2023 International Conference on Networking, Informatics and Computing (ICNETIC) | 979-8-3503-1331-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICNETIC59568.2023.00136

KNN-based Precision Marketing Data Mining Algorithm for E-business Platform

Xiao Gao
Xiangtan Institute of Technology, Xiangtan, Hunan, 410000 China
2545777640@qq.com

***Abstract*—The development of e-business industry is undoubtedly the needs of both the state and the society. With the continuous improvement of the online shopping system, consumers' consumption psychology and consumption form are changing, and they are more and more fond of businesses that can provide personalized and customized services. Based on the K-Nearest Neighbor (KNN) algorithm of artificial intelligence machine learning, this paper analyzes the potential consumption demand of customers, and realizes the prediction of the consumption behavior of target users, thus contributing to the precise marketing of e-business products. The results show that the algorithm is more accurate in the prediction of consumers' consumption tendency on e-business platform, which is 23.25% higher than the traditional CNN algorithm. Therefore, the precision marketing system model of e-business products designed in this paper has high precision in predicting consumers' purchase intention, and can be used as decision-making basis for enterprises in its precision marketing. Under the data mining technology, e-business enterprises should master the consumption rules and visiting time of consumers, assist them to formulate the most appropriate network marketing strategy, and enhance their own core competitiveness in the highly competitive market.**

Keywords*—E-business, Accurate marketing, Data mining, Consuming behavior

I. INTRODUCTION

Due to the continuous development of market economy, the traditional seller's market has already been transferred to the buyer's market. Moreover, based on the increasingly common personalized demand of consumers for products and services, the entire marketing environment has already undergone tremendous changes, which has led to the fact that enterprises can no longer use traditional marketing methods and means to meet the personalized needs of consumers, and thus obtain economic benefits. The development of e-business industry is undoubtedly the need of both the country and the society [1-2]. From the perspective of philosophy, the source of development of anything comes from the internal contradiction and external thrust of things. The development of business itself comes from people's own demand for more materials [3]. E-business is generated and developed under the action of people's complex demand contradiction for more material exchange and more commercial exchanges [4]. From the perspective of enterprises, the purpose of marketing is to find the target market for enterprises, and then develop a series of marketing strategies to serve the target customers and ultimately obtain the target customers, create profits and improve benefits [5]. With the continuous improvement of the online shopping system, consumers' consumption psychology and consumption form are changing, and their shopping is more and more like those who can provide personalized and customized services. Therefore, e-business enterprises can not meet the needs of consumers today by only adopting traditional marketing methods [6].

The innovation of e-business network marketing strategy changes with the change of the times and the progress of technology. Only by giving full play to the advantages of data mining technology from the perspective of innovation can it really help to play a positive role in the progress of e-business network marketing and truly promote the benign development of e-business network marketing [7]. In the daily marketing process, consumers will make daily consumption in the e-business platform and generate a large amount of data information. By analyzing these personalized data, we can effectively predict and analyze the consumer's consumption behavior, and then dynamically perceive the consumer's upcoming consumption activities to achieve consumption upgrading [8]. The core of accurate marketing is to accurately find users, which requires e-business enterprises to focus on users, fully understand the needs and preferences of each customer, and adopt different marketing methods for different users [9]. Mining and analysis of customers' potential consumption demand based on customer knowledge can help enterprises find out more factors affecting potential benefits and lay a foundation for fully realizing the double maximization of enterprises' own resources and customer demand [10]. In this paper, the KNN algorithm based on artificial intelligence machine learning realizes the prediction and analysis of the target user's consumption behavior, thus helping the accurate marketing of e-business products and promoting the development of e-business industry.

II. ADVANTAGES OF E-BUSINESS NETWORK MARKETING AND APPLICATION OF DATA MINING TECHNOLOGY

(1) Advantages of e-business network marketing

In the actual development process of e-business network marketing, the diversification of product forms is an important development feature. People can buy products and obtain corresponding services on the Internet. In the propaganda activities of contemporary enterprises, long-term tracking of the overall marketing process is realized, effective measurement and evaluation of marketing results are realized, and then adjustments are made according to the conclusions drawn [11]. At present, in the development of e-business network marketing, more and more emphasis is placed on product personalization, which is quite different from the traditional marketing activities. E-business network marketing puts more emphasis on personalized service, and can provide corresponding products without customers. In the early e-business business, there was little innovation in marketing mode, and the

979-8-3503-1331-4/23/$31.00 ©2023 IEEE
DOI 10.1109/ICNETIC59568.2023.00136

network was usually used as a platform for information release and collection. Relying on the new precision marketing system, enterprises will give timely feedback to the appropriate target groups, and respond quickly to make the next adjustment. The progress of contemporary network technology and the efficient application of big data system ensure the effect and process of accurate marketing services. Using the existing technology to monitor and analyze the effect of each stage in the whole marketing process, we can optimize the basic procedures of marketing in the process of realizing marketing objectives.

