Application of Big Data Technology in Product Selection on Cross-border E-commerce Platforms

LIANG JIA-HUI

School of Finance, Guangdong University of Science & Technology, Dongguan, Guangdong, 523083
592497460@qq.com

Abstract. With Amazon as the study case, the application of big data in product selection on this e-commerce platform is studied. Two big data analysis tools commonly used in commerce, the MPP distributed database and the Hadoop distributed database, were analyzed. Based on big data technology, the search function of the platform, the analytical tools, and third-party data analytical tools, this study compared different levels of comments of customers for the same type of products and analyzed the product selection mechanism.

1. Big data technology

Big data refer to massive data that are difficult to capture, manage or analyze by common software. Such data are massive and dynamic, which include data produced by the users’ behaviors on the website such as searching, clicking, browsing, placing orders, liking, refunding, commenting and giving feedback. Big data analytical tools can be used to store, mine, analyze, process and judge the collected data to improve the cross-border e-commerce managers’ skills in decision-making, product management, enhance user experience and aftersales service, thus providing new ideas to development of cross-border e-commerce enterprises and promote reforms of management modes.

As the big data industry grows rapidly, big data analytical tools are improving and provide functions including smart analysis and visualization. Commonly used analytical tools include MPP, Hadoop, Spark, Storm, Apache and Drill, among which the first two distributed databases are widely applied in commerce.

The MPP distributed database is a database that realizes optimization through analysis of the workload, a massive database for clustering and processing. MPP is usually used to store data and is a data warehouse or a distributed data warehouse. The data stored include Web tracking data, sales data, customer data, or inventory data or system data. This database stores each column as an object to realize fast and efficient processing of data.

The Hadoop distributed database is a software framework that can conduct distributed processing of large amounts of data; it is a popular data analytical tool that can categorize content as the users input keywords to search. For instance, on cross-border e-commerce platforms, Hadoop can categorize the content based on the key words input by users. Hadoop is used in different fields and has become a big data ecological system with comprehensive functions. The core of the Hadoop framework is HDFS and MapReduce, which can provide storage and calculation of massive data.

For cross-border e-commerce enterprises can make use of third-party big data platforms. For example, Google and eBay are excellent third-party big data platform providers. When the consumers browse or search on cross-border e-commerce websites, their behavior data that reflect their preferences and
consumption habits are uploaded onto Google and eBay to process and find infinite business opportunities.

2. **Impact of big data technology on product selection on cross-border e-commerce platforms**

The trade of cross-border e-commerce platforms reached 0.88 billion yuan in 2018. The trade volume of cross-border e-commerce enterprises increased by 30% from 2010 to 2020. For cross-border e-commerce enterprises, using big data to select products that meet the market needs is key to their success. “Product selection accounts for seven, while operation accounts for three” is a saying that reflects the importance of product selection in cross-border e-commerce. To succeed in product selection, cross-border e-commerce service providers need to analyze the target customer group, the type of products, the development stage of products, the attributes and price range of the products. Based on prudent analysis of market needs and precise judgement of the customers’ needs, e-commerce enterprises select products in a targeted and effective manner to maximize the enterprises’ operational profits. The impacts of big data technology on product selection of cross-border e-commerce platforms are reflected in the four aspects:

2.1. **Investigation of the needs of the target market**

The market for cross-border e-commerce is large. Before starting operation, e-commerce operators need first to identify the market needs for the products. The target markets of cross-border e-commerce enterprises are in different countries which feature spatial difference; thus, enterprises should analyze the local consumers’ consumption behaviors and preferences based on local culture. Via big data technology, they can collect data on the consumers’ behaviors like search, browsing, comparison, purchase on their platforms, use data crawling, mining and analysis to parse the search frequency, browsing frequency, clicks, payments and trading success rates of purchases, visualize product information using the consumer market analysis mode, track the buyers’ information to identify their preference and thus the needs of the target market.

