Cross-border E-commerce Course Construction Based on Data Mining Algorithm

Jianghua Deng1,*, Haohua Qing1

1Guangzhou College of Technology and Business, Guangzhou, Guangdong, 510850, China

*Corresponding author E-mail: djhua321@gzgs.edu.cn

Abstract. In recent years, Cross-border(CB) e-commerce(EC) has developed rapidly in our country, and the government has also introduced a series of measures to promote the development of CB EC. In the era of big data, recommendation systems are widely used in CB EC platforms, and more and more attention is paid to its personalized recommendation technology. Based on this, this paper studies the application of data mining algorithms in the construction of CB EC courses, which has certain guiding significance for the current market and enterprise development. By taking two CB EC companies as the research object, this paper uses data mining algorithms to mine and analyze the relevant data of these two EC companies, and find that a platform with a data mining system has more advantages in researching customer shopping preferences; comparing people’s satisfaction with EC companies with data mining systems and traditional companies, 83% of the participants in the experimental group were satisfied, of which 60% were very satisfied, while only 59% in the control group of people are satisfied. This shows that data mining algorithms are of great significance to the development of CB EC, and it also points out a new direction for the construction of CB EC courses, such as CB EC based on data mining algorithms. On the one hand, this will also play a positive role in the future development of CB EC.

Keywords: Data Algorithm, Cross-border e-commerce, Big Data, Data Mining Technology, Curriculum Construction

1. Introduction

In recent years, with the exponential development of science and technology, EC has great prospects and rapid development. In this process, a large amount of data circulates in the entire online world. How to effectively transform these data into EC" Kinetic energy", increase the potential value of the enterprise [1]. How to effectively organize and utilize the massive amount of EC information so that customers can accurately find the information they need in a complex ocean of data is a difficult problem and an opportunity for all enterprises, and it is also an urgent issue for EC websites[2]. And
data mining technology provides an effective solution to this problem. Data mining technology is a new commercial information processing technology. In the EC system, data mining technology is used to collect a large amount of information for enterprises and analyze them, and find information that is beneficial to themselves. Enterprises make decisions or judgments based on the corresponding information. The innovation and management capabilities of the enterprise are all improved at one time, thereby improving the efficiency of customer service [3-4].

Combining data mining algorithms with CB EC is not as simple as we imagined. Both of them are products of emerging developments in the new era. They must be used in a clever combination of their respective characteristics. Zhu Lizhi and Zhu Wuuzhe believe that most of the current data mining algorithms can only deal with data streams of one attribute, but there are few algorithms dealing with data streams of mixed attributes [5-6]. However, data mining algorithms provide great convenience for CB EC [7]. Niu Danqian believes that the development of CB EC has a good future, and the new policies on CB EC are also very conducive to its development. At the same time, he expressed the phenomenon that some colleges and universities have developed research on the construction of CB EC courses [8]. When planning the data mining algorithm into the business field, we can extract the key data that assists business decision-making from it, which is the focus of people’s attention [9]. These efficient decisions are an issue for every business structure. If there is room for progress, being able to take advantage of this opportunity and grasping this will put the company in an advantageous position in the fierce market competition [10], which is the basis for ensuring that it will not be eliminated by the times. Wang Xuerong proposed a dynamic forecasting model of export product sales based on CB EC controllable correlation big data driven by "Internet + foreign trade", and conducted an application experiment [11]. Yang Fang believes that with the rapid development of the Internet economy, CB EC marketing has become an important mode of trade for import and export companies, and data mining algorithms play an important role in the development of CB EC [12]. In the context of the rapid development of data mining and the continuous deepening of CB EC applications, data mining is not a simple process. A correct data mining model is essential, and appropriate data mining algorithms are the mainstay. After integrating these with data, we can extract some valuable business products from it, such as efficient marketing models, rationalized business rules, highly profitable customer distribution models, or the degree of correlation between products waiting is a business opportunity.

In business practice, we can usually use language, text, pictures, etc. to express these virtual products, and then hand them over to an expert team to process and analyze, make decisions, and provide constructive information for business activities or business planning. Suggest and test it in business practice. This is one aspect of the role. On the other hand, it can also analyze user behaviors, group and cluster some customers with similar behaviors, so as to obtain their public characteristics, to help themselves understand customers, understand their potential needs, and provide more information accordingly. Perfect service, which also provides great convenience for the rapid development of CB EC. Based on this theory, this paper takes two EC companies as the research object, uses data mining technology to conduct data mining analysis on them, and summarizes the development law of EC courses, which has certain positive significance for the development of EC companies. At the same time, it has greatly promoted the research on the construction of CB EC courses.

