Research Article

Qianqian Pan* and Gang Yang

Application of mining algorithm in personalized Internet marketing strategy in massive data environment

Abstract: Internet marketing requires a personalized marketing strategy. In this study, the application of data mining in personalized Internet marketing was studied. Based on the mining algorithm, a personalized marketing method was designed. Through the calculation of frequent closed item sets and support counts of positive and negative samples, the interval with a high success rate for marketing was obtained. With performance analysis, it was found that the success rate of the marketing method proposed in this study improved 8% compared with the traditional marketing method and had a better performance under the smaller interval number and smaller minimum success number. After applying the designed method in telecommunication enterprise A, it was found that after adopting the marketing method of this study, the marketing success rate of enterprise A increased from 2.72 to 6.31%, which indicated the effectiveness of the method. The research results of this study verify the role of data mining algorithms in Internet marketing, which is conducive to the further application of mining algorithms in personalized marketing and innovation of business modes.

Keywords: data mining, personalized marketing, internet marketing, big data

1 Introduction

With the development of information technology, the e-commerce model has occupied an increasingly important position in daily consumption and has become the main consumption method of many consumers [1]. Under this circumstance, marketing tools also change a lot. Driven by the big data environment, the development of Internet marketing is inevitable [2]. Online marketing can increase the visibility of products and thus influence users' purchasing decisions [3]. Compared with traditional marketing strategies, Internet-based marketing strategies have wider coverage and stronger interactivity, providing more audiences and improving marketing efficiency while reducing marketing costs. Internet marketing often requires a large amount of data support [4]. High-quality data analysis techniques can well improve the effectiveness of online marketing [5], and the emergence and development of data mining technology make Internet marketing more mature. With the help of mining algorithms, personalized marketing and precision marketing can be achieved better [6], and marketing efficiency is further improved. Personalized marketing can provide highly targeted services and improve marketing efficiency for enterprises [7], and more and more enterprises pay attention to it with the support of data mining. Huang et al. [8] analyzed the role of data mining technology in identifying customer behavior changes, proposed a transformation-mining

---

* Corresponding author: Qianqian Pan, College of Business Administration, Zhejiang Institute of Economics and Trade, Hangzhou, Zhejiang 310018, China, e-mail: p257o1@163.com
Gang Yang: College of Business Administration, Zhejiang Institute of Economics and Trade, Hangzhou, Zhejiang 310018, China

model for detecting changes in fuzzy quantitative sequence patterns, and verified the effectiveness of the model through real-world data sets. You et al. [9] selected the attributes through the rational function model, then identified the important attribute values through the Chi-squared Automatic Interaction Detector decision tree and Pareto values, differentiated the customer groups, implemented different marketing strategies for different customer groups, and conducted personalized marketing. Jiang et al. [10] studied the slope algorithm and designed a new one based on the combination of trusted data and user similarity fusion. The user similarity was used as the weight factor of the slope algorithm. Experiment on Amazon dataset suggested that the algorithm was effective in user recommendation and could realize the application on personalized marketing. Han [11] established a logical model of personalized marketing according to the characteristics of e-commerce, and then combined it with the filtering recommendation algorithm to improve the efficiency of personalized marketing. The implementation of personalized online marketing strategies cannot be achieved without the support of mining algorithms. What kind of mining algorithms can be used to obtain good results is the research subject of this study. In this study, a telecommunication enterprise was taken as an example to design a personalized marketing method based on mining algorithm. By mining the set of frequent closed items (FCIs), the marketing data characteristics of telecommunication companies were divided into different intervals, and then the algorithm found out the interval with the highest marketing success rate to realize personalized online marketing. The performance analysis verifies the effectiveness of the method and provides some theoretical bases for the further application of data mining technology in personalized marketing.

2 Personalized Internet marketing

In the context of the Internet, Internet marketing [12] has a tendency of accurate user targeting, diversified communication channels, and exponential increase of data. The way to obtain user information in traditional marketing mainly comes from market research, which takes a long time, costs a lot, and has strong subjectivity. Internet marketing can subdivide users and accurately locate target users with the help of data mining technology. In addition, there are many channels for Internet marketing, and various new media platforms are good bridges for Internet marketing. With the development of e-commerce and Internet marketing, its accumulated data are also increasing. Therefore, the processing of these massive data is the key to Internet marketing.

With the development of e-commerce, more and more enterprises need to conduct precise and personalized marketing to gain competitive advantage [13]. Personalized Internet marketing refers to the marketing strategy of e-commerce company that is accurate to the individual through the mining of different consumer demand, consumption characteristics, and other data. It can make full use of the value of every kind of marketing, and quickly and accurately present product information to users who need to buy. In the process of Internet marketing, if users continue to receive marketing information that they are not interested in, it is likely to cause users to feel less favorable to the business and conflicted, and affect the success rate of marketing, which requires personalized marketing.

