ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202202_5(2).0064

Research on Precise Demand Forecast of New Retail Based on Principal Component Analysis Model

Haobo Zhang[1], Hanshuo Song[2], Lingyun Song[3]

[1]School of Artificial Intelligence. North China University of Technology, Tangshan, 063210, China

[2]School of Mechanical Engineering. North China University of Technology, Tangshan, 063210, China

[3]School of Management. North China University of Technology, Tangshan, 063210, China

Abstract

Analyze the influence of various relevant factors on SKC target sales during the National Day, November 11, December 12 and New Year's Day in 2018. Firstly, FineBi and Excel were used to screen SKC and related data. Then, principal component analysis (PCA) was used to analyze related factors, and SPSS correlation matrix, general eigenvalue table and gravel map were obtained to determine the weight of each related factor. By weight analysis, the influence of holidays on sales could be obtained. The maximum discount sales feature is smaller than the discount but larger than the inventory effect on sales.

Keywords

Accurate demand forecasting; Principal component analysis.

1. Introduction

In the emerging retail industry, how to make multi-level accurate demand forecast according to the historical sales data of complex levels and various categories is a problem that most new retail enterprises need to pay attention to and think about. Based on the relevant data, this paper selects and preprocesses the data content, and analyzes the problems existing in the accurate prediction of demand in order to establish and optimize the mathematical model, and gives the corresponding prediction scheme

2. Organization of the Text

2.1. New Retail Industry Precise Demand Forecasting Problem Proposed

2.1.1. The Problem Background

With the continuous development of China's consumer market, the consumption mode in the market has gradually changed from "material oriented" to "customer oriented". In the new retail industry, value for money is no longer the only criterion for consumers to judge whether they want to buy. People's demand is not only the simple pursuit of practicality, more is to consider fashion, pay attention to "personalized, fashion, beauty" and other aspects. Due to such special needs, the new production mode of retail enterprises, gradually moving towards more varieties and small batches, makes the types of jewelry and toys in retail stores and shopping centers more and more dazzling, at the same time in the retail industry inventory management, increased great difficulties. How to accurately forecast demand at regional level, small level and even store SKC (Single Color) level according to the historical sales data of complex levels and various categories is a prob that most new retail enterprises need to pay attention to and think about at present.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202202_5(2).0064

2.1.2. Ask Questions

Try to analyze the influence of relevant factors on the target SKC sales during the National Day, November 11, December 12 and New Year's Day in 2018. Consider product sales characteristics, inventory information, holiday discounts and other factors.

2.2. Analysis of Precise Demand Forecasting in New Retail Industry

2.2.1. Problem Analysis

First of all, it is required to analyze the sales volume of target SKC in National Day, November 11, December 12 and New Year's Day. There is a large amount of data in the attachment, so the data must be processed first to get the required valid data. Filter the data to obtain the target SKC, and then filter the effective data of the target SKC during the four holidays and relevant factors. Finally, factor analysis and principal component analysis were conducted to evaluate the influence of screened data on related factors.

2.2.2. Model Assumes

1 .Assume that all data obtained after data pretreatment are valid and representative.

2. Assume that the retail environment won't change much in the future.

3. Assume that the development of retail industry in East China can represent the development of the whole retail industry.

2.2.3. Solution

It is required to evaluate the influence of various related factors on sales volume of target SKC during the four holidays of National Day, Double 11, Double 12 and New Year's Day. Firstly, we screened the target SKC from a large amount of data in the attachment by FineBi and Excel, and then screened and calculated the specific data of sales characteristics, inventory, holiday discounts and other related factors to obtain the specific data of various factors related to the four holidays. After get valid data, we use the principal component analysis (pca), by the SPSS software, based on various related factors to make a series of charts, finally established the comprehensive evaluation model, identified the four holidays impact on sales, the biggest discount sales characteristics influence on sales is less than the discount but is greater than the inventory's impact on sales.

3. The Establishment and Solution of Principal Component Analysis Model

3.1. Data Preprocessing

FineBi and Excel were used to screen the target SKC from the attached data, and then the sales, inventory and discount of SKC products were screened (the discount rate was obtained by dividing the actual cost by the sales volume and the price tag), and various related factors were screened out, among which some missing values were interpolated.

3.2. Basic Principles of Model

Principal component analysis (pca) is a statistical method of dimension reduction, it is by using a orthogonal transformation, the original random vector that are relevant to the component into its component is not related to the new random vector, this appears to be the original random vector on the algebra of covariance matrix transformation into a diagonal matrix, on the geometry of the original coordinate transformation into a new orthogonal coordinate system, Then, the multidimensional variable system can be converted into a low-dimensional variable system with a high precision by dimensionality reduction. Then, the low-dimensional system can be further transformed into a one-dimensional system by constructing appropriate value functions.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202202_5(2).0064

