

Internet User Consumption Behavior Based on Big Data

Yijie Yang¹

¹Statistics, North Carolina State University, Raleigh, 276957001, United States,

Abstract: With the vigorous development of the Internet economy, competition in the domestic market has become increasingly fierce. This research mainly discusses the consumption behavior of Internet users based on big data. Text preprocessing is required first after the target text has been obtained. Clean and delete content that is of no value or limited value in the text. First, define the Request network connection request. Secondly, write a function module for obtaining multiple commodity IDs. Traverse the full text to obtain the target content, which is often used to retrieve and replace the target text. Finally, the crawling of individual product information. According to the research product goals, big data technology is used to obtain data on two dimensions of product sales information and user experience information. Store the product sales specification text and product review text as divergent texts for the next stage of data cleaning to predict user consumption behavior. The function of the product, in the price range of 100-200 yuan, the user's attention is 27%. This research helps companies formulate precise marketing strategies.

1 Introduction

In contrast to traditional marketing methods, network marketing is a social activity management process that is carried out on the Internet and uses new ideas and methods for transactions between producers and consumers to meet both parties' needs. While online shopping changes people's living and consumption habits, it also subtly changes the marketing model of manufacturing enterprises.

Due to the various advantages of the Internet, such as strong penetration, wide-coverage, and low transaction costs, some companies have shifted their sales business from offline to online and become the direction of corporate development reform [1-2]. This trend has had an enormous impact on companies that are still marketing offline [3-4] ---judging from the telecommunications market's current operating conditions. The objective analysis done now is only for relatively successful businesses [5-6]. The information of failed users has not been thoroughly investigated. It cannot present a comprehensive understanding of the characteristics of users' consumption behavior and all aspects of users' needs. In the current development process, users will not fully support telecom operators' work [7-8]. The types of users cannot be reasonably divided, nor can they be adequately positioned to customers' value, and submit customers' consumer psychology for the detailed division. The differentiated marketing model cannot proceed smoothly and accurately grasp the consumer psychology of users. Indeed, the current production technology system cannot make full use of the database. It is also impossible to thoroughly track user changes [9-10].

In our country's operators' development process, the analysis of users' consumption behaviors extracts data from many aspects. For example, users need to conduct repeated investigations on the network by analysts in multiple dimensions using artificial energy for processing. After the current social development, the number of users is continuously increasing. The data that needs to be extracted is continuously increasing. The realms that require artificial power are becoming more and more complex and huge. Consequently, the number of staff required and the cost of operation will be increasing.

2 Online User Consumption Behavior

2.1 Big Data

Big data technology is developing fast due to open source and becoming more critical due to data growth. In the actual development of B2C e-commerce, all kinds of open source big data frameworks or secondary development tools are used as technical cornerstones to building big data. The system supports TB, PB, even EB-level data storage, data ETL, data computing, data mining, and artificial intelligence applications and realizes multi-dimensional data such as browsing, clicking, ordering, and consumer payment e-commerce platforms, including recording, storage, analysis, and processing.

The current development of big data will also bring great pressure to telecom and Internet merchants. Internet merchants must grasp the profit opportunity within the specified time. Get richer data resources. The acquired resources are extracted more deeply. Make deep use of background resources in the era of big data.

yyang54@ncsu.edu

2.2 Internet User Consumption

In an environment where the economic situation at home and abroad is deteriorating, and the networked development's model is in the ascendant, e-commerce and online shopping companies are undergoing a weak stage since their establishment. How to promote the transformation of e-commerce enterprises to network marketing and improve their operating capabilities under the network marketing mode are the critical issues for the future development of e-commerce enterprises. Based on network marketing theory, the current product network marketing strategies of e-commerce online shopping companies mainly have consumer demand, communication, convenience, and cost problems. These will have specific impacts on the results of e-commerce online shopping companies' network marketing. With the continuous development of Chinese country's social platforms, the Internet penetration rate is getting higher, and the e-commerce industry tremendous network marketing opportunities. Excavating and optimizing the original marketing strategy is an essential part of the e-commerce online shopping company's occupation of product online marketing. Therefore, e-commerce online shopping companies using social networking platforms and new online marketing methods to overcome existing marketing strategies' problems and increase their products' online sales focus on online shopping companies' product online marketing.

3 Experiments on the Acquisition of Consumption Behavior of Internet Users

3.1 User Query Information Text Preprocessing

Text preprocessing is firstly required when obtaining the target text. Clean and delete content that is of no value or extremely low value in the text. Avoid introducing word segmentation, word frequency statistics, text clustering classification, and even text modeling analysis, and carry out "noise reduction" processing to eliminate interference factors.