In Internet marketing, e-commerce companies can easily collect data such as users' browsing and purchases [14], which serves as a basis for the development of personalized marketing strategies. Such data are very large and complex, and it is necessary to require the help of data mining technology to obtain useful information. The mining algorithm can obtain valuable information through the steps of collecting, filtering, and mining massive data [15], thus personalized marketing is achieved.

For telecommunication enterprises, they have a tendency to recommend advertisements to as many users as possible when recommending packages. The success rate of this method is low, resulting in a huge waste of marketing resources. This study investigated the personalized Internet marketing of telecommunication enterprises.

3 Personalized Internet marketing strategies based on mining algorithm

3.1 Mining algorithm

$$I = \{i_1, i_2, i_3, \cdots, i_m\}, \tag{1}$$

is the collection of a all the items,

$$T = \{T_1, T_2, T_3, \cdots, T_n\}, \tag{2}$$

is the collection of all the transactions, and the item set in T_i is the subset of I. An item set having k items can be set as the k-item set. For example, {fried-chicken, beer} is a two-item set. The number of transactions of item set A is the support count (SC) and can be set as follows:

$$SC(A) = |\{t_i|A \subseteq t_i, t_i \in T\}|, \tag{3}$$

and the support of item set A can be set as follows:

$$Support(A) = \frac{Support(A)}{N}, \tag{4}$$

where N is the number of transactions.

Association rule can be shown as $A \to B$, $A \subset I$, $B \subset I$, and $A \cap B = \varnothing$. It is assumed that $C\%$ transactions in T all include A and B, then the support of association rule of $A \to B$ is $C\%$:

$$Support(A \to B) = \frac{Support(A \cup B)}{N} = C\%. \tag{5}$$

It is assumed that $C\%$ of transactions in T have A, at the same time, they have B, and then the confidence of $A \to B$ is $D\%$:

$$Confidence(A \to B) = \frac{Support(A \cup B)}{Support(A)} = D\%. \tag{6}$$

If the direct super set of A does not have the same SC, A can be called as a closed item set. If the support of the closed item set is larger or smaller than the threshold of minimum support (MS), it will be a FCI set.

3.2 Formalized description of telecommunication marketing

To formally describe the personalized Internet marketing of telecommunication enterprises, the marketing data characteristics of telecommunication enterprises include call charges, traffic, and text messages. It is assumed that there is n-dimensional characteristic X in samples. These characteristics are divided into different intervals and numbered, and then characteristic mapping is performed by the interval number; that is, the original characteristic X is mapped into the interval characteristic Y to get the following equation:

$$F : X \to Y. \tag{7}$$

It is assumed that if the number of positive samples satisfied the condition is a in the item set whose length is m, and the item set

$$I' = \{i_1, i_2, i_3, \cdots, i_m, i_{m+1}\}, \tag{8}$$

whose length is $m + 1$, and the number of negative samples is b and b'. Because of the existence of i_{m+1},

$$b \geq b', \tag{9}$$

can be obtained, then

$$\frac{a}{a+1} \leq \frac{a}{a=b'},\tag{10}$$

can be obtained, that is,

$$f_I \leq f_{I'}.\tag{11}$$

In the data sample of the telecommunication enterprise, the sample of users who selects the recommended package is called the positive sample, which is denoted as 1, and the number of it is N. The sample of users who does not select the recommended package is called the negative sample, which is denoted as 0, and the number of it is M. In the item set of the hit rate of package is as follows:

$$f = \frac{a}{a+b}.\tag{12}$$

The threshold of MS is set. If $a > msN$, then the item set is the frequent item set. The threshold of minimum count (MC) rate is set, and the frequent item set of $f \geq mc$ is the mining target.

3.3 Specific steps

The specific steps of personalized marketing based on mining algorithms were as follows:
(1) n-dimensional characteristic was divided into different intervals and numbered.
(2) The MS was set, and the positive samples were mined to get the FCI and SC. If a was smaller than the minimum SC, the combination would be discarded.
(3) FCI and SC for negative samples were calculated.
(4) Calculation was performed using the following equation:

$$f_i = \frac{\text{FCI}_{\text{SC}_a}}{\text{FCI}_{\text{SC}_a} + \text{FCI}_{\text{SC}_b}}.\tag{13}$$

(5) All the FCIs of $f \geq mc$ were sorted as f, and the obtained interval was the characteristic combination interval that satisfied the condition.

