



Copyright © 2017 American Scientific Publishers
All rights reserved
Printed in the United States of America

Prediction Model of Consumer Behavior and Volume with E-commerce

Min Duan^{1, 2, *}, Hao Wang³

¹ *School of Traffic and Transportation, XUCHANG University, Xuchang 461000, China*

² *School of Transportation Engineering, Beijing Jiaotong University, Beijing 100044, China*

³ *Chairman & CEO, Shanghai Jadecash Investment Management Co, Ltd, Shanghai 200433, China*

* *dm2308@163.com*

Abstract: This paper predicts the depth of the useful information such as the potential volume analysis, and establishes a customer e-commerce transaction model and volume forecast model based on being in specific e-commerce environment with large amounts of data for customer purchasing behavior. The paper relates to the study of calculation methods of transfer matrix, and the concrete prediction steps with an electronics store customer transaction data to verify the feasibility and validity of the model. Not only can offer the decision basis for enterprises to improve the quality of e-commerce business, the paper also improves the operation efficiency of e-commerce, which has important practical significance.

Keywords: E-commerce; sale volume; customer behavior.

1 INTRODUCTION

Compared with the traditional marketing by one-way mode, e-commerce provides the possibility of the direct contact between consumers and businesses, and through highly interactive network. This makes e-commerce operators or enterprise understand customer visit process and cycle by quickly using the network with high interaction characteristics and network data resources to, and clearly grasping the information such as customer transactions. Thus, e-commerce has changed the traditional marketing concept. The upstream suppliers or e-commerce operators in the supply chain begin to establish a supply of goods or services network market through the internet platform and information technology platform. The space of business activities is not a product swap space by traditional tangible physical, but it is a virtual space by breaking through the time and geographical restrictions, especially in the specific business situation, which can let the operator in the indefectible position in the competition^[1].

In the process of e-commerce, enterprises are facing with the different customer groups, the

different age, the different level of consumption and the different period. The huge market information of consumption network, such as the information of the sum of business transactions, it all led to the production of the new strategy, and it also is the secret of success by carrying out the electronic commerce for the current many enterprises.

The secret of success in electronic business for the enterprises is how to create useful database, and how to master the basic consumption demand information about the customer groups. Through these various ways, the enterprises will get the dynamic database of customers in many type markets including the type of shopping, the number of orders, the shopping time online, types of goods, payment amount, payment records, and so on. And this is a necessary work to forecast the huge trading information from the e-commerce customers. The e-commerce operators can understand customer actual transaction process and cycle by using existing database information, and predict the future behavior of customers with the collected data. It needs to count the customer log

by clicking on the website access in certain circumstances including the order time, the IP address and the specific e-commerce sites, which need to be in the different time, the different customer access to sales site records. So that can get a series of accurate report on customer behavior and the actual transaction amount.

Therefore, the upstream commodity suppliers in the supply chain can predictive parsing customers' behavior through e-commerce mode of operation, which can formulate the scientific and effective promotion strategies to improve the efficiency of suppliers including e-commerce enterprises, sales. At the same time, predicting the future behavior of customers has very important practical value of an increased corporate profitability for the associated enterprises in supply chain [2-3].

It is few studies about the forecast methods of e-commerce customer transactions and sum of business transactions of goods turnover in current academic and practical applications. This paper will establish the forecast model of e-commerce sites customer transactions and commodity turnover based on artificial neural network as a modeling tool. At the same time, the study will verify the validity and feasibility by the prediction model with some practical cases. Under certain circumstances, the research on the forecast methods of e-commerce customer transactions and the sum of business transactions of goods turnover can provide detailed data source for the unified statistical work of e-commerce transactions based on artificial neural network technology. It is convenient for suppliers in the upstream of commodity supply chain to establish and master the network sales database in a timely manner through the operation of e-commerce websites

2 RESEARCH ON CUSTOMER TRANSACTION FORECASTING WITH E-COMMERCE

2.1 Analysis of Prediction Model of Customer Transaction Behavior in E-commerce

This paper establishes a relatively effective forecasting model in order to realize the information about customer's access behavior. It can also provide an effective and simple approximate framework to analyze the transfer process of network customer transactions.

This study assumes n ($n = 1, 2, 3...$) be representative commodity types, using E to stand for visiting records in the web site on the end of the virtual state, and express the independent sequences for customers to visit the web site. The sequence can take for the visit records according to the customer access to certain times of sales pages stay without any click behavior, which access to

the customer behaviors to be divided into several independent access sequences [4-7].

