

## Research Article

# Application of Big Data Technology in the Impact of Tourism E-Commerce on Tourism Planning

Heqing Zhang , Tingting Guo, and Xiaobo Su

Management (Tourism) School of Guangzhou University, Guangzhou, Guangdong 510006, China

Correspondence should be addressed to Heqing Zhang; lyzqh8007@gzhu.edu.cn

Received 8 March 2021; Revised 21 April 2021; Accepted 17 May 2021; Published 28 May 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Heqing Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the improvement of the material standard of living, the demand of the people for spiritual culture continues to increase. In terms of tourism, people have gradually shifted from simple tourism needs to integrated tourism needs. Tourism has become an effective way for people to expand their horizons and enrich their spiritual world. Tourism is one of the first industries to apply network technology. After long-term exploration and innovation, tourism e-commerce has developed rapidly. Coupled with the advent of the era of big data (EBD), the concept of customized tourism has gradually entered people's lives. This paper mainly introduces the research on the impact of e-commerce tourism on custom tourism in the EBD and intends to provide some ideas and directions for the good development of custom tourism. This paper proposes a research strategy on the impact of tourism e-commerce on customized tourism in the EBD, including related theoretical research methods, random forest algorithms, support vector machine classification algorithms, and Bayesian estimation algorithms, which are used to customize tourism e-commerce in the EBD, and research experiment on the impact of tourism. The experimental results show that 79.84% of customers are willing to purchase related products again after experiencing travel customization services using big data technology. By using the various characteristics of tourism big data for data mining and classification, it provides users with personalized travel search services. At the same time, big data technology can provide basic technical support for customized tourism development, which shows that it can also provide customized services for users.

## 1. Introduction

With the development of the tourism industry and the upgrading of the consumer concept of the Chinese people, consumers have begun to seek a deeper tourism experience and the quality of tourism has gradually improved. It no longer considers popular tourism as the first travel option and pays more attention to the personalization and uniqueness of tourism. Many professionals and investors are looking for new tourism models, customized tourism is in line with market growth trends, and its scale is expanding day by day, making it a significant force in the tourism industry. The intelligent age and the rapid popularization of the Internet and the combination of the travel industry and

e-commerce have attracted many entrepreneurs to open up the online travel market. New business models, travel service platforms, and apps continue to emerge, such as Qunar, Ctrip, and Fliggy. The tourism e-commerce platform occupies an important position in the tourism market, driving the rapid development of personalized tourism e-commerce.

Big data is a kind of development based on information and communication technology. It is based on heterogeneous multisource data sets, supported by advanced information storage, analysis, and processing technology, and takes big data thinking as an opportunity. With the exponential growth of data volume and the generation of unstructured data, the data information that needs to be processed has greatly exceeded the ability

of traditional database tools and software to collect, store, manage, and analyze, and various new technologies for processing large amounts of data should be applied. The growth of data volume and the development of data technology have made it possible for big data to transform social life. Tourism, finance, medical care, and education have begun to use big data to analyze and solve problems and achieve better results than traditional ones. The e-commerce tourism chain relies on Internet technology to combine the resources accumulated by many carriers, hotels, and other suppliers into a huge database. Through excavating and analyzing the unique needs of consumers, high-tech data matches response resources and is committed to providing consumers with customized tourism products and services.

Kumoro and Hasanah found that, due to the rapid development of mobile social networks, mobile big data plays an important role in providing mobile social users with various mobile services. However, due to the inherent properties of mobile big data, the current MSN faces the challenge of providing satisfactory big data services to mobile social users. Therefore, Su proposed a novel framework for delivering mobile big data through content-centric mobile social networks. First, it studies the characteristics and challenges of mobile big data and then introduces a content-centric network architecture for transmitting mobile big data in MSN, where each data is composed of interest packets and data packets; next, by defining the priority of interest packets and data packets, how to choose a proxy node to forward interest packets and relay nodes to send data packets is given. This research step is relatively cumbersome and is not conducive to popularization in practice [1]. Liu et al. found that the use of Internet technology has become a new way to provide tourism services, which is also a new direction for the development of China's tourism industry. With the continuous improvement of technology, management, brand, and other factors, the development of China's tourism e-commerce has gradually accelerated. This research is theoretically strong but lacks experimental data support [2]. Wang believes that, with the continuous development of Internet technology, e-tourism e-commerce sites have also achieved rapid growth, with great growth potential. However, China's e-tourism sector is still growing slowly and is facing many problems. Therefore, Wang designed an e-commerce travel system based on service-oriented architecture (SOA) and B/S architecture in response to problems in the development of e-commerce tourism and based on the service-oriented design idea. Wang takes tourism companies as its main research object and relies on SQLService technology to conduct research. This method can be used for reference, but it is a bit one-sided [3].

The innovations of this paper are as follows: (1) proposing the use of Bayesian estimation algorithm for tourism big data analysis; (2) proposing the structure design of tourism big data mining; (3) carrying out the tourism e-commerce under the big data era custom travel design.

