08): 183-187.
- [2] Lv Wenyuan. GA-BP joint fuzzy PID for cutter arm swing speed control of roadheader [J]. *Coal Technology*, 2022, 41(10): 230-233.
- [3] Yuan Feng, Liang Yujia. Design of a particle swarm algorithm-based BP neural network electricity price prediction system [J]. *Integrated Circuit Applications*, 2022, 39(04): 132-133.
- [4] Wu Jiawen. Design and implementation of wind power prediction system based on BP neural network for wind farms [D]. Nanchang University of Aeronautics, 2019.
- [5] Chen C.W. Design and implementation of PM2.5 prediction system based on BP neural network [D]. South China University, 2019.
- [6] Zhai Xiaojuan. Design of BP neural network-based prediction of truck braking system performance parameters [D]. Shanghai University of Applied Science and Technology, 2018.
- [7] Luo Jinjin, Li Xiancan, Wu Mingqing, Li Fayong. Design and experiment of gin speed control system based on BP neural network prediction model [J]. *Modern Electronics Technology*, 2017, 40(09): 141-144.
- [8] Liu Shun. Research on the development of a system for predicting and optimizing the vertical bearing capacity of foundation piles based on genetic-BP neural network [D]. Shandong University of Construction, 2010.
- [9] Li Quanxi, Ma Xiaomiao. E-commerce model and its development trend [J]. *Intelligence Science*, 2005, (08): 1138-1142.
- [10] Zhao Hui, Shao Suhua, Xie Dongpo. The treatment of outliers in analytical data [J]. *Journal of Zhoukou Normal College*, 2004, (05): 70-71+115.