The paper uses D to stand for the transition probability matrix for customers to access the electronic commerce network in a period, which is a discrete parameter random parameter to express customer access to the site of interest or desire. Using c to express transfer times for network customers' access or transaction behavior. P_{ij} is used to represent the probability that customer will be interested in a network access, and the visitorial probability from the i th product to the j th product.

$$P_{ij} = \frac{n_{ij}}{\sum_{j=1}^n n_{ij}} \tag{1}$$

Thus, the probability matrix D of the customer access behavior of e-business network can be established under certain conditions.

$$D = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \tag{2}$$

It needs to write down the information such as the client's IP address, access time, access record, and the number of transactions under certain circumstances and in a specific time period by recording the transfer log of client's access to electronic commerce network behavior. That can predict the transfer information by the data of log files obtained from the server used to customers' access to electronic commerce network behavior.

The detailed analysis steps, which predict client's access to electronic commerce network behavior or trading behavior, are given below.

Step 1: The data log entry statistics for the customers' access to electronic commerce network behavior in a particular case including the access or transaction of the client's IP address, access time, access or transaction behavior times and other information.

Step 2: According to the information of the customers' access to electronic commerce network behavior, the number of customer visits or transaction conversion times is arranged in an integer according to the time series.

Step 3: According to the integer in the time series for the information of the customers' access to

electronic commerce network behavior, gathering the statistics for the number of time which the customer visits the site and the final access to state.

Step 4: To calculate or calculate the probability vector of customers' access or transactional behavior in the selected period.

Step 5: The probability matrix of the website access or transaction behavior of all customers is calculated based on the probability vector of customers' access or transactional behavior in the selected period, the matrix can be expressed specifically.

$$D = \begin{pmatrix} \frac{\sum_{i=1}^N c_{11}}{\sum_{i=1}^N c_{11} + \sum_{i=1}^N c_{12} + \dots + \sum_{i=1}^N c_{1m}} & \frac{\sum_{i=1}^N c_{12}}{\sum_{i=1}^N c_{11} + \sum_{i=1}^N c_{12} + \dots + \sum_{i=1}^N c_{1m}} & \dots & \frac{\sum_{i=1}^N c_{1m}}{\sum_{i=1}^N c_{11} + \sum_{i=1}^N c_{12} + \dots + \sum_{i=1}^N c_{1m}} \\ \frac{\sum_{i=1}^N c_{21}}{\sum_{i=1}^N c_{21} + \sum_{i=1}^N c_{22} + \dots + \sum_{i=1}^N c_{2m}} & \frac{\sum_{i=1}^N c_{22}}{\sum_{i=1}^N c_{21} + \sum_{i=1}^N c_{22} + \dots + \sum_{i=1}^N c_{2m}} & \dots & \frac{\sum_{i=1}^N c_{2m}}{\sum_{i=1}^N c_{21} + \sum_{i=1}^N c_{22} + \dots + \sum_{i=1}^N c_{2m}} \\ \dots & \dots & \dots & \dots \\ \frac{\sum_{i=1}^N c_{m1}}{\sum_{i=1}^N c_{m1} + \sum_{i=1}^N c_{m2} + \dots + \sum_{i=1}^N c_{mm}} & \frac{\sum_{i=1}^N c_{m2}}{\sum_{i=1}^N c_{m1} + \sum_{i=1}^N c_{m2} + \dots + \sum_{i=1}^N c_{mm}} & \dots & \frac{\sum_{i=1}^N c_{mm}}{\sum_{i=1}^N c_{m1} + \sum_{i=1}^N c_{m2} + \dots + \sum_{i=1}^N c_{mm}} \end{pmatrix} \quad (3)$$

2.2 The Prediction Model of E-commerce Transactions with Artificial Neural Network

The artificial neural network is based on the physiological research results of the brain, which is designed to simulate some mechanism of the brain to realize the function of some aspects.

International famous expert in neural networks, Hecht Nielsen, who is the founder and leader of the first neural computer company, defined the artificial neural network: "the artificial neural network is set up by the human to have to the graph topology structure of dynamic system, it through to the input of continuous or intermittent state corresponding to information processing." The prediction method of artificial neural network is an abstract mathematical model based on the modern neuroscience research results, which expresses the control system as a mapping from input to output characteristics, so that it is a complete system state by prediction modeling and control of the process in the future.