(2) Application of data mining technology

In the practical application process of e-business network marketing data mining technology, it plays a prominent role. On the basis of changing the traditional network marketing mode, it can effectively improve the overall quality effect of e-business network marketing. The application of data mining technology can combine users' information for targeted marketing, and effectively distinguish customers on the basis of the consumer transaction information recorded in the transaction database, thus achieving the role of market segmentation and finding valuable customer groups [12]. Through these data, it can help the relevant staff of e-business platform to improve the site location and density of network sites, optimize the user experience of e-business platform and improve the transaction security of e-business trading platform. During the development of e-business network marketing, we should pay full attention to data mining technology and play an active role. With the application of data mining technology, we can discover the past transaction records of customers in time and find new profit points.

III. E-BUSINESS PRECISE MARKETING STRATEGY BASED ON KNN ALGORITHM

In today's information age, the number of users of e-business platform is growing rapidly, and the original marketing model can't match the complex needs of today's society, so precision marketing came into being [13]. More and more enterprises are committed to how to do a good job in their marketing activities in the complex and changeable marketing environment, so as to achieve truly targeted and accurate marketing. For e-business platforms, the cost of opening up new consumer users is much higher than that of maintaining a consumer user. Therefore, many e-business platforms begin to turn their attention to precision marketing. Applying precision marketing can effectively capture target users, conduct purposeful marketing activities for specific users that meet their needs, and improve the success rate of marketing. Precision marketing relies on modern scientific information technology to segment the market on the premise of insight into consumers' behavior habits. Then, on this basis, select one or several suitable markets as target markets, and make accurate market positioning, and then provide targeted and differentiated marketing strategies for target customers to gain market competitive advantage. The composition of data mining system in e-business precision marketing is shown in Figure 1.

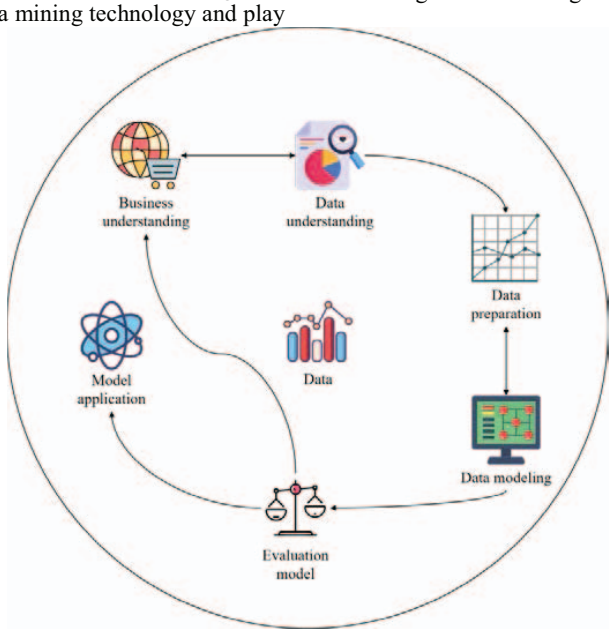

Figure 1. The composition of the data mining system

When customers buy a product, they not only pay money, but also pay time, energy, etc. These are the costs paid by customers to buy the product, and the sum of these costs is the total cost paid by customers when they buy the product. Similarly, when customers buy a product, they hope to get not only the value of the product itself, but also more service values other than the product, such as person value and value of the image. In this paper, a KNN algorithm-based prediction solution for e-business product precision marketing is proposed, that is, users' behaviors are hammed, and then the distance between users is calculated, so as to predict users' future consumption behaviors, and KNN algorithm is applied to the specific business scenario of e-business product precision marketing. A precise marketing system of e-business products based on KNN algorithm is constructed, and the extraction of feature items in e-business products and the distance calculation of feature items in vector space are studied. The system block diagram is shown in Figure 2.

