TELKOMNIKA, Vol.11, No.11, November 2013, pp. 6434~6440
e-ISSN: 2087-278X

A Novel Algorithm of Network Trade Customer Classification based on Fourier Basis Functions

Li Xinwu*[1], Guan Pengcheng[2]
School of International Trade and Economics, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China, 330013
*Corresponding author, e-mail: liyue7511@163.com[1], 407706483@qq.com[1]

Abstract

Learning algorithm of neural network is always an important research contents in neural network theory research and application field, learning algorithm about the feed-forward neural network has no satisfactory solution in particular for its defects in calculation speed. The paper presents a new Fourier basis functions neural network algorithm and applied it to classify network trade customer. First, 21 customer classification indicators are designed, based on characteristics and behaviors analysis of network trade customer, including customer characteristics type variables and customer behaviors type variables. Second, Fourier basis functions is used to improve the calculation flow and algorithm structure of original BP neural network algorithm to speed up its convergence and then a new Fourier basis neural network model is constructed. Finally the experimental results show that the problem of convergence speed can been solved, and the accuracy of the customer classification are ensured when the new algorithm is used in network trade customer classification practically.

Keywords: *BP neural network algorithm, Fourier basis functions, customer classification, calculation flow, algorithm structure*

1. Introduction

Artificial neural network is a hot researeh field in recent years, involving electronic scienee and technology, information and communication engineering, computer science and technology, control science and technology, and many other disciplines. The application fields of artificial neural network include modeling, time series analysis, pattern reeognition and control, and many other fields related.

And for E-commerce enterprises, as there are vast customers in network transaction, these customers differ in thousands of ways. For different customers, their demands are ever changing. It's impossible for E-commerce enterprises to meet the demands of all the customers, which is not only limited by self material conditions of enterprises, but also undesirable in the aspect of economic benefits. Therefore, E-commerce enterprises shall pick out the most valuable customers whom they can effectively serve; instead of hitting out in all directions, enterprises shall provide them with more individual service, and give consideration to each transaction customer. So to correctly and effectively classify transaction customers plays a significant role for enterprises to carry out individual service and marketing strategies for different customers [1].

2. Literature Review

The widely-used methods of enterprises for customer classification at present are mainly qualitative method and quantitative method. As the qualitative method for customer classification is just to classify all the target customers of enterprises in the macroscopic level, customer classification is carried out according to different value emphasis of different customers. The formation of customer value is simply expressed as: Value = Benefit-Cost. Qualitative classification method classifies customers in a simple way, only offering guidance for customer classification of enterprise in the macroscopic level, unable to provide specific and credible basis for enterprise decisions; furthermore, as there is no strict process of argumentation, the method depends on decider's subjective inference, there may be certain

Received April 20, 2013; Revised June 24, 2013; Accepted July 11, 2013

deviations in the analysis process, easily resulting in faulty decisions. For this reason, to truly provide customer classification information beneficial to enterprises should depend on quantitative technology for customer classification [2, 3].

Quantitative classification method is to apply quantitative analysis technology to conduct customer classification on the basis of some specific customer variables (credit level of customers, purchasing power of customers, characteristics of demand of customers, etc.). Currently, there are mainly two categories of data mining for quantitative customer classification research, which are traditional statistical method and non-statistical method. The former mainly includes cluster analysis, Bayesian classification, factor analysis method, etc.; this statistics-based method is unable to process a great deal of sophisticated customer data, and there are some problems on the accuracy of customer classification results, so to fundamentally solve the problem of customer classification needs to rely on non-statistical customer classification method, which mainly includes neural network, fuzzy set method, association rules, genetic algorithm, etc. The classification technology based on neural network is combined with certain information technology, which is a kind of mathematical method applicable to complex variables and multi influencing factors calculation, so it is more effective in solving complex customer classification problems with better classification accuracy, however, the convergence problem of the function itself greatly limits its application value in specific project practice. Secondly, classification is mainly based on such mathematical methods as fuzzy clustering, rough set, association rules, etc., although these methods offer classification reason explanation in a relatively clear way with better classification results under the circumstances of satisfactory data conditions, the modeling process needs to provide specific mathematical equations. As a result, these methods are limited by data conditions in specific application, always having problems like insufficient classification accuracy or poor “robustness”, limiting the application in customer classification. Due to lots of influencing factors related to customer classification, more often than not, the complicated relations are difficult to be expressed in mathematical equations [4].

Customer classification models based on data mining have high classification accuracy but leaves behind the question of slow convergence speed of its algorithm. Therefore, it is hard to put into effect in customer classification. Based on BP neural network, Fourier basis functions neural network is being constructed with Fourier basis functions in this paper. In so doing, not only the problem of convergence speed has been solved, but also the simplicity of the model structure and the accuracy of the classification are ensured.