First, this paper establishes the forecast model of the sum of business transactions of goods turnover by using artificial neural network in the prediction to electronic commerce network. That is a prediction matrix by setting up a good history of client access or transfer trading behavior as the basis of the prediction model of electronic commerce network.

This paper sets a data vector for commodity sales in the history, $C=(c_1, c_2, c_3, \dots, c_n)$, c_1 stands for the

current sales amount, c_2 is suitable for the previous sales amount, c_3 stands for the two-stage sales amount. C , the data vector which is standardized into vector $Sc=(sc_1, sc_2, sc_3, \dots)$, and constructs the four X_i sample vectors and the output Y_i by corresponding target. The sales amount in some period is associated with the sales amount of the front 4 stages. And then the four samples and the corresponding target output are put into artificial neural network to calculate and adjust the weights between neurons and closing value until the output of the error of all samples reach the lower limit, and the switching phase is completed.

In this way, the paper can accurately predict the sales amount of the most recent issues with the artificial network technology, and predict the sales amount sc_0 of the next issue based on that forecast quantity. As long as the sales amount of the front 4 stages (sc_1, sc_2, sc_3, sc_4) as sample are put into the system, the output of the network will be looked as the forecast amount of the next issue sales quantity sc_0 .

Among them, the output number of neuron is set m , and the input number of neuron is set n , and a stands for the constant between 1 and 10.

Then the hidden layer node is expressed as:

$$n_1 = \sqrt{n + m} + a \quad (4)$$

The second, the forecast function to the customers' spending amount is set in this paper.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

Based on artificial neural network, the prediction steps of e-commerce transaction network are described as follows:

Step 1: In the artificial neural network structure, the correlation values are initialized, and the weighted values of w_{ij} and the threshold values are lied in the interval (0, 1) as the random number. At the same time, the maximum number of iterations M and the target error are set, and the network error squared and the initial value are 0, and the following iteration is calculated according to the number of iterations.

Step 2: Providing the sales forecast with artificial neural network by random selection of a pattern for $X_i=(sc_{n+1}, sc_{n+2}, sc_{n+3}, \dots)$ and output phase $Y_k=sc_n$.

Step 3: The connection weight value w_{ij} and the threshold value θ_j are connected by the input value a_j , and the output value b_j of the middle layer unit is

calculated by the forecast function $f(x)$ to the customers' spending amount. s_j stands for neuron units.

$$s_j = \sum_{i=1}^n w_{ij} a_i - \theta_j \quad (6)$$

$$b_j = f(s_j) = \frac{1}{1 + e^{-s_j}} = \frac{1}{1 + e^{-\sum_{i=1}^n w_{ij} a_i - \theta_j}} \quad (7)$$

Step 4: The input value l_t of the output layer unit and the response value C_t of the output layer unit are calculated based on the connection weight value w_{jt} , the threshold value θ_j and the output value b_j .

$$l_t = \sum_{j=1}^6 v_{jt} b_j - \gamma_t \quad (8)$$

$$c_t = f(l_t) = \frac{1}{1 + e^{-l_t}} = \frac{1}{1 + e^{-\sum_{j=1}^6 v_{jt} b_j - \gamma_t}} \quad (9)$$

Step 5: The error value d_t of the output layer and the error value e_j of intermediate layer are calculated based on the connection weight value w_{jt} , the threshold value θ_j , the output value b_j and the response value C_t of the output layer unit.

$$d_t = (Y_t - c_t) c_t (1 - c_t) \quad (10)$$

$$e_j = (d_t) (v_{jt}) b_j (1 - b_j) \quad (11)$$

Step 6: The connection weight value w_{jt} and the threshold value θ_j are modified by using the error value d_t of the output layer and the error value e_j of intermediate layer.

Step 7: Random selecting a the next transfer calculation model to the supply network to go back to step 3, the whole transfer calculation will be finished, until the network error function E is less than the predetermined minimum value.

Tab 1. The data of customer behavior in the some toy shop by e-commerce.

	A	B	C	D	E	F	G	H	SUM
A	1	30	15	17	7	11	4	9	94
B	11	0	10	5	0	9	5	0	40
C	22	4	2	15	2	8	1	1	55
D	8	3	35	0	3	27	5	2	83
E	5	0	3	4	0	0	8	10	30
F	13	7	9	4	1	1	12	0	47
G	5	7	1	6	4	5	0	1	29
H	2	2	1	13	6	9	24	0	57
SUM	67	53	76	64	23	70	59	23	433

3 SIMULATION

This paper takes a toy sales website as an example by e-commerce with the first year of customer visits and sales and other log data. The toy sales

website mainly deals with various kinds of smart toys, and selects eight representative products, which are displayed by the customer access behavior records in the e-commerce website (shown in Table 1).