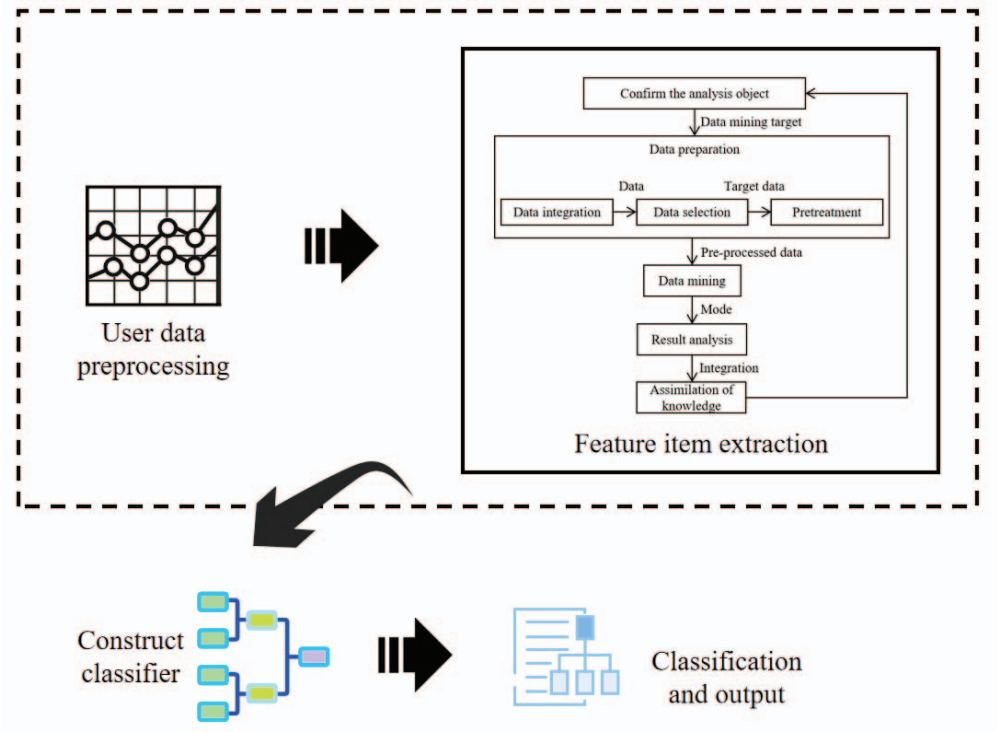

Figure 2. E-business precision marketing system

In KNN algorithm, Hamming distance can be used to measure the distance between sample data. After coding the features of users, the Hamming distance between codes can be calculated, and the magnitude of Hamming distance can represent the differences and similarities among users.

The amount of information is inversely proportional to the frequency of information, and the amount of information contained in different information can be accumulated. Set sample set:

$$S = \{s_1, s_2, \ldots, s_m\} \quad (1)$$

The sample categories are:

$$C = \{c_1, c_2, \ldots, c_k\} \quad (2)$$

Then the calculation formula of the sample information entropy is as follows:

$$H(S) = \sum_{j=1}^{k} \sum_{i=1}^{m} p(s_{ij}) \log \frac{1}{p(s_{ij})} = -\sum_{j=1}^{k} \sum_{i=1}^{m} p(s_{ij}) \log p(s_{ij}) \quad (3)$$

Where $p(s_{ij})$ represents the probability that the sample point in the sample set belongs to the category c_j. Let the sample attribute set:

$$A = \{a_1, a_2, \ldots, a_t\}^T \quad (4)$$

There are t different values. The sample set is divided into t sample subsets according to the attribute A, and the information entropy of the sample set is as follows:

$$H(S|A) = -\sum_{j=1}^{k} \sum_{i=1}^{t} p(a_{ij}) \log p(a_{ij}) \quad (5)$$

Among them, $p(a_{ij})$ represents the probability that the sample point belongs to the category c_i when the attribute A is a_i.

User portraits need a lot of user data, and if there is not enough user data, the portraits created can't really depict real people. Using crawler technology, we can easily collect all kinds of visible information left by users on the website. This kind of data acquisition has the advantages of being fast and reliable. Before data collection, you should list the types of data you want to collect, and then divide different types into different modules, each of which has a code to record the corresponding data. When the user enters the website, all the behaviors of the user will be recorded to form the website log, thus forming the data set for constructing the user portrait. The construction process of user portrait is shown in Figure 3.

