Journal of Web Systems and Applications (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/jwsa.2022.040101
ISSN 2560-6913 Vol. 4 Num. 1

Sales Forecasting of Auto Retail Parts Based on BP Neural Network Analysis Model

Hao Zhang[1], Zhehua Zhang[*1], Jun Jiang[1], Shuting Huang[2], Xinyan Wang[1]

[1]Tibet University, Lhasa 850000, Tibet, China

[2]Hainan University, Haikou 570000, Hainan, China

**Corresponding Author*

Keywords: BP neural network, Sales forecasting, Spss

Abstract: The economic system reform, automobile manufacturing industry has reached the peak of production, accompanied by the ensuing demand and supply of auto retail parts, how to achieve the purpose of forecasting the precise demand for auto retail parts, is to solve the current small class level of most auto retail parts enterprises and even store auto retail parts level to provide basic demand forecasting, and for enterprises and even manufacturers inventory management has a more It also has a faster categorization management method for enterprises and even manufacturers' inventory management.

1. Introduction

With China's booming economy and rising incomes, the consumption model in the market is gradually changing from the traditional “object-oriented” retail model to the new “customer-oriented” retail model. New retailing is a new retailing model in which enterprises use the Internet as the backbone and apply advanced science and technology such as big data and artificial intelligence to upgrade and transform the production, circulation and sales processes of commodities, thereby reshaping the business structure and ecosystem and deeply integrating online services, offline experiences and modern logistics. Under this influence, people's consumption is developing in the direction of convenience, personalization, humanization, experience and diversification. These special consumption trends promote the production mode of the new retail industry to gradually move towards a multi-species, small batch mode, which also adds considerable difficulty to the inventory management of the retail industry, and the amount of inventory directly affects the level of sales of enterprises.

The sales, total sales and inventory of auto retail accessories for the four holidays of National Day, Double 11, Double 12 and New Year's Day 2021 are investigated through survey data thus obtaining the sales and total sales weights and the weights of sales and inventory, and the auto retail accessories division of the top 50 auto retail accessories in terms of cumulative sales within July 1, 2021 to October 1, 2021 Inventory thus obtaining sales and total sales weights and weights of sales and inventory for comparison and comparing for correlation analysis of the impact of sales characteristics, inventory information, holiday discounts, and other factors on sales volume.

2. Model Building

2.1 BP Neural Network Analysis Model

Based on a comprehensive analysis of the relevance of each relevant factor on the sales volume of retail auto parts, an empirical study on the sales volume of the target category can be conducted considering inventory information, product sales characteristics, holiday discounts, and consumer preferences, among other factors. According to a survey conducted by Baidu on Internet users, consumers often use search engines to obtain product information before purchasing the target category, focusing on brand, manufacturer, product price and performance information, and the most commonly searched for influencing factors are sales characteristics. Therefore, the purchase desire of such consumers is stronger than those who search for inventory information or holiday discounts as an influencing factor.

This paper uses the Cross Correlation function of SPSS software to conduct correlation test analysis between the influencing factors Baidu index and the monthly sales data of the target subcategory.

Firstly, the influencing factors with absolute value of correlation coefficient greater than or equal to 0.45 and the leading ratio relative to the monthly sales data were retained, and there were 14 influencing factors that passed this screening process. In order to improve the quality of the research data and simplify the complexity of the model, the keywords with correlation coefficients less than 0.5 were eliminated and finally 13 influencing factors were retained for later model training, as shown in Table 1.

Table 1 Correlation Coefficients And Ratios of Influencing Factors

Influencing Factors	Correlation coefficient	Percentage
Sales Characteristics	0.653	0.12
Inventory Information	0.513	0.14
Holiday Discount	0.623	0.15
Nature of the product	0.502	0.05
Fashionability of the product	0.621	0.13
The size of the product's value	0.514	0.06
Word of mouth of the product	0.601	0.05
The size of the target market range	0.507	0.04
Potential demand of the target market	0.503	0.03
The degree of concentration and fragmentation of the market	0.514	0.03
Buying characteristics of customers	0.625	0.08
Market competition situation	0.518	0.05
Consumer preferences	0.604	0.07

2.2 Model Structure and Analysis of Prediction Results

The number of layers of BP neural network is determined as 3, in which the number of input layer nodes is 13 and the number of output layer nodes is 1. According to the empirical formula f=1.5 mn (f is the number of neurons in the hidden layer, n, m are the number of input neurons and the number of output neurons respectively), the number of nodes in the hidden layer is initially determined as 5. Subsequently, the network is trained with different numbers of hidden layer nodes using the trial-and-error method, and it is found that the network is optimal when the number of nodes is 10. The node transfer function uses tansi tangent S-type transfer function, losi logarithmic S-type transfer function and linear transfer function purelin; the training function uses momentum backpropagation with dynamic adaptive learning rate of gradient descent function BP algorithm

trainlm [1], and the rest uses the system default. The neural network structure is shown in Figure 1 .