3. Selection of Customer Classification Indicators

The selection of reasonable classification variables is the basis of correct and effective customer classification, namely establishing scientific and reasonable classification indicators system. In view of the nature of trading and own characteristics of online trading, this Paper adopts customer characteristics type variable and customer behaviors type variable in the specific selection of customer classification variables [5].

3.1. Selection of Customer Characteristics Type Variable

Customer characteristics type variable is mainly used for getting the information of customers’ basic attributes. Such variable indicators as geographical position, age, sex, income of individual customer play a key role in determining the members of some market segment. This kind of variables mainly comes from customers’ registration information and customers’ basic information collected from the management system of banks, the contents of which mostly indicate the static data of customers’ basic attributes, the advantage of which is that most of the contents of variables are easy to collect. But some of the basic customer-described contents of variables are lack of differences at times [2, 3]

Based on analyzing and summarizing existing literatures, the customer characteristics type variables designed in this paper include: Customer No., Post Code, Date of Birth, Sex, Educational Background, Occupation, Monthly Income, Time of First Website Browsing, and Marital Status.

3.2. Selection of Customer Behaviors Type Variables

Customer behaviors type variables mainly indicate a series of variable indicators related to customer transacting behavior and relation with banks, which are used to define the

orientation which enterprises should strive for in some market segment, and are the key factors for ascertaining target market. Customer behaviors type variables include the records of customers buying services or products, records of customer service or production consumption, contact records between customers and enterprises, as well as customers' consuming behaviors, preferences, life style, and other relevant information [4].

Based on analyzing and summarizing existing literatures, the customer behaviors type variables designed in this paper include Monthly Frequency of Website Login, Monthly Website Staying Time, Monthly Times of Purchasing, Monthly Amount of Purchasing, Type of Consumer Products Purchased, Times of Service Feedback, Service Satisfaction, Customer Profitability, Customer Profit, Repeat Purchases, Recommended Number of Customers, Purchasing Growth Rate.

4. Research Method

4.1. Working Principle of BP Neural Network Algorithm

BP neural network algorithm has uique advantages to descript the non-linear relationship and strong function simulating capability. It not only has input and output-layer node, but also hidden-layer node. Its hidden-layer neurons adopt S type variation function and output-layer neurons use pure linear conversion function so that BP neural network algorithm can be close to the corresponding relationship between any functions and data if there are enough hidden layers and neurons theoretically. Therefore, in the study of color management, the mapping relations among the color spaces of different equipments can be derived through the training of standard output data and measurement data to complete their conversion. BP neural network algorithm is generally consisted of three layers of neurons as shown in Figure 1 [6, 7].

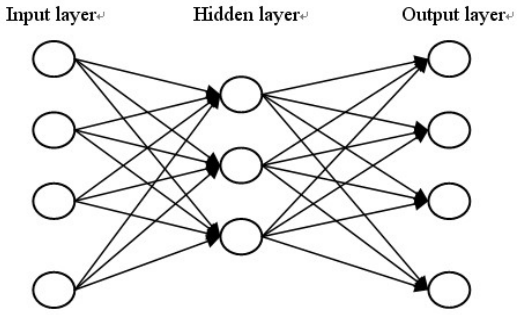

Figure 1. Working Principle of BP Neural Network Algorithm

4.2. Continuous-time Fourier Series of Periodic Signal

As we all know, for signal $f(t)$ that the period is T, it can be showed by continuous-time Fourier series, i.e. formula 1 [8].

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega_0 t) + \sum_{n=1}^{\infty} b_n \sin(n\omega_0 t) \quad (1)$$

Of formula 1, $\omega_0 = \frac{2\pi}{T}$ is fundamental angular frequency, a_0 is DC component, and a_n、b_n are Fourier series, i.e. formula 2.

$$a_0 = \frac{1}{T}\int_0^T f(t)dt \qquad a_n = \frac{2}{T}\int_0^T f(t)\cos(n\omega_0 t)dt \qquad b_n = \frac{2}{T}\int_0^T f(t)\sin(n\omega_0 t)dt \quad (2)$$

For time-limited nonperiodic signal $f(t)$, $0 \le t \le T$, the periodic signal that $f(t)$ is via continuation of period T is $f_p(t)$, i.e. formula 3.

$$f_p(t) = \sum_{m=-\infty}^{\infty} f(t - mT) \tag{3}$$

Of formula 3, m is a positive number. $f_p(t) = f(t)$ occurs obviously when time t is $0 \le t \le T$. Therefore, the continuous-time series of periodic signal $f_p(t)$ can be also showed by Formula 1 within the principal value period $0 \le t \le T$.