## 2. Strategies on the Impact of Tourism E-Commerce on Customized Tourism in the Big Data Era

### 2.1. Theoretical Research Method

*2.1.1. Literature Analysis Method.* Obtain printed material through bibliography retrieval and Internet reading from a large number of bibliographic materials. Learn from domestic and foreign big data technologies, cloud platforms, tourism websites, e-commerce tourism, custom tourism, and other pieces of related literature, to adapt e-commerce tourism data. In-depth analysis and impact summary will provide theoretical guidance for this study [4].

*2.1.2. Field Trip Method.* In order to truly understand the development status of customized tourism in tourism e-commerce in the EBD, this paper went to relevant tourism e-commerce companies to conduct on-site inspections to learn about customized tourism related information on tourism websites, conduct research on major tourism e-commerce platforms, and collect relevant information on-site. And then after analysis and induction, the obtained data will be used for in-depth analysis of the research topic [5].

*2.1.3. Interview Method.* Choose major e-commerce sites, conduct structured interviews with selected travel site managers and employees, collect and organize data, and provide realistic guidance and research base. Structured interviews, also called standard interviews or closed-ended interviews, mean that the researcher asks questions to the respondent in a consistent manner according to a predetermined outline and the respondent chooses from alternative answers, which is in fact a closed verbal questionnaire. The advantages are the high controllability of the research, the high response rate, the strong structure, and the easy quantification [6].

### 2.2. Related Algorithms

*2.2.1. Clustering Algorithm.* The grid-based method quantifies the object space into a limited number of units, forming a grid structure [7, 8]. All clustering operations are performed on this grid structure. The main advantage of this method is its fast processing speed, and its processing time is independent of the number of data objects and only related to the number of units in each dimension in the quantization space [9, 10].

There are a series of classifiers  $h_1(x), h_2(x), \dots, h_k(x)$ :

$$F(x, y) = \frac{1}{K} \sum_k l\left(jx - \frac{Y}{d_1(x)}\right) - \frac{1}{K} (h(x) - j), \quad (1)$$

where  $Y$  is the clustering function value [9].

$$p(c | I, x) = \frac{1}{T} \sum_{t=1}^T p_t(c | I, x). \quad (2)$$

Each tree must be trained on different random data sets and a group of randomly selected split candidate parameters  $\phi = (\theta, \gamma)$  ( $\theta$  is the feature parameter;  $\gamma$  is the threshold) [11].

Divide instance  $q = \{(I, x)\}$  into two left and right subsets  $q_l(\phi)$  and  $q_r(\phi)$ :

$$q_l(\phi) = \{(I, x) \mid f_\theta(I, x) < \gamma\}, \quad (3)$$

$$q_r(\phi) = \frac{q}{q_l(\phi)}. \quad (4)$$

Calculate the maximum obtained information given by  $\varphi$ :

$$\begin{aligned} \phi^\Delta &= \arg \max_\phi G(\phi), \\ G(\phi) &= H(\phi) - \sum_{s \in \{l, r\}} \frac{|q_s(\phi)|}{q} H(q_s(\phi)). \end{aligned} \quad (5)$$

**2.2.2. Support Vector Machine Classification Algorithm.** Assume that the two-dimensional linear discrete function is expressed as

$$f(x) = \langle w, x \rangle + b. \quad (6)$$

The normalized form of the equation of the optimal classification line  $L$  is

$$\langle w, x \rangle + b = 0. \quad (7)$$

We can express the classifier as follows:

$$y_i(\langle w, x_i \rangle + b) \geq 1, \quad i = 1, 2, \dots, n, \quad (8)$$

$$\begin{aligned} \min J(w) &= \frac{\|w\|^2}{2}, \\ L(w, b, a) &= \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i (y_i (w x_i + b) - 1) a_i. \end{aligned} \quad (9)$$

The Lagrangian function can be expressed as

$$\frac{\partial L}{\partial w} = 0, \quad \text{so : } w = \sum_{i=1}^n a_i y_i x_i, \quad (10)$$

$$\frac{\partial L}{\partial b} = 0, \quad \text{so : } \sum_{i=1}^n a_i y_i = 0. \quad (11)$$

Finally, the optimization problem of the function is obtained:

$$W(a)_{\min} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j a_i a_j (x_i x_j) - \sum_{j=1}^n a_j. \quad (12)$$

Get the optimal solution  $a' = (a'_1, a'_2, \dots, a'_h)$ , and then the modulus of the optimal classifier  $w'$  has the following formula:

$$\|w'\|^2 = 2w(a') - \sum_{SV} a'_i a'_j (x_i x_j) y_i y_j. \quad (13)$$

The finally obtained tourism data classification function has the following expression:

$$f(x) = \text{sgn} \left( \sum_{SV}^n y_i a'_i (x_i x) + b \right). \quad (14)$$

**2.2.3. Bayesian Estimation Algorithm.** Bayesian network is a probabilistic graph model. When learning the Bayesian network structure, it is impossible to evaluate the scores of all networks in the space. You can first use the search algorithm to search for a specific possible space based on the rating function for structure selection. Commonly used search algorithms include hill climbing algorithm, taboo search, and greedy search [12, 13].