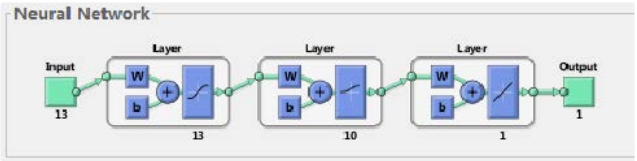
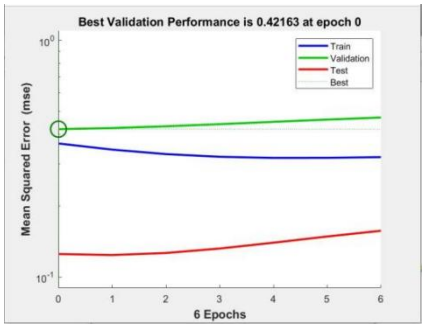


Fig.1 Neural Network Structure

According to the three influencing factors identified in the previous section for analysis, the monthly search data of each influencing factor were aligned with the monthly sales data staggered according to the time difference relationship, and the sales data of the target subcategory whose historical sales time was within June 1, 2021 to October 1, 2021 and whose cumulative sales ranked in the top 10 [3] and its corresponding sales ranked in the top 10 of the target The monthly sales data of the 13 influencing factors in the subcategory were output as training data to the Matlab program. The training results are shown in Fig. 3, where the curves represent the actual sales for each year and the dots represent the fitted values of the neural network, and the network training error is very small.



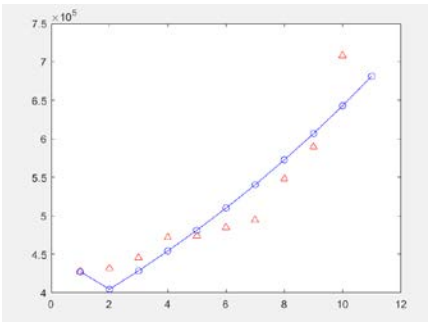


Fig.2 Training Process Error

Fig.3 Comparison of Sample and Network Output

The trained neural network was used to predict sales for three months from October to December 2021, and the network search data for three consecutive months corresponding to the 13 influencing factors and target subcategories were input, and the prediction results are shown in Table 2.The error analysis results show that the absolute average error of the model used to predict the sales data for three months is 5.2% ,with Absolute Percent Error (APE) Percent Error, APE) and Mean Absolute Percent Error (MAPE) is the average of the absolute error, the mean absolute error can better reflect the actual situation of the prediction value error, with MAPE indicates the prediction accuracy of the model, the smaller the error value, the higher the prediction accuracy of the model [2], defined respectively as follows:

MAPE calculation equation.

$$MAPE = \sum_{i=1}^{n} \frac{\left| y_i - \hat{y}_i \right|}{n * y_i} = \sum_{i=1}^{n} \frac{1}{n} * \frac{\left| y_i - \hat{y} \right|}{y_i} \quad (1)$$

$$APE = \frac{\left| y_i - \hat{y}_i \right|}{y_i} \quad (2)$$

Table 2 Prediction Results And Error Analysis

		Months			
	Oct.2021	Nov.2021	Dec.2021	MAPE	

Actual Sales	5493612.5	4915310	5598175		
Forecast Sales	5806748.41	5151244.88	5883681.92		
APE	5.7%	4.8%	5.1%	5.2%	

3. Improvement of the Model

Since there are 13 selected influencing factors, there may be certain linear or nonlinear relationships among these influencing factors, which cause interference of repeated information for model training and prediction carrying out. Therefore, it is considered to synthesize the covariance indicators through principal component analysis, and transform multiple indicators into a small number of comprehensive indicators with higher information and stronger explanatory power, so as to achieve the purpose of simplifying the model and improving the prediction accuracy.

In the process of principal component analysis using SPSS software, the sample data need to be standardized first due to the large differences in magnitudes among the indicator data. Before the principal component analysis, KMO and Bartlett's sphericity test are conducted to quantitatively test whether the variables are correlated, which contains two indicators: the first is the KMO value, the closer the value is to 1, the more suitable for factor analysis, 0.756 is not enough to make a judgment; the second is the Bartlett's sphericity test, accompanied by the probability Sig less than 0.05. Both were evaluated comprehensively and were able to do principal component analysis.