For bandlimited signal $f(t)(0 \le \omega \le N\omega)$, formula 1 can be changed as formula 4.

$$f(t) = a_0 + \sum_{n=1}^{N} a_n \cos(n\omega_0 t) + \sum_{n=1}^{N} b_n \sin(n\omega_0 t) \tag{4}$$

For the numerical computation, formula 4 is separated into formula 5.

$$f(k) = a_0 + \sum_{n=1}^{N} a_n \cos(n\omega_0 kT_s) + \sum_{n=1}^{N} b_n \sin(n\omega_0 tkT_s) \tag{5}$$

Of formula 6, T_s is a sampling period, and $T_s \le \frac{\pi}{N\omega_0} = \frac{T}{2N}$. When $T_s = \frac{T}{2N}$, formula 5 can be changed as formula 6.

$$f(k) = a_0 + \sum_{n=1}^{N} a_n \cos(\frac{\pi}{N} nk) + \sum_{n=1}^{N} b_n \sin(\frac{\pi}{N} nk) \tag{6}$$

In formula 6, $k = 0,1,2...2N-1$.

4.3. Improving BP Neural Network with Fourier Basis Function

In formula 6, neural network model based on Fourier basis function is p roduced if $f(k)$ is a neural network output, $f_d(t)$ is a neural network training sample, a_n、b_n are neural network training weights, and $\cos(\frac{\pi}{N} nk)$ and $\sin(\frac{\pi}{N} nk)$ are neural network excitation functions. See Figure 2 [9, 10].

The algorithm of neural network model based on fourier basis founction is as follows:

1. See formula 6 for neural network output.
2. See formula 7 for error function of network model.

$$e(k) - f_d(k) - f(k) \tag{7}$$

3. See formula 8 for network model performance index.
4. Weight adjustment by gradient descent algorithm, See formula 8 and 9 for weight adjustment quantity.

$$\Delta a_n^k = -\eta \frac{\partial J}{\partial a_n^k} = \eta e(k) \cos(\frac{\pi}{N} nk), \quad n = 0,1,2...N \tag{8}$$

$$\Delta b_n^k = -\eta \frac{\partial J}{\partial b_n^k} = \eta e(k) \sin(\frac{\pi}{N} nk), \quad n = 0,1,2...N \tag{9}$$

5. See formula 10 and 11 for weight adjustment, in which, η is a learning rate, and $0 \prec \eta \prec 1$.

$$a_n^{k+1} = a_n^k + \eta e(k)\cos(\frac{\pi}{N}nk), \quad n = 0,1,2...N \tag{10}$$

$$b_n^{k+1} = b_n^k + \eta e(k)\sin(\frac{\pi}{N}nk), \quad n = 0,1,2...N \tag{11}$$

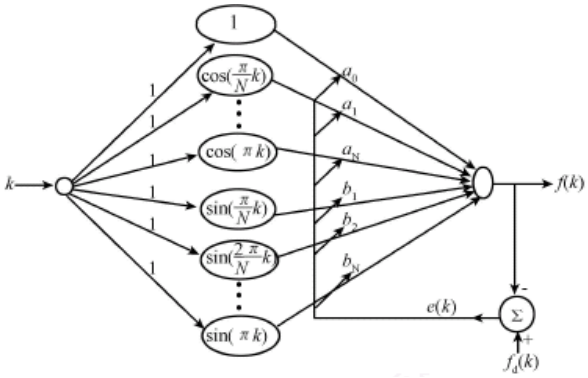

Figure 2. Working Principle of Fourier Basis Neural Network Model

4.4. Solution of Improved Algorithm

(1) Network training: The network training employs BP neural network algorithm algorithm by assigning the values of all the classification indicators from the training database as input value and that of classification weight as output. In this algorithm, both weight value and threshold value are randomly picked out in the range of -0.5~0.5, with adequate adjustment with regard to the real convergence.

(2) Initialization: to initialize the weight coefficient with a small random number.

(3) Circulation: to set an iteration number and load data to undergo network training. The weight coefficient required is acquired once the accuracy of designated customer is reached.

(4) Keep the value of weight coefficient of Fourier basis neural network and conclude the training.

4.5. Convergence Analysis of the Improved Model

As we all know, the size of learning rate η affects neural network convergence significantly. If too small, the convergence speed of neural network is slow and the computation amount and time are increased; if too large, neural network shocks not to reach the convergence. For absolute convergence of neural network, a theorem of neural network convergence is given as below.

Only when the learning rate η satisfies $0 \prec \eta \prec \frac{4}{3N+1}$, neural network algorithm is convergent. Here $2N$ is the number of neural network training samples. For the space limitation, see Reference 8 for the detailed proof of Theorem 1.

5. Results and Analysis

In order to test the effectiveness of improved algorithm in this thesis, simulation hardware is Dell Poweredge R710, in which processor is E5506, memory 2G, hard disk 160G; software platform is Windows XP operating system, Matlab710 programming language environment.