After mining, the result was as follows: the probability of the user whose charge was $a–b$, traffic was $c–d$, and short message was $e–f$ selected a package was $X\%$, and it could be known which kinds of combination package would have higher marketing success rate according to the mining result.

4 Performance analysis of the mining algorithm

To verify the effectiveness of the data mining algorithm, the marketing method based on the mining algorithm was compared with the traditional marketing method. The success rate of the package marketing of the two methods is shown in Table 1.

It was found from Table 1 that compared with the traditional method, the designed method had significantly improved success rate of package marketing, and the average success rate of package marketing reached 12.32%, which was about 8% higher than the traditional method. The above results suggested that the data mining-based personalized marketing could greatly improve the marketing efficiency. In addition, the data mining-based personalized marketing did not need to spread the net widely, which greatly reduced the marketing cost, decreased the possibility of wasting marketing resources, and maximized the value of marketing.

Table 1: Comparison of marketing success rate between the data mining-based method and the traditional method

		Positive sample	Negative sample	Success rate of package marketing (%)	Average success rate (%)
The data mining-based method	Data set one	3,924	26,076	13.08	12.32
	Data set two	4,785	35,215	11.96	
	Data set three	5,962	44,038	11.92	
The traditional method	Data set one	1,256	24,562	4.86	4.67
	Data set two	1,569	38,431	3.92	
	Data set three	2,615	47,385	5.23	

To further understand the optimal design of the algorithm, the influence of interval division on results was studied, and the number of characteristics divided was analyzed. The minimum success number of each interval was set as 30, and the characteristics were divided into 10, 20, 30, 40, and 50. Then the success rates of the package marketing were compared. The results are shown in Figure 1.
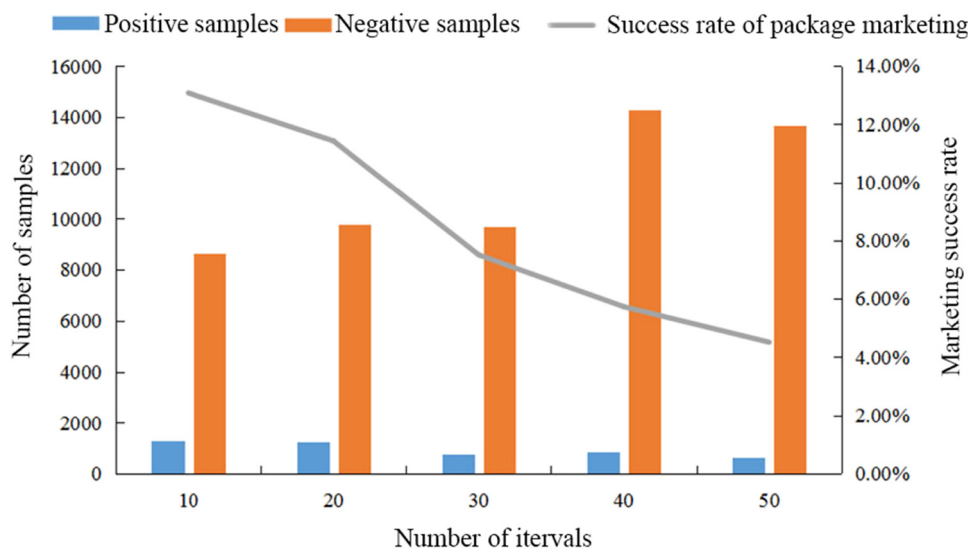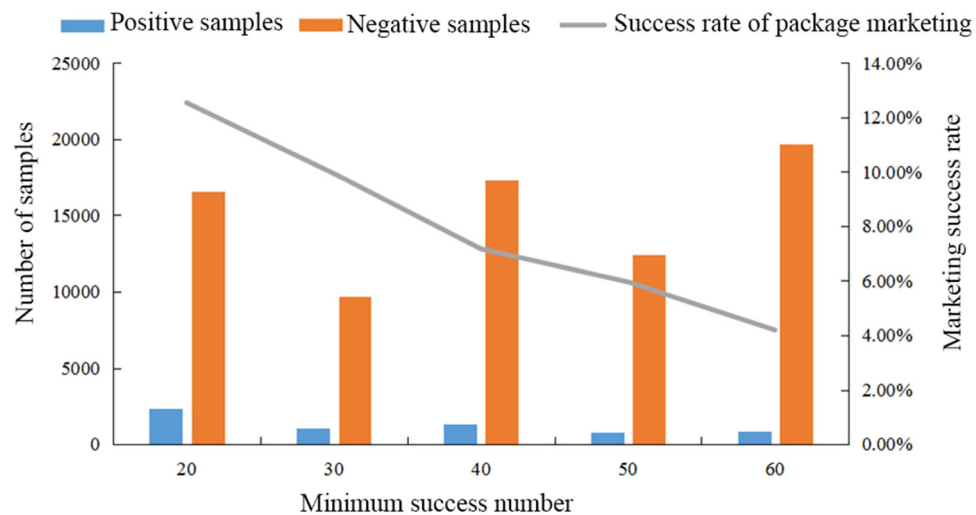

Figure 1: Marketing success rate under different intervals.