According to the customer access or the probability vector of transaction behavior, the probability matrix of website access or the transaction behavior can be calculated. The probability matrix can be expressed specifically.

$$D = \begin{bmatrix} 0.01 & 0.32 & 0.16 & 0.18 & 0.07 & 0.12 & 0.04 & 0.10 \\ 0.28 & 0 & 0.25 & 0.12 & 0 & 0.23 & 0.12 & 0 \\ 0.40 & 0.07 & 0.03 & 0.27 & 0.04 & 0.15 & 0.02 & 0.02 \\ 0.10 & 0.04 & 0.42 & 0 & 0.04 & 0.32 & 0.06 & 0.02 \\ 0.17 & 0 & 0.10 & 0.13 & 0 & 0 & 0.27 & 0.33 \\ 0.28 & 0.15 & 0.20 & 0.08 & 0.01 & 0.02 & 0.26 & 0 \\ 0.17 & 0.25 & 0.03 & 0.20 & 0.14 & 0.17 & 0 & 0.03 \\ 0.04 & 0.04 & 0.02 & 0.22 & 0.10 & 0.16 & 0.42 & 0 \end{bmatrix}$$

Based on the current situation of the customer's access, and combined with the interest transfer matrix, the connection weight matrix between the input layer and the hidden layer can be predicted to the approximate access range of the future customers.

$$W_{ij} = \begin{bmatrix} 3.6771 & -8.3036 & -11.6897 & 6.4435 & 11.5693 & -3.4291 \\ -3.2147 & -2.3645 & -1.4056 & -1.6092 & -0.7096 & 2.4820 \\ 2.187 & 1.5365 & 0.8722 & 1.4609 & 1.4761 & -1.9429 \\ 1.5700 & 3.4698 & -3.8485 & 4.9407 & -3.9260 & 3.9787 \end{bmatrix}$$

Through the above numerical reflect: this toy sales websites in e-commerce within a year, can predict the sales amount of the products with using the forecasting method of artificial neural network to e-commerce sales forecast. At the same time, the result is satisfactory, and the average prediction error is 2.10%, the maximum error of the prediction for 4 to 30%, and the minimum error index of the network is 0.0006, which less than 0.0001. This explains that the prediction error about the site visiting and sales of customers' behavior lie in the expected range.

4 CONCLUSION

This paper studies the customers' transfer and sales prediction based on the prediction technology with artificial neural network to e-commerce sales of client access to certain situations. The main purpose of this research supports the marketing decisions for the e-commerce operators, and it also proves that is feasible to predict the sales forecast of electronic commerce by using artificial neural network with the scientific prediction model.

REFERENCES

- [1] Z. C. Li, M. L. Liu, 2012. Research on consumer behavior in e-commerce environment. *Chinese Journal of Management Science*, vol. 10, no. 6, pp. 88-91.
- [2] R. Sen, M. H. Hansen, 2013. Predicting a Web user's next request based log data. *Journal of computational Graphics and statistics*, vol. 12, no. 1, pp. 143-155.
- [3] Q. M. Huang, Q. Yang, J. Huang, X. Zhe, et al., 2014. Mining of web-page Visiting patterns with Continuous Time Markov Models. *Lecture Notes in Computer Science*, vol. 3056, pp. 549-558.
- [4] S. Jespersen, T. B. Pedersen, J. Thorhauge, 2013. Evaluating the markov assumption for web usage mining. In: *Proceedings of the 5th ACM International Workshop on Web Information and Data Management*, ACM Press, pp. 82-89.
- [5] Y. Wang, 2009. Research and implementation of electronic business customer behavior and sales forecasting model. Beijing: Capital Normal University.
- [6] S. Li, A. Montgomery, K. Srinivasan, et al., 2012. Predicting online purchase conversion using Web path analysis. Graduate school of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA.
- [7] E. Magdalini, V. Michalis, K. Dimitris, 2015. Web path recommendations based on page ranking and Markov models. In: *Proceedings of the 7th annual ACM International Workshop on Web Information and Data Management*, Bremen, Germany, pp. 2-9.