3.3. Modeling

The evaluation model of sales characteristics, inventory, discount and other related factors of target SKC products was established, and the principal component analysis method was used to make comprehensive evaluation. Principal component analysis (PCA) has a good dimension reduction technology, which can transform multiple indexes into multiple unrelated comprehensive factors. The comprehensive factor variables can reflect most of the information of the original index variables, and can solve the problem of multi-index evaluation better. By principal component analysis, m principal components are selected F1,F2,F3...Fn, and the comprehensive evaluation function is constructed with the relative contribution rate a_i of each principal component as the weight:

$$F = a_1F_1 + a_2F_2 + a_3F_3 + \cdots + a_iF_i$$

Which Fi(i=1,2,3...m) in the case of principal component score, when we calculate each index after scoring the first principal component, we can during the holiday, through the principal component scores, in various related factors affecting sales target SKC, the principal component scores generated comprehensive evaluation function, to calculate comprehensive score of related factors, in order to score, Score SKC for all relevant factors in holiday sales orders. Calculate the correlation coefficient matrix $R = (r_{ij})_{m\times n}$

$$\boldsymbol{r_{ij} = \frac{\sum_{k=1}^{n} a'_{ki}a'_{kj}}{n-1}, (i,j = 1,2,\dots,m)}$$

Calculate eigenvalues and eigenvectors

Calculate the eigenvalues of the correlation coefficient matrix R: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_m \geq 0$ and the corresponding eigenvectors: $\mu_1, \mu_2, \dots\dots, \mu_m$.

Among them $\mu_j = \left(\mu_{1j}, \mu_{2j}, \dots, \mu_{mj}\right)^T$ are m new index variables composed of eigenvectors.

$$y_1 = \mu_{11}\tilde{x}_1 + \mu_{21}\tilde{x}_2 + \cdots + \mu_{m1}\tilde{x}_m$$
$$y_2 = \mu_{12}\tilde{x}_1 + \mu_{22}\tilde{x}_2 + \cdots + \mu_{m2}\tilde{x}_m$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\boldsymbol{y_m = \mu_{1m}\tilde{x}_1 + \mu_{2m}\tilde{x}_2 + \cdots + \mu_{mm}\tilde{x}_m}$$

Where y_1 is the first principal component, y_2 is the second principal component, ... $\boldsymbol{y_m}$ is the MTH principal component.

The eigenvalue $y(j = 1,2,\dots,m)$ is the variance contribution rate and cumulative contribution rate.

$$b_j = \frac{\lambda_j}{\sum_{j=1}^{m}\lambda_k}(j = 1,2,\dots,m)$$

Variance contribution rate of principal component Yj

$$\alpha = \sum_{j=1}^{p} b_j$$

Main components Y1, Y2... , the cumulative contribution rate of yp.

When the eigenvalue is less than 1 or the cumulative contribution rate is greater than 85%, the first P index variables Y1, Y2... , Yp as P principal components, instead of the original M index variables, so p principal components can be comprehensively analyzed.The synthesis formula is as follows:

$$M = \sum_{j=1}^{p} b_j y_j$$

Where bj is the variance contribution rate of the JTH principal component, which can be evaluated according to the comprehensive score value.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202202_5(2).0064

4. Conclusion

Principal component analysis was used to treat the three related factors as three variables. SPSS was used to analyze the correlation coefficient between the three variables, and SPSS was used to process the data to obtain the correlation matrix, as shown in Table 1 (taking the correlation coefficient matrix of National Day as an example)

Table 1. Correlation matrix table

		Inventory	Sales characeristics	sales	discount
correlation	Inventory	1.000	0.320	0.192	0.231
	Sales characeristics	0.320	1.000	0.131	0.049
	sales	0.192	0.131	1.000	0.028
	discount	0.231	0.094	0.028	1.000

SPSS software was used to analyze the variance contribution rate of the feature vector of the correlation coefficient matrix, as shown in Table 2(Taking the variance contribution rate of feature vector of correlation coefficient matrix on National Day as an example)

Table 2. Variance contribution rate table

	Initial eigenvalue			Extract the sun of square of loads		
ingredients	total	Percentage ovariance	cumuative%	total	Percentage of variance	cumulative %
inventory	1.467	36.684	36.684	1.532	32.342	35.752
Sale scharacteristics	1.093	27.324	64.008	1.268	28.663	62.423
sales	0.887	22.170	86.178	0.754	21.154	88.468
discount	0.553	13.822	100.000	0.432	12.682	98.772

References

[1] Research on retail Marketing Strategy Mix and Retail Business Diversification [J]. Sun Pengtao. Brand (second half), (11): 33-34, 2015.

[2] Zhang Y, Zhang Y, Zhang Y, et al. A new approach to target threat assessment based on principal component analysis and MPSO-ELM [J]. Xi Zhifei, Xu An, KOU Yingxin, LI Zhanwu, Yang Aiwu. Acta Aeronautica Et Astronautica Sinica :1-21, 2019.

[3] Li Ying, Wang Haihe, Fu Chengcheng, Zhu Ling, Xu Wei. 2018(03). Analysis and Evaluation of water quality in Hongfeng Lake based on principal component Analysis [J].

[4] Wang Y, Wang Y, Wang Y, et al. Comprehensive evaluation of maize inbred lines based on principal component analysis and grey relational degree analysis [J]. Ye Kaimei, Chen Zehui, Zhu Yunfang, Wang Angui, Song Bi. Seeds. 2019(10).