2. Method

2.1 Data Mining

Data mining is not mining in the traditional sense. It aims to extract information that is valuable to people from redundant data. It is not a blind extraction. These valuable information or laws are provided for individuals or enterprises. For example, in today's society, it is widely used in item recommendation on online shopping platforms, as well as personalized web pages on some websites. Tracing its source is inseparable from data warehouse technology, a technology used to process huge business data, looking for the hidden regularity behind it, and at the same time modeling it to provide a
reference technology for decision makers, so as to be in the market to take the upper hand and gain an advantageous position. Compared with traditional data analysis, it has no hypothesis, that is, it extracts and discovers useful information under the premise of no hypothesis.

2.2 E-commerce
EC refers to the use of modern information network electronic technology as the main means. It is also a form of business centered on traditional commodity exchange. It can also be understood as a traditional EC on the mobile Internet, enterprise internal question-and-answer network and enterprise value-added network. The way of trading is to directly carry out commodity trading operations and commercial activities that provide related information services, which is also the key to realizing traditional electronic commercial activities, but its business is more electronic, networked, and informatized. EC is also usually defined as broadly referring to a wide range of electronic commerce trade operations around the world. Under the global network economy environment where the global Internet is rapidly opening up, buyers and sellers conduct various commerce and trade activities online without meeting to realize consumers' online shopping, various online transactions between merchants and online EC payments by merchants, as well as various EC activities, transaction agency activities, financial service activities, and various comprehensive financial services related to transactions. EC activities refer to EC activities that use modern microcomputer computing technology and modern network communication information technology. Governments, scholars, and business analysts at all levels in various countries have given many different industry definitions according to their market position and their different angles and degrees of attention to participation in China's electronic information business industry. EC can be divided into: abc, b2b, b2c, c2c, b2m, m2c, b2a, c2aO and other modes.

2.3 Construction of Cross-border E-commerce Courses
Because many companies of different sizes are currently in urgent need of CB EC and other compound talents, in order to develop foreign trade business and seek overseas markets. Therefore, in order to conform to the development of the times, major universities have also opened CB EC courses. The general course content is to introduce students to the current development trend of CB EC and the operating rules of several popular mainstream platforms, so that students can understand the knowledge related to CB EC, and let students pass the usual courses Practice, and cultivate the ability to use the knowledge learned and multiple platforms to open up the international market.

2.4 Data Mining Algorithms and E-commerce
In the EC marketing activities of Chinese CB enterprises, the product content, title, reward solutions, services and other aspects of the website are very likely to attract a large number of customers. Due to the numerous websites of Chinese electronic enterprise business companies, customers can usually easily switch between different websites. Therefore, the website design of Chinese electronic enterprise business companies should be able to analyze and mine the information content of each customer's visit, by analyzing the different browsing of customers habit behavior, so as to fully understand each customer's browsing loyalty, preferences and different needs, and quickly and automatically adjust the corresponding page to meet the different needs of each customer. For example, Amazon, Express, etc., use data to analyze customer needs, quickly locate products, and prioritize products that meet their needs to facilitate customers' selection.

2.5 Data Mining Algorithm
Generally, data mining integrates basic theories and applied technologies in various professional fields such as various databases, artificial intelligence, robotics learning, and statistics. The typical enterprise data mining analysis methods mainly include data classification, correlation point and rule structure analysis, cluster analysis and outlier rule analysis. Among them, clustering analysis needs to subdivide
several basic steps. The first is to calculate the similarity. If the cosine angle similarity is used, suppose that there are two books $i$ and $j$ in a system $j$, and the interactive information is $u$ and $r$, then the weight calculation formula of cosine angle similarity is expressed as:

$$sim(i, j) = \cos(i, j) = \frac{\sum_{u \in U} \left( R_{i,u} \cdot R_{j,u} \right)}{\sqrt{\sum_{u \in U} R_{i,u}^2}}$$

(1)