This paper uses the most commonly used hill climbing algorithm. The basic idea is to start from the initial model, get a series of candidate models through search patterns, and then calculate the results based on the candidate models. Compare the scores of the current model, select the model with the highest score as the initial model for the next search, continue the search or stop the search, and return to the current model [14]. The scoring function uses three functions (k2, BIC, and BDe) for comparison and analysis [15].

In the complete model learning of the data set, maximum likelihood estimation and Bayesian estimation are two commonly used parameter learning methods. The assumptions of these two methods are that the data sample set needs to be independent and identically distributed; that is, for a data set  $B$  containing  $n$  variables and containing  $x$  samples, it can be expressed as

$$b_i = \{b_{i1}, b_{i2}, \dots, b_{im}\}, \quad i = 1, 2, \dots, x. \quad (15)$$

Then each sample in data set  $B$  satisfies the following two conditions. One is that each sample in  $B$  is independent of each other after a given parameter  $\alpha$ :

$$p(B | \alpha) = \prod_{i=1}^x p(b_i | \alpha). \quad (16)$$

The other is that each sample  $b_i$  has the same conditional probability distribution  $p(b_i | \alpha)$  [16].

(1) *Maximum Likelihood Estimation.* Maximum likelihood estimation is based on parameter estimation based on the likelihood function of data samples and parameters [17]. There is a formula:

$$\alpha^* = \arg \max_{\alpha} L(\alpha | B), \quad (17)$$

where  $L(\alpha | B)$  is the likelihood function of the parameter  $\alpha$  after the model is determined, and the logarithm of it can be obtained:

$$\begin{aligned} \log(\alpha | B) &= \log \prod_{i=1}^x p(b_i | B) \\ &= \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} x_{ijk} \log(\alpha_{ijk}). \end{aligned} \quad (18)$$

(2) *Bayesian Estimation*. Regarding the question of uncertainty, Bayesian estimation and maximum likelihood estimation have different views. Bayesian estimation believes that parameters are also random variables and need to be assigned specifically. Therefore, Bayesian is reasonable in terms of the estimated maximum probability. It is generally believed that the prior distribution of the parameter  $\alpha$  is the Dirichlet distribution, and it can be known from the coupling that the posterior distribution of the parameter  $\alpha$  is also the Dirichlet distribution [18, 19]. For a given data set  $D$ , combined with the decomposability of the Bayesian network, the estimated value may appear:

$$\begin{aligned}\alpha_{ijk}^* &= \int p(y_i = k | pa(y_i) = j, \alpha_{ijk}) p(\alpha_{ijk} | B) b\alpha_{ijk} \\ &= \int \alpha_{ijk} p(\alpha_{ijk} | B) b\alpha_{ijk}.\end{aligned}\quad (19)$$

Considering that  $P(\alpha | B)$  also obeys the Dirichlet distribution, the Bayesian estimate of parameter  $\alpha$  can be obtained as

$$\hat{\alpha}_{ijk} = \frac{x_{ijk} + \alpha_{ijk}}{\sum_{k=1}^{r_i} (x_{ijk} + \alpha_{ijk})}.\quad (20)$$

The method part of this paper uses the above method to study the impact of tourism e-commerce on customized tourism in the EBD. The specific process is shown in Table 1.

### 3. Experiment on the Impact of Tourism E-Commerce on Customized Tourism in the EBD

#### 3.1. Attractions Data Mining

*3.1.1. Data Management System*. In the recommendation system, the corresponding data needs to be retrieved from the database at each step, such as scenic spot information database, user registration database, and pattern database. The database management system is responsible for the operation, management, and maintenance of the data in the database [20, 21]. The database management system is actually the middle layer between the online recommendation application and the database, and a certain database management system (DBMS) software can usually be used directly [22].

*3.1.2. Data Preprocessing*. Data preprocessing is an indispensable stage in the data mining process, because most of the data in the real world are incomplete, noisy, and inconsistent, and the formats of these data are diverse. Data mining algorithms may only mine part of the data in the database. For this, we need to extract useful data. For semistructured data on the web, data preprocessing is even more important. The preprocessing of web usage data mainly includes data cleaning (Data Cleaning) and session identification (Session Identification). Due to the existence of client-side caching and proxy server-side caching, it is necessary to complete the path for the more demanding personalized recommendation service (Path Completion). If

the website does not use cookie technology or embedded session identification technology, user identification (User Identification) is required; finally, sequence identification is required for the user's access sequence [23].

*3.1.3. Mode Decomposition*. The data analysis basis of the generated pattern is the user purchase database and web server log. The pattern analyzer uses data mining technology to analyze web server log data to obtain relevant information; from the analysis of the user's purchase database, it can find similar patterns of associations between scenic spots when they are selected [24, 25].

