Research on Consumer Behavior Analysis and Recommendations for Cross-border E-commerce based on Big Data

Shasha Wu*

* Corresponding author e-mail address: 6843599@qq.com

School of Foreign Languages, Wuhan Qingchuan University, Wuhan, China

Abstract. At present, cross-border e-commerce is becoming an important transaction method for international trading. Therefore, the reasonable analysis for purchasing behaviors and precise recommendations are essential for the trading platforms. However, existing recommendation system relies on the collaborative filtering from users and ignore the extraordinary characteristic information from users including gender, hometown, and historical preferences. In this work, we initially determine the review scoring in consumer behavior, the influence of external factors such as time and geography on consumer behavior is analyzed, and the influence of time, space and other factors on the quality of merchants is eliminated. Subsequently, a training neural network is established to obtain the recommendations results for different consumers from the previous analysis module. Finally, we estimate our model in real-world commerce data and compare with existing recommendation algorithms. From our simulation results, we can observe that our proposed model can achieve the precise recommendations with reasonable time complexity.

Keywords: Cross-border E-commerce; Consumer Behavior Analysis; Recommendations; Neural Network.

1   Introduction

Cross-border e-commerce represents a convenient method for international tradings where can improve the digital and global marketplace without physical borders. This sector, fueled by the rapid evolution of technology and the increasing interconnections of the world, enables businesses and consumers to transcend geographical and cultural barriers, offering a platform for international trade and cultural exchange [1]. Cross-border e-commerce is characterized by transactions involving buyers and sellers from different countries, completing transactions in the marketplace through digital channels. With the advancement of internet technology, the popularity of mobile devices, and the rise of social media, e-commerce platforms have driven a shift in the form of global commerce, and these factors have combined to democratize access to global markets.

In the development of cross-border electronic platforms based on Internet technology, the massive amount of information gradually accumulated and generated has caused most of the user-generated content to be buried, making the value of the platform not fully highlighted [2].

On the one hand, it makes it difficult for most businesses to be discovered by consumers, so they cannot get users' purchases and reviews. On the other hand, the exploration of massive data cannot satisfy consumers who pursue high efficiency, thus affecting the user's experience of using the platform. In short, when the Internet platform is under data overload, the inability to effectively match users and merchants is a vicious circle for both merchants and users. Therefore, it is necessary to analyze consumers and make effective recommendations [3].

Consumer behavior analysis is a complex data processing process in which analytical models delve into diverse and intricate patterns of how individuals select, buy, use, and dispose of goods and services. The research on consumer user behavior and the recommendation of merchant products are inseparable from the attributes and characteristics of cross-border e-commerce platforms [4]. At present, cross-border e-commerce platforms have two major characteristics, including user-generated content platforms, which support the development of the platform is the user's participation content, which not only contains the information value expressed by the user, but also contains the user's unique views and preferences, which has valuable reference value for analyzing user behavior; On the other hand, the unique characteristics of the platform are high active user participation and large interaction space, and because of this, high user stickiness is generated, so social networking has become the top priority of the platform [5].

Massive user information is the treasure of cross-border e-commerce platform, merchants and consumers through the information to generate feedback, so as to produce the foundation of operation and development, and the ultimate goal is to obtain satisfactory information for users, thereby increasing user stickiness, the platform can rely on user traffic agency promotion and advertising services, so as to obtain realization [6]. How to use user-generated high-quality content, analyze the characteristics and needs of users, match users with information, increase user experience, promote the frequency of use of old users and new users, and generate high-quality information is an important issue affecting the development of the platform [7].

Recommendations play a pivotal role in decision-making processes across various domains, from business strategies and public policy to personal choices and academic research. They serve as guided suggestions, informed by expertise, analysis, or experience, and are tailored to address specific goals, challenges, or questions [8]. The consumer behavior of the existing cross-border e-commerce platform is mainly for consumers to go to the platform for review and score feedback after consumption, and the research on consumer behavior of online reviews mainly includes the characteristics of consumers' review scores, text language characteristics of reviews, and review intentions.

As recommendation systems become more important in many areas, many methods are being used to improve recommendation systems, more often based on customers reviews in e-commerce platforms, which interpret users' ratings and help us understand their rating behavior more effectively [9]. In addition, review scoring also involves different aspects of the product that are important to the user and are often expressed in the reviews of customers.