Figure 1 shows that the success rate of package marketing gradually decreased as the number of divided segments increased. When the number of intervals was 10, the success rate of package marketing was 13.07%; when the number of intervals was 50, the success rate of package marketing was 4.51%, which was 8.56% less than when the number of intervals was 10. The reason for the above results was because the larger the number of interval divisions was, the smaller the number of samples in the interval was, the fewer the qualified users selected were, and the lower the success rate of package marketing was.

Similarly, the minimum success number was analyzed. The number of intervals was set as 10, the minimum success number was set as 20, 30, 40, 50, and 60, respectively, and the success rate of the package marketing was compared. The results are shown in Figure 2.

Figure 2: Marketing success rate under different minimum success number.

Figure 2 shows that the success rate of package marketing became lower as the minimum success number increased under the same number of intervals. When the minimum success number was 20, the success rate of package marketing was 12.53%, and when the minimum success number increased to 60, the success rate of package marketing was 4.19%, which was 8.34% less than when the minimum success number was 20. The reason for the above results was because the larger the minimum success number was, the smaller the number of samples required by the load was, and the lower the success rate was.

In conclusion, the data mining algorithm was practical in the personalized marketing of telecommunication enterprises. Telecommunication enterprise A was taken as an example. The marketing success rate of the traditional method in September 2018 was 2.72%. In October, November, and December, the mining algorithm was used for personalized marketing, and the results obtained are shown in Table 2.

Table 2: Marketing situation of telecommunication enterprise A from September to December

	September	October	November	December
Total number of customers receiving marketing	25,678	24,687	25,987	26,128
Number of successful marketing	698	1,536	1,542	1,648
Marketing success rate (%)	2.72	6.22	5.93	6.31

It was found that after implementing the personalized marketing strategy based on the mining algorithm, the number of users who bought the packages of enterprise A increased significantly, and the marketing success rate increased significantly, from 2.72 to 6.31%, which suggested the reliability of the method.

5 Discussion

The Internet has become an important tool to promote business development [16]. In the network environment, due to the convenience of data acquisition, the marketing method of companies is more precise and personalized. Personalized Internet marketing can accurately target the development of marketing strategies to individuals through the understanding of user behavior, psychology, and other data,

thereby reducing marketing costs, greatly improving the success rate of marketing [17], and promoting the development of e-commerce companies better and faster. The formulation of marketing strategies is inseparable from the processing and analysis of massive data [18]. At present, many e-commerce companies do not use data in a traditional way to make decisions when they make decisions. Fuzzy marketing cannot effectively control marketing costs and efficiency. Therefore, data mining technology is essential to obtain valuable information from fragmented, complex, and large data [19]. Through mining algorithms, e-commerce companies can turn useful information into knowledge and serve marketing to make the right decisions and create value. It can be found that data are the basis for personalized marketing, and formulating personalized strategies through data mining can give full play to the potential value of each user, help improve the competitive advantage of enterprises [20], and maximize the interests of enterprises.

This study focused on the marketing strategy of telecommunication enterprises. The competition among telecommunication enterprises becomes increasingly serious with the growth of mobile users. The emergence of various value-added services makes it more and more difficult for telecommunication enterprises to win customers. In this environment, how to properly market becomes more and more important. At present, when telecommunication enterprises recommend packages to users, they often use as many advertisements as possible to market them to all customers. However, this marketing strategy has a very low success rate and consumes many marketing resources, which is not conducive to the long-term development of the telecommunications enterprise. Thus, only the personalized marketing strategy can improve the marketing effect of the enterprise.