#### 3.2. Customized Tourism Design in Tourism E-Commerce in the Big Data Era

##### 3.2.1. Design Principles

(1) *Adhere to the Principle of Demand-Oriented*. According to the analysis of Maslow's theory of the hierarchy of needs and the theoretical value of tourists, as well as the reasons for the rise of custom tourism, it is concluded that the design of custom tourism should pay attention to the needs of tourists. By enhancing communication with tourists, understand that tourists are in this adjustment. On the subject, elements of tourism (food, housing, transportation, travel, entertainment, and shopping), and psychological needs, specific plans are made on this basis to achieve the purpose of adaptation by tourists.

(2) *The Principle of Individualization*. The most fundamental purpose of tourists to choose custom tours is to get rid of the travel methods that cut cookies and look for a unique travel form that fits their heart. Therefore, when adapting tourism design, the uniqueness of custom design must be reflected to meet the individual needs of tourists.

(3) *The Principle of Precision*. In order to reflect the advantages of customized tourism, it is necessary not only to accurately subdivide customized tourism objects and accurately analyze the needs of tourists but also to reflect the principle of precision in the design from time to time, so that the design can be highly tailored to customer needs and accurately design the plan to realize customized travel.

(4) *The Principle of Tourist Participation*. The design of customized tours should focus on customer participation throughout the entire process and strengthen customer communication. At the beginning of the design, it is necessary to determine the customized needs and travel details of the tourists through communication; in the design process, it is necessary to communicate with the tourists whether the initial plan meets the expectations of the tourists and find inappropriate places to modify and improve; after the end of the tour, it is also necessary to strengthen the participation of tourists encourages and guide tourists to evaluate this trip. Only when tourists participate in the whole process, can the customized travel plan be perfected and recognized by tourists.

TABLE 1: Part of the technical process of the method in this paper.

Research strategy on the impact of tourism E-commerce on customized tourism in the big data era	2.1	Theoretical research methods	1	Document analysis
			2	Fieldwork
			3	Interview method
	2.2	Related algorithms	1	Random forest algorithm
			2	Support vector machine classification algorithm
			3	Bayesian estimation algorithm

### 3.2.2. The Realization of Customized Tourism in the EBD

(1) *Optimize the Source of Customers.* To optimize the source of customers, personalized travel companies must first provide excellent service and sophisticated products to retain tourists, that is, to improve the repurchase rate of old customers through good customer relationship management and on this basis through old customer friends circle to attract customers. Secondly, with the advent of the information age, new technologies such as the Internet, 5G, big data, and artificial intelligence must be fully utilized for the development of online businesses and the good use of the OTA platform. At the same time, as you grow online, do a good job combining online and offline articles. Finally, it actively develops business, develops its own resources, and accumulates customers by expanding service extensions, such as accumulating reputation through other member businesses, developing its own islands, and being a good supplier.

(2) *Enhance the Application Ability of New Technical Means.* With the development of science and technology, it is necessary to increase the application of new technical means in terms of obtaining tourist needs and integrating resources, strengthen the extraction and analysis capabilities of big data, and make full use of big data to complete the collection, summary, and analysis of tourist travel choices. For example, use big data to collect information on tourists' flights, railway trips, hotel dining, and other pieces of information to make preliminary positioning of tourists. It is also possible to obtain tourist preferences for tourist destinations and various elements in tourist activities through tourist photos posted in the circle of friends by tourists. The application of new technology is also reflected in the integration of resources by tourism companies. Through big data technology, the integration of resources can be accelerated to provide customers with more and more tailored customized services.

The method part of this paper uses the above method to study the impact of tourism e-commerce on customized tourism in the EBD. The specific process is shown in Table 2.

## 4. Impact of Tourism E-Commerce on Customized Tourism in the EBD

4.1. *Scoring Function Comparison Result.* For the Bayesian estimation algorithm, the three scoring functions of k2, BIC, and BDe are tested for different standard nodes, the accuracy of all nodes is averaged and variance, and the relevant results

are statistically sorted. The specific conditions are shown in Table 3 and Figure 1.

Comparing the overall accuracy rate of the graph and the accuracy rates of cases 1 and 2, the results of the BDe score correlation are better than k2 and k2 is better than BIC. In the case of 0, the accuracy rate is just the opposite, BIC is better than BDe, and BDe is better than k2. In terms of variability, BDe has the lowest variability, followed by BIC, and k2 has the highest variability. In fact, this is related to the sample size of the k2 rating function. When the number of nodes changes, the sample size will change, resulting in a change in prediction accuracy.

4.2. *Development Status of Tourism E-Commerce Platform in the EBD.* The Internet has become the most important way for tourists to choose customized tourism. Table 4 shows the status of my country's major tourism e-commerce platforms.

The downloads of China's major tourism e-commerce platforms that appeared in the search for the keyword "travel" on the mobile phone are shown in Figure 2 (take the Android system search results as an example).

### 4.3. Experimental Results

- (1) In order to study the development of customized tourism in the EBD, this paper uses data mining and other related technologies to classify, collect, and organize user data in the tourism e-commerce platform and divide the data into the main goals of customized tourism. This classification standard is based on tourists' specific needs and preferences for customized tourism (such as preferences for tourism elements).