3.2 User Behavior Data Collection

Firstly, define the Request network connection request. Secondly, write a function module for obtaining multiple commodity IDs. Traverse the full text to obtain the target content, which is often used to retrieve and replace the target text. Finally, the crawling of individual product information. According to the research product goals, big data technology is used to obtain data on two dimensions of product sales information and user experience information. Store the product sales specification text and product review text as separated texts for the next stage of data cleaning to predict user consumption behavior. The commodity data set is shown in Table 1.

Table1. Product data set

Table Name	Remarks
Consumer Behavior Information Form	Recorded consumer clicks, additional purchases, purchases
Product category information table	The level of product category
Product Sales Data Sheet	Daily sales data for each category

4 Online User Consumption Behavior

4.1 Commodity Sales Data

Product sales data is shown in Table 2. The commodity sales data table presents the sales volume, tag price, and average sales price of each product corresponding to each type of commodity every day during the statistical period. The table contains 7.32 million rows of records. This table records each product's daily actual sales volume; therefore, it can be used for the historical sales analysis of time series and the source data of feature engineering. More importantly, because this data mining case predicts future demands, it is necessary to generate the training samples' explained variables according to each category's historical sales volume. The so-called mining technology is to carry out data mining in the system. In other words, it automatically selects the most valuable information from the vast database. When applying it in the business model, the company will lock the company's project goals, then search in the company's massive data and find that the data has been discovered or not. To achieve company goals by mining technology, we can design a system that analyzes customers' specific behaviors. Its characteristics can select the data that meets user consumption behavior analysis from the massive data, thereby helping the workers analyze.

Table2. Product sales data

Commodity ID	Sales volume (unit: piece)	Average sales price (unit: A)	Tag price (unit: A)
A	1	1.99	1.99
B	1	10.99	15.99
C	3	30.99	44.99
D	2	6.99	14.99

4.2 Sales Statistics in Different Price Ranges

Figure 1 shows sales in different price ranges. It gives the results that the sales volume of low-end products in the price range of 0-50 is about 146,800, the sales volume of mid-end products in the price range of 50-100 is approximately 257,500, and the sales volume of high-end products in the price range of 100-200 corresponds to 140,700. Simultaneously, the sales of high-end products in the price range of 200 yuan or more were around 127,700. Corresponding to the trend curve of sales volume, first rises and then declines, the corresponding sales volume is the most mid-end products, followed by low-end products,

followed by mid-to-high and high-end products. In terms of overall sales, the sales of mid-end products accounted for 38.27%, the sales of low-end products accounted for 21.82%, the sales of mid-to-high-end products accounted for 20.91%, and the sales of high-end products accounted

for 19%. Mid-range products are close to twice the sales volume of the other three categories. The market share of mid-end products, mid-to-high-end products, and high-end products is close to 80%.

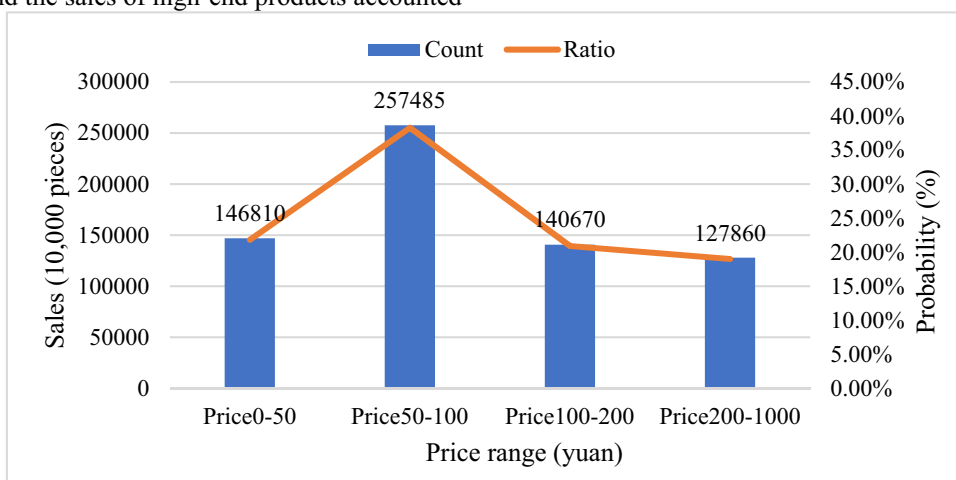


Figure 1. Sales in different price ranges

User attention in different price ranges is shown in Figure 2. The user's focus on the product is mainly researched from six aspects, including the brand, quality, function, appearance, price, and user experience. For each price range, users will first pay attention to the product's appearance and function when purchasing products, followed by the user experience of the product, product quality, and the last two are the corresponding product

price and product brand. For example, the function of the product, in the price range of 0-50 yuan and the price range of 50-100 yuan, the user's attention is 26%, in the price range of 100-200 yuan, the user attention is 27%, and the price is above 200 yuan. The interval is 21%, that is, in the process of high-end products, users' attention to functions is increasing.