5.1. Process of Experimental Verification

The process of the experimental verification can be listed as follows.

(1) what is to be processed during the classification is the numeric data, so the numeric coding on character data should be conducted first;

(2) if the value number of certain attribute is equal to sample number, it means that it has little effect on classification, hence, remove such attribute first. Three attributes as Customer No., Post Code and Date of Birth are removed in this case.

(3) establish training sample set according to domain (prior) knowledge. Times of purchasing and total amount of purchasing of each customer are two major factors of customer classification (this is the prior knowledge of domain), so select 400 pieces of typical data among all the customers to form training sample set. And divide them into four types as Gold Customers, Silver Customers, Ordinary Customers, Potential Customers according to ABC management theory.

(4) use the customer classification algorithm above-mentioned, and the traditional BP neural network algorithm to classify customer.

5.2. Experimental Results

Experimental data come from the customer database registered by certain E-commerce enterprise. Relevant data of 10041 customers are randomly selected from the database to serve as the basis for data mining, and abstract the required customer value data to evaluate according to the designed evaluation indicator system. According to customers' overt value and potential value, test clusters customers into golden customer, silver customer, copper customer, general customer and ignorable customer; evaluation results are as shown in Table 1.

Table 1. Customer Classification Results of Some Website

Customer Type	Number of Customers	Percentage %	Profit Contribution Proportion
Gold Customers	786	7.83	53.13
Silver Customers	1278	12.73	30.92
Copper Customers	2622	26.11	13.94
General Customers	3456	34.42	6.13
Negligible Customers	1899	18.91	-4.12
Total	10041	100.00	100.00

From Table 1, we can see that Gold Customers occupies 7.83% of the total customer while the profits from them occupies 53.13% of the total. So the gold customers play an important role for the enterprises and they be treated with special service. However, the negligible customers account for 18.91%, who make minus profits for the enterprise.

In order to test the advantages and disadvantages of improved algorithm in this paper the improved model, the traditional BP neural network algorithm [7] and K-means algorithm [4] is realized in th paper, and specific experimental results of all these three algorithms is shown in Table 2. From Table 2 we can see that the algorithm presented in this paper has higher classification accuracy than that of the ordinary BP neural network algorithm and K-means algorithm.

Table 2. Classification Performance Comparison of different Algorithms

Algorithm	Algorithm in This Paper	Ordinary BP Neural Network Algorithm	K-means Algorithm
Accuracy Rate	99.23 %	93.07%	84.36%
Time Consuming(S)	18	503	17

6. Conclusion

The research of neural network in theory and application is still developing. And how to correctly and effectively carry out correct and reasonable classification on network transaction customers, reform network marketing mode and improve customer management and service level is also a key to increase the competitiveness of E-commerce enterprises. This paper, on account of the shortage of BP neural network in data mining, puts forward a new Fourier basis

neural network model to classify network trade customers. Experimental results show that the improved network trade classification algorithm has enhanced the accuracy of customer classification, more reasonable in classification results.

Acknowledgement

This work is supported by scientific research project of the education department of Jiangxi province (No. GJJ13300) and 52nd Chinese postdoctoral fund (No. 2012M521284).

References

[1] Arwa M, Safi A. CRM Scorecard - CRM Performance Measurement. *International Journal of Networked Computing and Advanced Information Management*. 2012; 2(1): 8-21.
[2] Deng WB, He MS. B2C Customer Classification Algorithm Based on Based on 3DM. *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*. 2012; 24(4): 568-572.
[3] Riehar SF, Norma TD. Study of Classifying Customers Method in CRM. *Journal of Computer Simulation*. 2011; 28(8): 249-254.
[4] Sulma LE, Stalk E. Research on Customer Classification of E-Commerce Website Based on Customer Value Analysis. *Journal of Computer Engineering*. 2011; 24(6): 289-294.
[5] Teece DJ, Stalk E. Application of Commercial Bank Customer Subdivision Based on K-means. *Journal of Information Management*. 2011; 7(3): 199-209.
[6] Mardiyono, Reni S, Azlan A. Intelligent Monitoring System on Prediction of Building Damage Index Using Neural-Network. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2012; 10(1): 155-164.
[7] Budi R, Supriyadi. Early Model of Traffic Sign Reminder Based on Neural Network. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2012; 10(2): 713-722.
[8] Post LP, Chiristopher C. An Evaluation of Methods for Producing CRT Monitors. *Color Research and Its Application*. 2007; 14(4): 172-186.
[9] Roy SK, Mark E. CRT Colorimetry Based on Improved BP Neural Network. *International Review on Computers and Software*. 2009; 38(4): 299-313.
[10] Li JJ, Rui L. Construction Equipment Control Research Based on Predictive Technology. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(5): 960-967.