The main target groups of customized tourism are shown in Table 5 and Figure 3.

It can be seen from the chart that tourists have good satisfaction after experiencing customized travel products designed based on big data mining and analysis. A total of 87.13% of customers indicated that they were in line with their needs, and 79.84% of customers were willing to purchase related products again.

In the sample of tourism e-commerce platform, tourists' demand for tourism elements when accepting customized tourism is shown in Table 6 and Figure 4.

As can be seen from the chart, when customizing tourism consumption, tourists have the largest demand for food and tourist attractions,

TABLE 2: Part of the technical process of the experiment in this paper.

Research experiment on the impact of tourism E-commerce on customized tourism in the EBD			
3.1	Data mining structure design	3.2	Customized tourism design in tourism e-commerce in the EBD
1	C/S structure	1	Feature construction
2	B/S structure	2	Design principles
		3	The realization of customized tourism in the EBD

TABLE 3: Comparison results of the three scoring functions.

Standard Scoring function	Mean			Variance		
	k2	BIC	BDe	k2	BIC	BDe
Overall accuracy (%)	81.46	83.07	82.58	32.21	34.17	36.52
The accuracy of the prediction result is 0 (%)	22.17	25.24	23.36	27.56	31.25	29.28
The accuracy of the prediction result is 1 (%)	90.42	92.15	93.62	19.43	21.08	24.37
The accuracy of the prediction result is 2 (%)	72.31	74.56	71.29	42.59	40.71	39.74

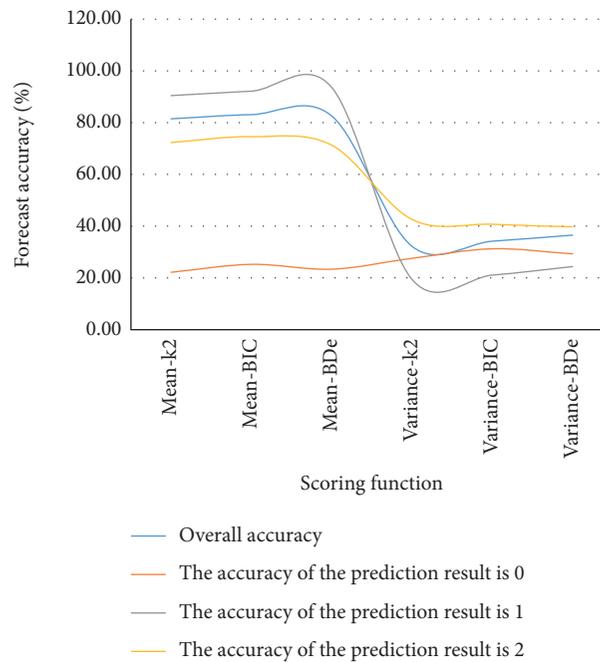


FIGURE 1: Comparison of the three scoring functions.

TABLE 4: Chinese major tourism e-commerce platforms.

Platform name	Release time (year)	Advantages
Qunar	2005	Qunar is the world's largest Chinese-language online travel website. Qunar provides consumers with real-time search of air tickets, hotels, venues, and holiday products, as well as group purchases of travel products and other travel information services, and provides online technology and mobile technology solutions.
Meituan	2010	In April 2018, Meituan travel and UnionPay International reached in-depth cooperation, which will deepen the exploration of technology, big data, and shopping experience to make travel and shopping more convenient. As China's leading e-commerce platform for life services, Meituan travel services cover more than 200 categories, such as catering, takeaway, fresh food retail, taxis, shared bicycles, hotel travel, movies, leisure, and entertainment.

TABLE 4: Continued.

Platform name	Release time (year)	Advantages
Fliggy	2016	Fliggy provides online booking services for more than 200,000 domestic and overseas hotels, homestays, inns, and serviced apartments, provides multiple domestic and outbound tourist routes, and provides one-stop services such as phone cards, overseas Wi-Fi, car rental, and visas; full coverage tickets for many paid attractions at home and abroad can be booked online with one-click reservations; online bookings for domestic and most mainstream international routes and all domestic train tickets are provided.
Tongcheng	2004	Tongcheng tourism is a national high-tech enterprise and the first batch of e-commerce demonstration enterprises by the Ministry of Commerce. Tongcheng tourism has been selected as one of the “top 20 China tourism groups” for three consecutive years. It ranked 9th in 2014 and 8th in 2015. It is one of the three major enterprise groups in China’s online travel industry. In the new ten years, the company takes “the first place in leisure tourism” as its strategic goal, uses scenic spot tickets as the entrance, and actively deploys peripheral tourism, long-term tourism, cruise tourism, and other business sectors.
Ctrip	1999	Ctrip provides hundreds of holiday product lines, including more than 20 holiday specialty stores, each of which has multiple lines of different product combinations. Guests can choose to depart from Beijing, Shanghai, Guangzhou, Shenzhen, Hangzhou, Chengdu, Shenyang, Nanjing, Qingdao, Xiamen, and Wuhan. In addition, tourism is divided into themed tourism, surrounding tourism, domestic tourism, Hong Kong, Macao, and Taiwan tourism, and other different forms of tourism. In each form, the content is detailed and rich, which can be suitable for customers to make different choices.