2.2. **Accurate product selection**

Cross-border enterprises should use big data technology to analyze all consumers on their platforms to realize accurate product selection. To select products that consumers is halfway towards success. Currently, cross-border e-commerce platforms have myriad customers, different brands and products ranging from household appliances, electronic products, books, cosmetics to daily supplies and food. Big technology can help accurately identify the important target consumers, their psychological features, their acceptable range of prices, the development phase of products and the profits of products. To realize product selection based on big data, cross-border e-commerce platforms should first determine the catalogue of products, ensure smooth marketing channels and stable supply of products, make in-depth analysis of the content and the type of products, the product selection mode and learn from the product selection principles of competitive platforms. Next, they should identify the target customer group. Based on the selected product catalogue, they need to grasp the acceptable brands and consumption points of different consumers and analyze the market share of different brands.

2.3. **Providing optimized customer experience**

In this big data age, cross-border e-commerce enterprises should use advanced big data technology to improve customer experience to increase their competitiveness. Improving customer experience via big data technology means providing more targeted and effective product design and marketing activities for consumers. eBay is the world’s largest e-commerce service provider and the earliest adopter of the e-commerce mode in business. On eBay, massive data on trading behaviors are produced every minute, including the consumers’ browsing logs and clicks; to improve customer experience, eBay use the results of big data analysis to understand the customers’ behaviors, preference and acceptable price range to adjust its website design, the prices of products and improve the experience of customers abroad.
2.4. Improving efficiency of business decision-making
In this big data age, decision-makers of cross-border e-commerce enterprises rely on reforms for their decision-making. Using big data, they collect customer data on consumers, and analyze their target market and target customers. Meanwhile, with MPP and Hadoop databases, they can share data in real time, decompose the data, conduct calculation on independent nodes, and re-combine the data based on the calculation results. With big data tools, e-commerce service providers can obtain analysis results in merely four to five seconds and realize in-depth analysis of customers on their platforms. Indeed, as complexity of the calculation increases, the analysis time increases. Big data technology can streamline the framework of the data system, help enterprises grasp the dynamics of customers, analyze the needs of market needs, improve the business innovation capacity of cross-border e-commerce platforms and improve the business decision-making efficiency.

3. Application of big data technology in cross-border e-commerce platforms.

3.1. using the self-equipped search function of Amazon to search hot products
Search hot search words on Amazon to find hot products. In March 2020, the author obtained big data by searching keywords on Amazon and found the 8 hottest products, including toilet paper, disinfectant, mask, clinical thermometer, disinfectant spray, puzzles for adults, clinical thermometer for adults, disinfectant wet tissue. The first place, i.e. toilet paper, became the hottest product during the covid-19 epidemic.

It is worth noting that the product “puzzles for adults” was a hot product that ranked top among the hottest products. The changes in the ranking are shown in Table 1.

| Table 1. Changes in the ranking of puzzles for adults in March 2020 |
|---------------------------------------------------------------|
| Product | 3/1—3/7 | 3/8—3/14 | 3/15—3/21 | 3/22—3/28 |
|-------------------|---------|---------|-----------|-----------|
| Puzzles for adults | 1268 | 229 | 13 | 8 |
| Puzzles | 2325 | 568 | 42 | 26 |
| Puzzles for adults 1000 pieces | 4598 | 1026 | 63 | 33 |
| Jigsaw puzzles 1000 pieces | 5799 | 1757 | 136 | 60 |

Source: Amazon

The four types of puzzles above did not rank top in early March, but at the end of March, the ranking increased rapidly. These products are products in the “fulfilment by merchant” (FBM) mode and have advantages in product supply. During the epidemic, the puzzles sell well as a recreational tool for those confined at home in the lockdowns, thus becoming a hot product.

3.2. using the data analytical tools of Amazon to realize differentiated product selection
Amazon Brand Analytics is a tool Amazon provides for registered sellers to produce valuable analytical reports for products and marketing to help brand sellers to make rational decisions. The author searched the keyword “scrubs” on Amazon and found that the top search result was “scrubs cup”, and by clicking the ASIN code, all results were B018LOLMSO. This product was procured by Amazon and the price was low ($6.98/item); it was in short supply and was in pure colors. The product had room for improvement – it could be provided in mixed color patterns and the price could be increased properly. When searching for scrub cups in mixed color patterns, the results were from Chinese sellers in the FBM mode, priced at $13.99. The shelving date was 25th March, and ranked 25th among products in the same category. The daily sales reached 20, and was identified as a hot product.