Sometimes group parameters are also used, and these parameters in the group process can usually be considered as a good design solution for process simulation laboratory. The group variance refers to the group variance between the mean and a single point. It is mainly used to reflect the overall development of a social group. The definition can be described as follows:

$$\mu = \frac{1}{N} \sum_{i=0}^{N} x_i$$

(2)

Or it can sometimes be calculated with the help of differential equations:

$$\frac{dx}{dt} = x k [\mu(k,s) - \mu(s,s)], k = 1,2,\ldots,n$$

(3)

Where $x_k$ is the proportion of strategy $k$ used in a population, $\mu(k,s)$ represents the payment when strategy $k$ is adopted, $\mu(s,s)$ represents the average payment, and $k$ represents different strategies.

3. Experiment

According to the foregoing, in this chapter we will specifically apply data mining algorithm technology to CB EC. Through data analysis experiments, we will summarize and analyze the advantages of using data mining technology in CB EC compared to no application. We can also find that in the CB EC courses offered by many colleges and universities, the application of data mining algorithms in CB EC accounts for a large proportion of them. This also shows that data mining algorithms are used in CB EC. The importance of EC course construction.

3.1 Selection of Experimental Subjects

Take two EC companies of the same type as examples to analyze their sales performance. Among them, company A uses intelligent search.

3.2 Experimental Test Indicators

Within the selected time, we observe the attractiveness of similar products from both sides to customers, take the page stay time as the experimental data, and take a random questionnaire survey method to conduct questionnaire surveys of the people to investigate their business in the two modes. The degree of satisfaction of the operation, combined with feedback, to compare and analyze the effects of the two sets of experimental data.

3.3 Processing of Experimental Data

We use the statistical Bayesian formula to verify. Its significance is that the more events that support a certain attribute, the greater the probability that the attribute will be established. The formula is as follows:
\[ P(A|B) = \frac{P(B|A)P(A)}{\sum_j P(B|A_j)P(A_j)} \quad (4) \]

4. Result

4.1 Comparison of Data Mining Results of Two EC Companies

In the field investigation, regarding the actual impact of data mining, two EC companies with very similar scales and systems constitute the experimental group and the control group. The experimental group has formed a relatively complete data mining system, and the control group is a traditional company. The average length of stay of the same type of items on their respective online platforms in the same day is the experimental data. The results are shown in Table 1 and Figure 1.

**Table 1.** Average length of stay in various items

|               | Clothing | Book | Electronic product | Food | Furniture | Virtual goods |
|---------------|----------|------|--------------------|------|-----------|--------------|
| Experiment group | 23.5     | 15.5 | 39.7               | 19.3 | 54.3      | 8.6          |
| Control group  | 19.6     | 14.2 | 38.6               | 15.6 | 43.2      | 8.7          |

**Figure 1.** Average length of stay in various items

It can be seen from Table 1 and Figure 1 that the average length of stay of consumers in the experimental group in clothing is 23.5 seconds, the average length of stay in books is 15.5 seconds, and the average length of stay in electronic product pages is 39.7 seconds. The average stay in the purchase of physical objects was 19.3 seconds, the average stay in the furniture purchase page was 54.3 seconds, and the average stay in the purchase of virtual products was 8.6 seconds. The average length of stay of consumers in the control group was 19.6 seconds in clothing, 14.2 seconds in books, 38.6 seconds in electronic product shopping pages, and in-kind purchases. At 15.6 seconds, the average stay time on the furniture purchase page was 43.2 seconds, while the average stay time when
purchasing virtual products was 8.7 seconds. This shows that compared to traditional EC platforms, platforms with a data mining system have certain analysis and evaluation functions, and then interpret the data, can understand customer needs, and provide more suitable products with higher satisfaction. To customers.

4.2 People’s Satisfaction with EC Companies and Traditional Companies with Data Mining Systems

In the questionnaire survey, we conducted statistics on the public’s satisfaction with EC companies and traditional companies with data mining systems. The results are shown in Table 2 and Figure 2.