This study designed a personalized marketing method based on mining algorithm. It calculated interval classification, numbering, and FCIs of marketing data characteristics to find out the interval of the highest success. It was found in the performance analysis that compared with the traditional marketing method that spread the net widely, the data mining-based marketing method had a significantly higher success rate. The average marketing success rate of the traditional method was only 4.67%, but the data mining algorithm-based method had an average success rate of 12.32%, which was 8% higher than the traditional method. Then, in the study of optimal marketing methods, it was found that a large number of intervals and a large number of minimum successes would reduce the marketing success rate, which was because the number of eligible samples decreased. Thus, it was concluded that it was necessary to control the number of intervals and the minimum number of successes to achieve an efficient marketing. In the case analysis, the marketing mode in enterprise A was improved by the proposed data mining algorithm. It was found that the data mining-based personalized marketing method improved the marketing success rate of enterprise A from 2.72 to 6.31% and kept it stable at 6%.

6 Conclusion

In this study, the application of data mining in personalized marketing was studied. A data mining algorithm-based personalized marketing method was designed for telecommunication enterprises, and its performance was analyzed. It was found that compared with the traditional method, the data mining algorithm-based personalized marketing had a higher success rate. The marketing success rate increased from 2.72 to 6.31% when the designed personalized marketing mode was used in enterprise A, which demonstrated the effectiveness of the method and provided some theoretical support for the application and development of mining algorithms in personalized marketing.

Conflict of interest: Authors have no conflicts of interests.

References

[1] Zitkiene R, Gircys V, Zitke M, Bartuseviciene I. Model of impact of social networks on internet marketing of enterprises. SHS Web Conf. 2021;92(2):09020.

[2] Yau HK, Tang HYH. Analyzing ecology of Internet marketing in small- and medium-sized enterprises (SMEs) with unsupervised-learning algorithm. J Mark Anal. 2018;6:1–9.

[3] Soomro HJ, Khaskheli GA, Jumani AK, Hussain A. Intervening analysis of internet marketing between the relationship of brand awareness and online purchase decision of mobile products among the students of HEIS in Pakistan. Ind J Sci Technol. 2020;13:1535–45.

[4] Elsalamony HA. Bank direct marketing analysis of data mining techniques. Int J Comput Appl. 2014;85:12–22.

[5] Ikumoro AO, Jawad MS. Assessing intelligence conversation agent trends-chatbots-ai technology application for personalized marketing. Test Eng Manag. 2020;81:4779–85.

[6] Gao H, Kuang L, Yin Y, Guo B, Dou K. Mining consuming behaviors with temporal evolution for personalized recommendation in mobile marketing apps. Mob Netw Appl. 2020;25:1233–48.

[7] You Z, Si YW, Zhang D, Zeng X, Leung S, Li T. A decision-making framework for precision marketing. Expert Syst Appl. 2015;42:3357–67.

[8] Huang CK, Chang TY, Narayanan BG. Mining the change of customer behavior in dynamic markets. Inf Technol Manag. 2015;16:117–38.

[9] You Z, Si YW, Zhang D, Zeng X, Leung S, Li T. A decision-making framework for precision marketing. Expert Syst Appl. 2015;42:3357–67.

[10] Jiang L, Cheng Y, Yang L, Li J, Yan HY, Wang XQ. A trust-based collaborative filtering algorithm for E-commerce recommendation system. J Amb Intel Hum Comp. 2018;10:1–12.

[11] Han M. The design and implementation of e-commerce personalized services based on collaborative filtering recommendation system. Appl Mech Mater. 2014;687–691:2039–42.

[12] Vynogradova O, Drokina N, Yevtushenko N, Darchuk V, Irtlach M. Theoretical approaches to the definition of Internet marketing: Ukrainian dimension. Innov Mark. 2020;16:89–103.

[13] Luo FL, Li S. Precision marketing and modern information technology. Appl Mech Mater. 2014;687–691:4683–6.

[14] He J, Liu H, Xiong H. SocoTraveler: travel-package recommendations leveraging social influence of different relationship types. Inf Manag. 2016;53:S0378720616300362.

[15] Zhao X, Huangfu YH. Research on data mining in the information network application. Appl Mech Mater. 2015;713–715:1839–42.

[16] Andreopoulou Z, Tsekouropoulos G, Koliouska C, Koutroumanidis T. Internet marketing for sustainable development and rural tourism. Int J Busin Inf Syst. 2014;16:446–61.

[17] Sun L. The application design of personalized recommendation system based on data mining to e-commerce. Adv Mater Res. 2014;989–994:4538–41.

[18] Ozyirmidokuz EK, Uyar K, Ozyirmidokuz MH. A data mining based approach to a firm's marketing channel. Proc Econ Financ. 2015;27:77–84.

[19] Abessi M, Yazdi EH. Marketing data mining classifiers: criteria selection issues in customer segmentation. Int J Comput Appl. 2014;106:5–10.

[20] Shen AY. Recommendations as personalized marketing: insights from customer experiences. J Serv Mark. 2014;28:414–27.