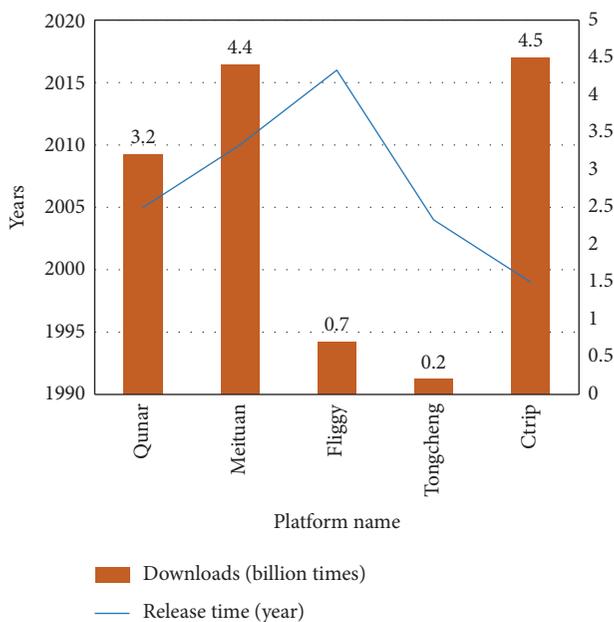


FIGURE 2: Downloads of Chinese major tourism e-commerce platforms.

TABLE 5: Main target groups for customized tourism.

Target population	Percentage (%)
Student	7.40
Corporate staff	19.50
Government and institution	17.80
Entrepreneur	18.20
Retirees	26.70
Freelance	10.40

accounting for 32.7% and 36.5%, respectively; accommodation 11.2%; shopping 7.9%; travel mode 7.1%; entertainment 4.6%, the least proportion of demand. The analysis shows that food, scenic spots, and accommodation are the most concerned issues for customers of customized tourism when choosing tourism elements. Therefore, these elements need to be the first consideration when making customized design. It can be seen that the related technologies generated in the EBD are conducive to customized tourism to be closer to the needs of tourists, improve customer satisfaction, and promote the development of tourism e-commerce.

- (2) The results of the comparison of the improved method applied to different tourism data sets are shown in Table 7 and Figure 5.

According to the data in the chart, it can be seen that when the support threshold is the same, the frequent set mining effects of different data sets are basically similar.

- (3) After using big data technology to design customized tourism products for tourists, conduct online surveys, return visits to customers, investigate and analyze customer satisfaction and feedback for the products, and organize relevant data, as shown in Table 8 and Figure 6.

It can be seen from Figure 2 that the current downloads of major tourism e-commerce platforms are all over 100 million. The market for tourism e-commerce platforms is vast and the demand for tourists is large, which is conducive to the development of customized tourism products.

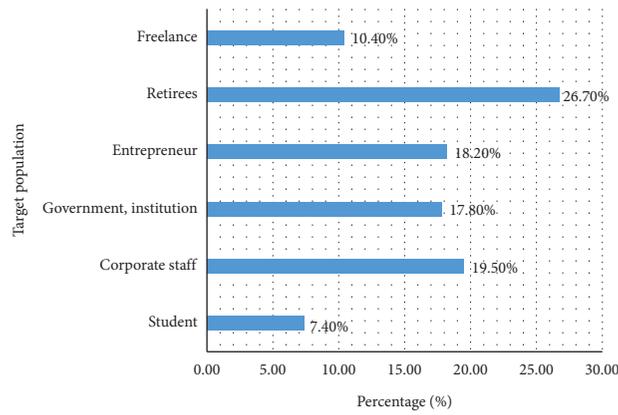


FIGURE 3: The main target groups of customized tourism.

TABLE 6: Tourists' demand for tourism elements.

Tourist demand	Eat (%)	Accommodation (%)	Travel mode (%)	Scenic spot (%)	Entertainment (%)	Shopping (%)
Percentage	32.7	11.2	7.1	36.5	4.6	7.9

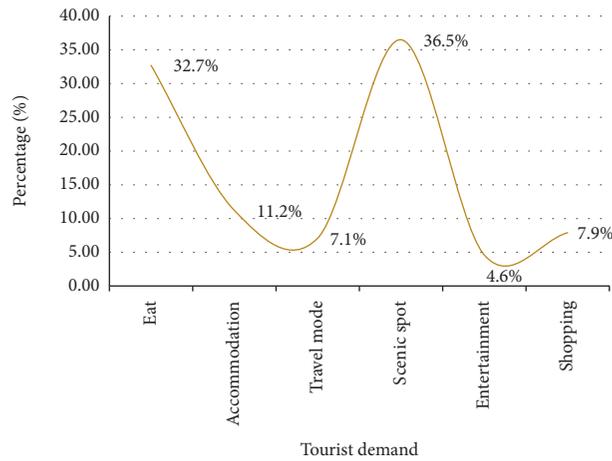


FIGURE 4: Tourists' demand for tourism elements.