The author also used the Jungle Scout tool to make analysis. Jungle Scout is a tool developed in 2015 by a statistics brand in the US for Chinese Amazon sellers. Use it to analyze the brand GUOER, and the
author found that the monthly sales of hit 4000, priced between $12.98 and $14.98; by testing the profit on the website http://www.amz123.com by considering the normal delivery fees, the author found the net profit was at least $4.8, whether it was in the FBA mode or the FBA mode, with a profit rate of 34.31%, showing great prospect for this product.

The author selected products by referring to the four list on Amazon’s Movers & Shakers, and identified the hot search index and recent search volume. For instance, the product set of “2 neck strapped safe filter and multifunctional masks + ten pasting patches” was released on 7 April and ranked 2nd among products of the small category, and ranked the 43rd in the brand store. This product also had much room for improvement but ranked 88th in hot search results. Also, by using Movers & Shakers, the author crawled search results for the product of “set of 3 stainless steel tweezers”, which moved up from the 100596th place to the 97th place, the ranking of which moved up by 103607%. Currently, it was in short supply. Movers & Shakers could be used to explore the buyers’ potential needs and realize differentiated product selection.

3.3. Using third-party data analytical tools to summarize the product’s sale increase
Adobe Analytics can also be used for product selection. This tool is a rule and algorithm-based attribution model for big data analysis. Enterprises can use Adobe Analytics to make prediction to make correct investment decisions. During the epidemic, the author used the tool to analyze the hot products: the data on 11th March showed that the sales of protection tools increased by 807%, toilet paper by 231%, and OCT by 271%, food by 87%, household fitness facilities by 55% and desktops, laptops and office appliances by 40%, toys and reactionary gadgets by 50%~100%, household items and home appliances by 20%~50%, furniture and lighting products by over 20%. The analyses above show that protection items increased the most rapidly in sales and became hot products during the epidemic.

3.4. Selecting products by comparing the buyers’ comments of different levels on Amazon
Enterprise can also compare buyers’ comments of different levels on Amazon to realize product selection. This method is suitable for sellers who already identified the category of products but have not identified the features of products. By comparing different levels of buyers’ comments, they can grasp the products’ features and factors for disapproval to improve their products and realize product selection. In practice, 90% buyers review comments before buying, and favorable comments are key to the success of Amazon sellers. Power Reviews is a salient company that has developed excellent comment review software. The comments include Amazon Reviews (about products) and Amazon Feedback (about the sellers). Reviews and Feedback are different. Amazon Reviews are comments provided by buyers for the products they have bought on Amazon; while Amazon Feedback is feedback provided by buyers about the sellers, including their delivery speed, the package of products, service attitude and their intention to buy again. The reviews and feedback can be compared to provide a basis for product selection.

4. Conclusion
Using big data technology for product selection of cross-border e-commerce platforms is a hot research topic in these two years. The self-equipped search functions, data analysis tools of cross-border e-commerce platforms or third-party analytical tools have provided ample help for enterprises in product selection. However, this technology is still in its infancy and many technical bottlenecks are to be addressed. It is advisable that by better understanding the users’ needs and improving the e-commerce platforms’ functions, the cross-border e-commerce service providers can find an efficient way to use big data to realize accurate product selection.

Acknowledgement
Funded project: this project is funded by the key science research program for universities in Guangdong in 2017 “training of cross-border e-commerce talents for innovation and entrepreneurship systems by universities, industries and enterprises” (Grant No. 2017GXJK199), and the 2019 science research
program of Guangdong University of Science and Technology – “Development of a scientific innovation center in the Guangdong- Hong Kong- Macau Greater Bay Area”.

References
[1] Wu Y. Y, Han S., Jin X. Application of big data technology in cross-border e-commerce [J]. Inner Mongolia Science technology & Economy, 2019(22) 54-56.
[2] Sun A. T., Sun M. Y. Research on customer information of O2O cross-border e-commerce websites based on big data technology [J]. China Management Informationization, 2016 (19): 154-156.
[3] Hu Z. F. Strategies for successful product selection of cross-border e-commerce buyers [J]. Practice in Foreign Economic Relations and Trade, 2018 (08): 67－70.
[4] Deng Z. C. Analytical strategies for product selection on cross-border e-commerce platforms based on big data technology [J]. Special Zone Economy, 2019 (365): 135-137.
[5] Huang Y. W. Research on application of big data technology in the product chain of cross-border e-commerce [J]. Foreign Economic Relations Trade, 2019 (7): 75-77.