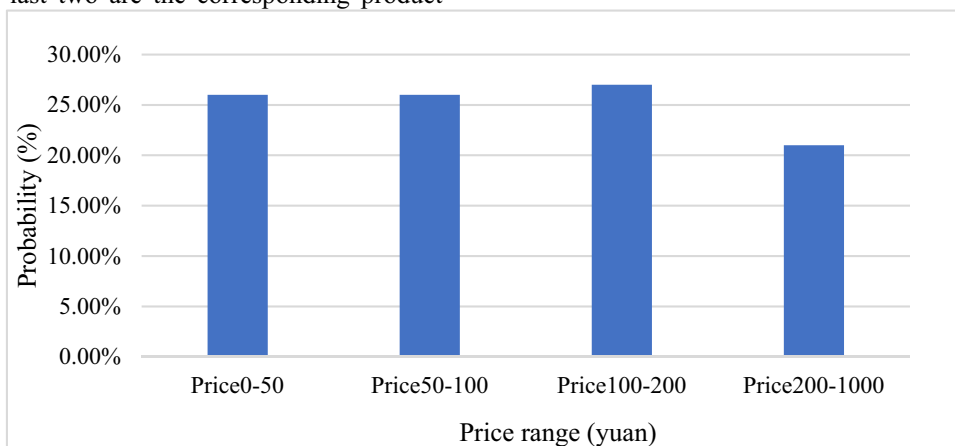


Figure 2. User attention in different price ranges

5 Conclusion

This research mainly discusses the consumption behavior of Internet users based on big data. When the target text has been obtained, the text preprocessing is required first. Clean and delete content that is of no value or extremely low value in the text. First define the Request network connection request. Secondly, write a function module for obtaining multiple commodity IDs. Traverse the full text to obtain the target content, which is often used to retrieve and replace the target text. Finally, the crawling of individual product information.

According to the research product goals, big data technology is used to obtain data on two dimensions of product sales information and user experience information. Store the product sales specification text and product

review text as separate texts for the next stage of data cleaning to predict user consumption behavior. This research helps companies formulate precise marketing strategies.

References

1. Gu K , Tao D , Qiao J F , et al. Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data[J]. IEEE Trans Neural Netw Learn Syst, 2018, 29(4):1301-1313.
2. Wu C , Zapevalova E , Chen Y , et al. Time Optimization of Multiple Knowledge Transfers in the Big Data Environment[J]. Computers Materials & Continua, 2018, 54(3):269-285.

3. Mann L . Left to Other Peoples' Devices? A Political Economy Perspective on the Big Data Revolution in Development[J]. *Development and Change*, 2018, 49(1):3-36.
4. Andre J C , Antoniu G , Asch M , et al. Big Data and Extreme-Scale Computing: Pathways to Convergence - Toward a Shaping Strategy for a Future Software and Data Ecosystem for Scientific Inquiry[J]. *International Journal of High Performance Computing Applications*, 2018, 32(4):435-479.
5. Prasad S , Zakaria R , Altay N . Big data in humanitarian supply chain networks: a resource dependence perspective[J]. *Annals of Operations Research*, 2018, 270(1):383-413.
6. Neggers J , Allix O , Hild F , et al. Big Data in Experimental Mechanics and Model Order Reduction: Today's Challenges and Tomorrow's Opportunities[J]. *Archives of Computational Methods in Engineering*, 2018, 25(1):143-164.
7. Duan J , Tian R , Xing Y , et al. A Collaborative Pricing Framework for In-Network Caching in Information-Centric Networking[J]. *IEEE Access*, 2018, 6(1):40485-40493.
8. Martinez B , Adelantado F , Bartoli A , et al. Exploring the Performance Boundaries of NB-IoT[J]. *IEEE Internet of Things Journal*, 2019, 6(3):5702-5712.
9. Hendry H , Chen R C . Consumer Behavior based on APP use for Food and Beverage Consumption[J]. *AITI*, 2018, 15(1):1-13.
10. Shahini A , Ansari N . NOMA Aided Narrowband IoT for Machine Type Communications with User Clustering[J]. *IEEE Internet of Things Journal*, 2019, 6(4):7183-7191.
11. Cao X , Wang F , Xu J , et al. Joint Computation and Communication Cooperation for Energy-Efficient Mobile Edge Computing[J]. *IEEE Internet of Things Journal*, 2019, 6(3):4188-4200.