Using SPSS statistical software to calculate the principal component analysis for each factor affecting car sales, it can be seen in Table 4 that after the original web search data are analyzed by principal component analysis, there are two principal components with characteristic roots greater than 1. According to the default selection of SPSS, the cumulative variance contribution of the first two principal components is over 73%, which can better explain the original data information, and the rest of the components contain less information, so they are discarded.

Table 3 Extraction of Principal Components of Variance Decomposition of Keyword Search Index

		Initial Eigenvalue			Extraction of squares and loading	
Ingredients	Total	Variance%	Cumulative%	Total	Variance%	Cumulative%
1	6.898	53.059	53.059	6.898	53.059	53.059
2	2.625	20.195	73.252	2.625	20.159	73.253

The matrix of score coefficients of the principal components in Table 4 enables to obtain the mathematical expressions of each principal component of the search index influencing factors respectively, and the expression of the first principal component Z is shown below:

$$Z_1 = 0.860X_1 - 0.958X_2 + 0.790X_3 + 0.834X_4 + 0.926X_5 - 0.881X_6 + 0.943X_7 - 0.172X_8 + 0.573X_9 - 0.431X_{10} + 0.598X_{11} + 0.616X_{12} + 0.343X_{13}$$

On the basis of the BP neural network prediction model proposed above, the principal component analysis is further performed on the keyword search indexes affecting car sales to eliminate the covariance between different search indexes. Each extracted principal component index is used as the new network input, and two principal components are extracted, so the number of nodes in the input layer is modified to 2,and the number of nodes in the hidden layer is modified to 2 according to the empirical formula and comparative analysis.The rest of the structure and key parameters are not adjusted.

The model is trained again by inputting new data, and the model structure with reduced data dimensionality is more simplified. The training efficiency is improved, and the data fitting effect is shown in Fig. 4.

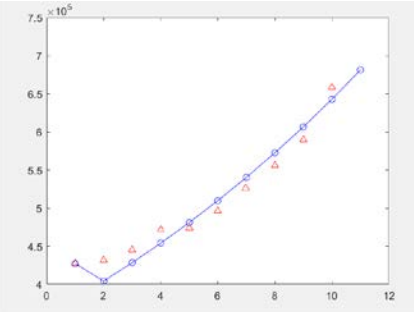


Fig.4 Improved Fitting Effect

The comparison results of the improved model are shown in Table 4. The MAPE index decreased by 0.5% after subjecting the keyword network search data to principal component analysis, and the prediction accuracy was improved.

Table 4 Principal Component Analysis BP Neural Network Prediction Results

		Months		
	Oct.2021	Nov.2021	Dec.2021	MAPE
Actual Sales	5493612.5	4915310	5598175	
Forecast Sales	5593613	5015310	5698175	
APE	0.021	0.018	0.031	2.3%

4. Conclusion

Using the Cross Correlation function of SPSS software, the correlation test was conducted between the image factor Baidu index and the monthly sales volume data of the target subcategory, 13 influencing factors were screened out for later model training, the number of BP neural network layers, input nodes and output nodes were determined, and the network was trained by taking different numbers of hidden layer nodes using the trial-and-error method. After reaching the set error range, the trained neural network is used to predict sales for three months from October to December 2021, and the network search data for three consecutive months corresponding to the 13 influencing factors and the target subcategories are input to obtain the prediction results, and the prediction is made for the specific target subcategories. Therefore the model has high reference value in practical application. It should be noted that the prediction timeliness of the model in this paper depends on the ratio of network influencing factors to sales, and since the minimum ratio chosen for the empirical evidence is 1%, the lead time that the model can predict is one month.

Acknowledgement

2019 Cultivation Fund of Tibet University (ZDTSJH19-02); 2021 Wuhan University of Technology-Tibet University “Tibetan Economic and Social Development and Highland Scientific Research Co-Building Innovation Fund” (lzt2021007).

References

[1] Ma Huan. Research on auto sales prediction model based on Baidu index and BP neural network [D]. Wuhan University of Technology, 2018. 16-17

[2] Cheng Li. Comparison of GDP forecasting models in Hunan Province [D]. Xiangtan University, 2021. DOI:10.27426/d.cnki.gxtdu.2021.001444.

[3] Zhang Shuo, Liu Kun, Li Xichang. Accurate demand forecasting of new retail products [J]. China collective

economy, 2021(19):55-56.
[4] Wei F. Y., Li Q. F., Yan Y. T., and Zhang J. X. Research on sales prediction of new retail target products based on genetic neural network[J]. China Business Journal, 2021(18):32-35.DOI:10.19699/j.cnki.issn2096-0298.2021.18.032.