TABLE 7: Frequent set mining results under different tourism data sets.

Data set	Support threshold			
	0.02	0.04	0.06	0.08
id1_Gender	112	41	25	32
id1_Language	173	59	39	43
id1_Releasedate	140	63	45	49
id1_songs_1	237	78	51	52
id2_Gender	161	46	62	64
id2_Language	342	117	59	53
id2_Releasedate	254	138	125	121
Id2_songs_2	328	121	113	92

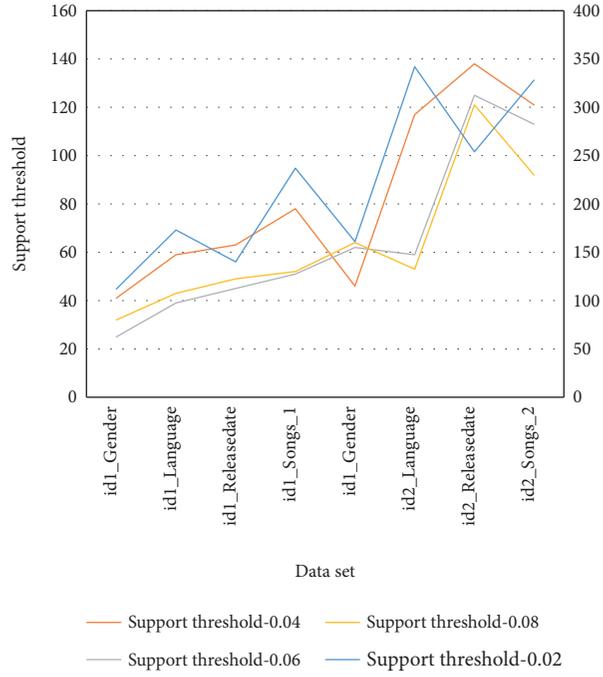


FIGURE 5: Frequent set mining results under different tourism data sets.

TABLE 8: Customer feedback.

Feedback	Percentage, yes (%)	Percentage, no (%)
1 Does the product meet your needs	87.13	12.87
2 Is the product easy to experience	64.26	35.74
3 Are you willing to experience related products again	79.84	20.16
4 Is the product cost-effective	22.48	77.52

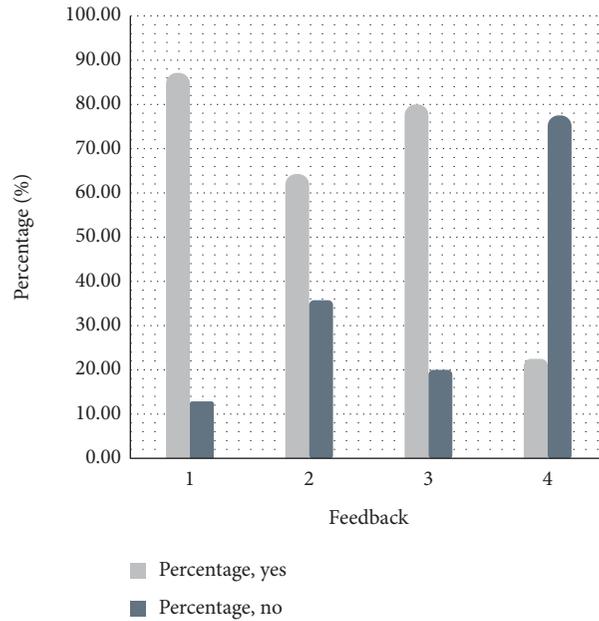


FIGURE 6: Customer feedback.

## 5. Conclusions

With the rapid development of social networks, huge amounts of travel data have been created on the Internet, resulting in the emergence of information overload. It takes a lot of energy for users to get effective information from it, which makes users increasingly in demand for effective travel information search. The study of tourism e-commerce big data has important theoretical and application significance. The research work in this paper is mainly focused on the customized tourism of tourism e-commerce in the EBD. By using the various characteristics of tourism big data for data mining and classification, it provides users with personalized travel search services. At the same time, big data technology can provide basic technical support for customized tourism development, which shows that it can also provide customized services for users.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Planning Office of Philosophy and Social Science Foundation of China: Research on Cultural Heritage Conservation and the Activation of South China Historical Trail (no. 19FSHB007)