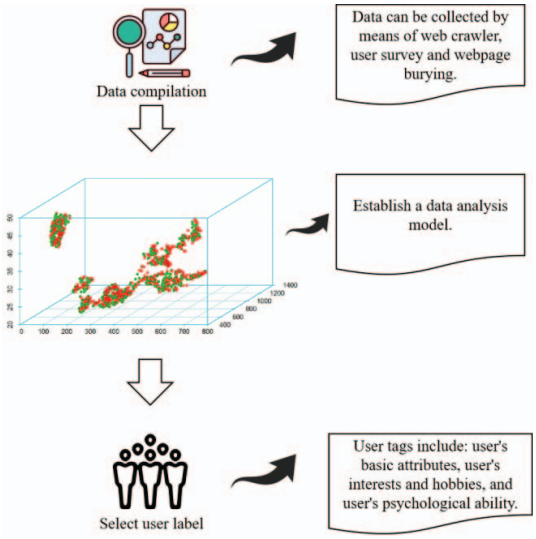

Figure 3. User portrait construction process

In the real society, there are many and complicated factors involved in the making of customers' purchasing decisions, and besides the qualitative factors of customers' purchasing characteristics themselves, the interrelation among various characteristics is obvious, which even has a joint effect on customers' purchasing decisions. At the same time, in the process of discovering more influencing factors of customer value with the help of data mining technology, the characteristics of data itself are relatively complex, and the influence mechanism of customer purchase characteristics on customer value will become more complex under the dual effects of qualitative and quantitative.

IV. Result Analysis and Discussion

Marketing activities of enterprises can't be separated from communication. Although traditional marketing also emphasizes keeping communication with customers, there are always various unnecessary intermediate links between enterprises and consumers due to the limitations of social conditions at that time. If consumers want to communicate with enterprises, they need to communicate with dealers first. The sellers feed back consumers' wishes to the next level. Through such layers of communication, consumers' wishes can't be timely and accurately conveyed to the enterprise level. Therefore, the marketing formulated by enterprises can't reach the level of accuracy. Communication theory requires enterprises to establish direct contact with consumers and fully and quickly understand the needs of consumers. The last consumption time of a customer refers to the time when the customer purchased the goods in the previous time. Theoretically, the closer a customer's previous consumption time is to the present time, the greater the customer's trust in the enterprise, which proves that the customer is a high-quality customer. If this kind of customer is marketed, the more likely it is to make a second purchase.

A rough observation of data sets shows that there are two main problems: the first is that there are a large number of missing values in some indicators in user attribute data and sales data. The second point is that there are refund users in the sales data. This part of user data is meaningless for the calculation of user value. If this part of data is retained, it will have a huge impact on the model results. For quantitative attributes such as the actual payment amount of customers, due to the complexity of the actual transaction situation, in order to avoid the loss of important information caused by the elimination of special data, the quantitative data with large difference in data distribution interval can be discretized. Data outlier removal processing is shown in Figure 4.

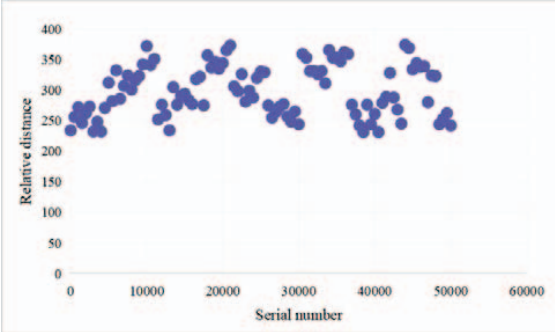

Figure 4. Data to remove outliers

Consumers tend to prefer to buy products that can get more customer value, so that they can be satisfied to the greatest extent. Compare the rate of return and accuracy of the algorithm in the prediction of consumers' consumption tendency on e-business platform, as shown in Figure 5 and Figure 6.

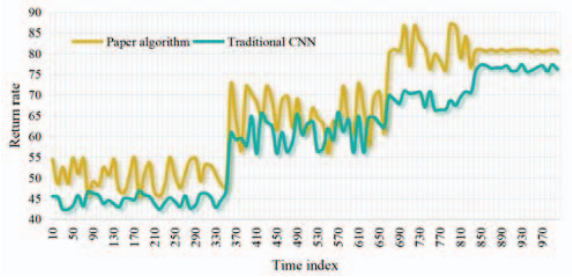

Figure 5. Comparison of returns predicted by propensity to consume

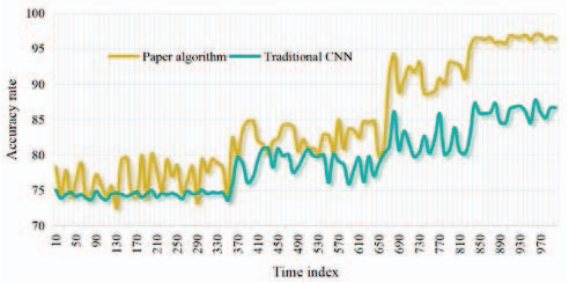

Figure 6. Comparison of prediction accuracy

From the subjective test results, it can be seen that this algorithm is more accurate in the prediction of consumers' consumption tendency on e-business platform, which is 23.25% higher than the traditional CNN algorithm.