Table 2. The final comprehensive results of the two groups of people in degree of interest in art-derived literature

|                    | Very satisfied | Quite satisfied | Dissatisfied | Unclear |
|--------------------|----------------|-----------------|--------------|---------|
| Experiment group   | 60%            | 23%             | 12%          | 5%      |
| Control group      | 40%            | 19%             | 34%          | 7%      |

According to the data in Table 2 and Figure 2, 60% of the people in the experimental group are very satisfied, 23% are quite satisfied, 12% are not satisfied, and 5% are not clear. In the control group, 40% of the people are very satisfied, 19% are quite satisfied, 34% are not satisfied, and 7% are not clear. This shows that compared with the traditional system, people prefer this kind of human and intelligent shopping experience. Through the mining and analysis of background data, they can compare products efficiently and are more popular with consumers.

5. Conclusion

It can be said that with the rapid development of modern technology, EC is an inevitable product of the high-speed operation of future business. Under the influence of globalization, various fields should strengthen the corresponding network construction. This is a clear way of development to quickly and effectively promote the development of enterprise informatization, equip with corresponding security system, and create a good EC. Every company should pay attention to its development atmosphere. At the same time, in the area of network information, there are two points that we must break through. One is to improve the ability of network data mining and multimedia data mining, and the other is to solve existing problems, data quality, data security and confidentiality. And the integration of data mining with other commercial software are also problems that we need to overcome. Utilize modern
information technologies such as data warehouse and data mining to give full play to the unique advantages of enterprises, promote management innovation and technological innovation, and make enterprises invincible in the trend of EC. It is necessary to truly realize the development process and development of business activities of electronic enterprises in our country. Among them, service electronization, digitization and business networking are undoubtedly an inevitable technological achievement of the rapid development of China's modern business information era, and it is undoubtedly an inevitable choice for the future innovation mode of electronic business service operation. Business data mining analysis based on digital EC technology will be a very large and promising business field in the future. Research on the construction of CB EC courses based on data mining algorithms is becoming more and more popular among colleges and universities. We think colleges and universities should pay special attention to the theoretical and practical abilities of students while cultivating CB EC talents. The combination of this should not only teach students the knowledge in this area, but also give them more opportunities to exercise, encourage students to actively participate in extracurricular practice, boldly communicate and communicate, and allow students to conduct business operations as a qualified professional, to cultivate students' practical business ability and cultivate excellent talents for CB EC, which should be the goal of a qualified CB EC course.

Acknowledgments

Construction of theory and practice teaching of cross border EC Foundation Course (Project No: 20180227007); Practice teaching base of EC Industrial Park in MeiDong.

References

[1] LV Zhifeng. Research on efficient data mining algorithm for large scale data sets [J]. Informatization construction, 2016, 25 (006): 111-112.
[2] Li Yan. Group user access control model for mobile EC based on data mining algorithm [J]. Modern electronic technology, 2020, 043 (004): 153-156.
[3] Luo yangqianzi. Research on efficient data mining algorithm for large scale data sets [J]. Journal of Hunan City University: Natural Science Edition, 2016, 25 (4): 111-112.
[4] Siyu. Cross border EC report released in the first half of the year is still the first choice for users [J]. China food, 2018, no.749 (13): 112-113.
[5] Zhu Lizhi, Zhu wuzhe. An algorithm for clustering mixed attribute data streams [J]. Computing technology and automation, 2016, 035 (002): 34-37.
[6] Chen Shiquan. Building Shiyan EC college service regional marketing strategy [J]. Education, 2016 , 23(11): 00219-00219.
[7] Shou en Dai. A survey on the demand for high-end CB EC talents in the cultivation of Applied Undergraduate Business English entrepreneurship and innovation talents [J]. Scenic spots, 2018, no.352 (09): 122-123.
[8] Niu danqian, Zhang Baoming. Analysis on the influence trend and Countermeasures of cross border EC new policies on import retail platform [J]. EC, 2017, 1 (006): 22-23.
[9] Sun Liang. Research on efficient data mining algorithms for large data sets [J]. automation and instrumentation, 2016, No.197 (03): 192-193.
[10] Zhang Zhen. Research on Data Mining Algorithm in cloud computing environment [J]. Network security technology and application, 2019, 221 (05): 61-62.
[11] Wang Xuerong, Wan NianHong. A dynamic prediction model of export product sales based on CB EC controllable relevance big data [J]. Computer applications, 2017 (04): 130-135 + 142.
[12] Yang Fang, Zhong Jiamin, Zheng Zhuo, et al. Measurement of the application level of cross border EC marketing websites of import and export enterprises [J]. Consumer guide, 2018, 14 (007): 81-82.