## References

- [1] D. T. Kumoro and U. Hasanah, "Tinjauan desain interface website E-commerce wisata Mototravel.id menggunakan evaluasi heuristik," *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, vol. 2, no. 1, pp. 43–49, 2020.
- [2] Y. Liu, J. Wang, X. Shao, and J. Li, "The current situation and development trend of China's tourism e-commerce," *International Journal of Tourism Sciences*, vol. 18, no. 4, pp. 312–324, 2018.
- [3] D. Wang, "Research on design and realization of SOA-based tourism E-commerce system," *Revista de la Facultad de Ingenieria*, vol. 32, no. 13, pp. 732–736, 2017.
- [4] E. Baccarelli, N. Cordeschi, A. Mei et al., "Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: review, challenges, and a case study," *Computers & Chemical Engineering*, vol. 91, no. 2, pp. 182–194, 2016.
- [5] J. Xu, Q. Hou, C. Niu, Y. Wang, and Y. Xie, "Process optimization of the university-industry-research collaborative innovation from the perspective of knowledge management," *Cognitive Systems Research*, vol. 52, pp. 995–1003, 2018.
- [6] D. A. Mohan, "Big data analytics: recent achievements and new challenges," *International Journal of Computer Applications Technology and Research*, vol. 5, no. 7, pp. 460–464, 2016.
- [7] H. Stevens, "Big data, little data, no data: scholarship in the networked world," *Journal of the Association for Information Science & Technology*, vol. 67, no. 3, pp. 751–753, 2016.
- [8] M. Zaharia, R. S. Xin, P. Wendell et al., "Apache spark," *Communications of the ACM*, vol. 59, no. 11, pp. 56–65, 2016.
- [9] Z. Obermeyer and E. J. Emanuel, "Predicting the future—big data, machine learning, and clinical medicine," *New England Journal of Medicine*, vol. 375, no. 13, pp. 1216–1219, 2016.
- [10] D. Specht, "The data revolution: big data, open data, data infrastructures and their consequences," *Media Culture & Society*, vol. 37, no. 7, pp. 1110–1111, 2016.
- [11] B. S. Kim, B. G. Kang, S. H. Choi, and T. G. Kim, "Data modeling versus simulation modeling in the big data era: case study of a greenhouse control system," *Simulation*, vol. 93, no. 7, pp. 579–594, 2017.
- [12] J. Aikat, T. M. Carsey, K. Fecho et al., "Scientific training in the era of big data: a new pedagogy for graduate education," *Big Data*, vol. 5, no. 1, pp. 12–18, 2017.
- [13] M. Martinez, P. D. Nauta, and D. Sarno, "Real and apparent changes of organizational processes in the EBD analytics," *Studi Organizzativi*, vol. 2, pp. 91–107, 2017.
- [14] C. Wu, G. Xue, J. Li, K.-L. A. Yau, and J. Qadir, "Computational intelligence for Internet of things in the big data era (part II) [guest editorial]," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 22–23, 2020.
- [15] M. Kang and F. Ampornstira, "Research on data analysis of Chinese public accounting firms in the big data era," *Open Journal of Accounting*, vol. 10, no. 1, pp. 1–8, 2021.
- [16] D. Bertsimas, N. Kallus, and A. Hussain, "Inventory management in the era of big data," *Production and Operations Management*, vol. 25, no. 12, pp. 2006–2009, 2016.
- [17] T. Ermakova, A. Hohensee, I. Orlamünde, and B. Fabian, "Privacy-invading mechanisms in e-commerce—a case study on German tourism websites," *International Journal of Networking and Virtual Organisations*, vol. 20, no. 2, pp. 105–126, 2019.
- [18] T. J. Mapeshoane and S. Pather, "The adoption of E-commerce in the Lesotho tourism industry," *The Electronic Journal of Information Systems in Developing Countries*, vol. 75, no. 1, pp. 1–24, 2016.
- [19] J. Neidhardt, N. Rümmele, H. Werthner, and H. Werthner, "Predicting happiness: user interactions and sentiment analysis in an online travel forum," *Information Technology & Tourism*, vol. 17, no. 1, pp. 101–119, 2017.
- [20] X. Su, "Analysis of the development prospects of rural tourism e-commerce system under the background of big data," *Revista de la Facultad de Ingenieria*, vol. 32, no. 9, pp. 492–499, 2017.
- [21] R. Zhou, "Research on the framework of tourism E-commerce platform," *Boletin Tecnico/technical Bulletin*, vol. 55, no. 18, pp. 649–655, 2017.
- [22] S.-Y. Park and S.-H. Kim, "Examining the predictivity of Chinese inbound tourism demand on the stock return of each tourism industry," *Journal of Tourism Studies*, vol. 29, no. 4, pp. 153–177, 2017.
- [23] H. Yoon, "Tourist behavior between current and upcoming older adults: focus on the benefits sought from the travel experience and constraints," *Journal of Tourism & Leisure Research*, vol. 28, no. 3, pp. 5–21, 2016.
- [24] A. Mhadgut, G. Bhagwat, and P. Dhule, "Survey on customized tourism and travel recommendation system based on user-location vector approach," *International Journal of Computer Applications*, vol. 182, no. 14, pp. 16–18, 2018.
- [25] A. J. Chorin, F. Lu, R. N. Miller et al., "Sampling, feasibility, and priors in Bayesian estimation," *Discrete & Continuous Dynamical Systems—Series A (DCDS-A)*, vol. 36, no. 8, pp. 4227–4246, 2017.