Compared with the traditional data analysis and prediction methods, KNN algorithm has higher decision accuracy and efficiency, and is more intelligent. Accurate marketing requires not only the planning of the

decision-making level, but also the effective management of the whole activity by a huge team and the immediate implementation of the executive level. A perfect enterprise structure is not only easier for enterprise management, but also bound to maximize the value of precision marketing on the premise of saving enterprise costs. If an enterprise wants to attract consumers, it is necessary to provide them with more customers' value goods. For this reason, the enterprise can reduce the time cost of customers by reducing the time they spend buying goods, so as to achieve the purpose of increasing customers' value. Precision marketing is to adopt this idea. Enterprises can find their own target customers through precision marketing. Formulating marketing means and services according to the needs of target users can greatly reduce the time and energy spent by customers in purchasing goods and improve their customer-giving value.

V. Conclusions

With the continuous development and maturity of the e-business industry, more and more enterprises have started to stay in the e-business platform, hoping to further expand their profits through e-business channels. However, the entry of a large number of enterprises has led to greater competitive pressure among e-business. Under the influence of computer technology, e-business is developing rapidly, and the demand for precision marketing is constantly increasing. For a huge customer group, the traditional marketing method will have some defects, such as high input cost and uncontrollable return effect. Therefore, precision marketing has become a more reasonable marketing method. In the accurate marketing analysis of e-business products based on KNN algorithm, it is necessary to analyze from the perspective of model establishment, and then in the future marketing decisions, the accuracy of marketing can be effectively improved and marketing can be carried out accurately. For potential users, we should let them know about the store's commodity information as much as possible, and adopt appropriate price reduction and profit-making methods to encourage users to further increase their consumption. Through continuous targeted publicity and promotion, this kind of customers will eventually be transformed into high-value users. Because e-business involves a wide range of areas, the ways to achieve precise marketing are also very rich, and some of the contents involved in trade secrets are not easy to obtain, so there are still many places in this article that are not comprehensive and in-depth, and I hope that a more in-depth analysis can be achieved in the follow-up.

References

[1] Prasetya C. Sistem Rekomendasi Pada E-Business Menggunakan K-Nearest Neighbor. Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 4, no. 3, pp. 194, 2017.

[2] Sahabi Y A, Che H, Busalim A H. Influence of e-WOM engagement on consumer purchase intention in social commerce. Journal of Services Marketing, vol. 32, no. 4, pp. 493-504, 2018.

[3] Kim Y, Peterson R A. A Meta-analysis of Online Trust Relationships in E-business. Journal of Interactive Marketing, vol. 38, no. 5, pp. 44-54, 2017.

[4] Barrutia J M, Paredes M R, Echebarria C. Value co-creation in e-business contexts: does product type matter? European Journal of Marketing, vol. 50, no. 3-4, pp. 442-463, 2016.

[5] Paredes, Mario, R, et al. Value co-creation in e-business contexts: does product type matter? European journal of marketing, vol. 50, no. 3/4, pp. 442-463, 2016.

[6] Mcguigan L. Direct Marketing and the Productive Capacity of Commercial Television: T-commerce, Advanced Advertising, and the Audience Product. Television & New Media, vol. 16, no. 2, pp. 196-214, 2015.

[7] Pei Z, Yan R. Cooperative behavior and information sharing in the e-business age. Industrial marketing management, vol. 76, no. 1, pp. 12-22, 2019.

[8] Currim I S, Mintz O, Siddarth S. Information Accessed or Information Available? The Impact on Consumer Preferences Inferred at a Durable Product E-business Website. Journal of Interactive Marketing, vol. 29, no. 2, pp. 11-25, 2015.

[9] Dolores Gallego M, Bueno S, Lopez-Jimenez D. Impact of B2C e-business codes of conduct on sales volume: lessons from the Spanish perspective. Journal of Business & Industrial Marketing, vol. 31, no. 3, pp. 381-392, 2016.

[10] Micu A, Micu A E, Geru M, et al. The Impact of Artificial Intelligence Use on the E-Business in Romania. Amfiteatru Economic, vol. 23, no. 56, pp. 137-154, 2021.

[11] Javier A Sánchez-Torres, Juarez F. Modelling SME e-business with IMAES. Journal of Business & Industrial Marketing, vol. 34, no. 1, pp. 137-149, 2019.

[12] Sharma G, Wang L. The effects of online service quality of e-business Websites on user satisfaction. Electronic Library, vol. 33, no. 3, pp. 468-485, 2015.

[13] Escobar-Rodriguez T, Bonson-Fernandez R. Analysing online purchase intention in Spain: fashion e-business. Information Systems and e-Business Management, vol. 15, no. 3, pp. 1-24, 2017.