2 Related Works

Initially, the research on the consumption behavior of cross-border e-commerce platforms shows that online reviews have an important impact on the market and sales, and focuses on the characteristics of consumers' review ratings, including the length of the score and the characteristics of the rating, indicating that the length of the review is related to the quality of the product and reflects the willingness of consumers to comment.

Consumer psychology focused on the anxiety and behavior of consumers, particularly college students, in the context of cross-border e-commerce. It highlighted that rapid e-commerce development led to keen interest in overseas fashion brands, often resulting in excessive consumption and subsequent anxiety and regret [10].

Subsequently, a comprehensive study analyzed the interaction of customers' personal traits, like national culture and trust disposition, on purchase intention within different e-commerce contexts. It emphasized the moderating role of the country-of-origin in the relationship between national culture, trustworthiness, and consumer behavior [11].

Marios studied whether the user's bias affects the user's rating and comments, the article selects the user's hometown as a variable, and the analysis model believes that the bias comes from the difference in popularity between the current place and the reviewer's hometown, which can be explained as the difference in popularity, and is defined as a function of two opposing forces, with the score to represent the user's satisfaction, and the popularity bias as the user's expectation, the popularity difference deviation model is constructed to prove that the user's hometown popularity affects the user's review behavior and review score in a new city [12].

For the product recommendation model of the trading platform, researchers Ghasemi N and MomtaziS conducted a study on Internet retailers to help users provide recommendations based on their tastes and preferences, and improve the recommendation system by finding similar users based on reviews and ratings. Using seven different methods to calculate the similarity of users' comments, two of which are lexically based, two of which benefit from the neural representation of words, and three of which are based on the neural representation of text, it is concluded that the model based on the long-term short-term memory network achieves the best results [13].

3 Methodologies

3.1 Notions

Above all, we conclude the primary used parameters in following Table 1, which involves the big data technology for customers behavior analysis component.

Table 1 Notions Description

Parameter Symbols	Explanations
COR	Correlation coefficient
X	Customer characteristics set
Y	Consumption amount

F	Extracted features
δ	Learning rate
β	Connection weights
ϵ	Training epoch

3.2 Behavior Analysis

Before the analysis for customers, the model will extract the features and transfer these features into readable information. Following Equation 1 describes the feature selections for customers by calculating the correlation analysis. Calculating the correlation coefficient between purchase frequency and user consumption, features with high correlation may be more valuable.

$$COR = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2(Y_i - \bar{Y})^2}} (1)$$

Feature building is the process of constructing unique features from existing data to enhance the capabilities of a recommendation model. Tremendous features can cause the model to overfit, while poor quality features can interfere with the model's learning. Proper feature engineering can dramatically improve the performance of a model and the accuracy of its predictions. Following Equation 2 describes the establishment process for customers.

$$F = \{f_{COR} > 0.7\}(2)$$

In consumer behavior analysis, it is important to be able to accurately identify and leverage these characteristics to extract valuable insights and develop strategies accordingly. Through a comprehensive analysis of these characteristics, it is possible to better understand consumer behavior, predict future trends, and develop corresponding marketing and product strategies.

3.3 Commendation Training

After introducing the basic features extractions and analysis model, following Figure 1 demonstrates the recommendation method by deploying a neural network with double layers and active functions.
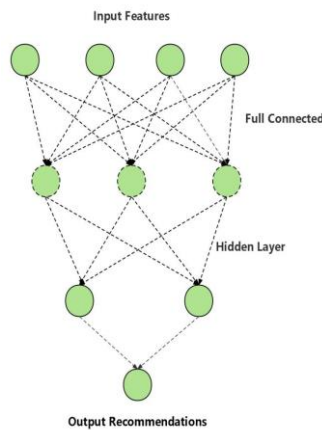
Figure 1 Illustration framework of recommendation neural network.

Neural networks predict items that may be of interest to users based on their historical data and current behavior. This could include recommending new products to users that they've never seen before but might be interested in, or recommending related products based on the user's shopping cart and browsing history. The whole system is designed to improve user satisfaction and purchase conversion rate, while also bringing higher revenue to e-commerce platforms, and the neural network is highly adaptable and learning due to its high adaptability.

4   Experiments

4.1   Experimental Setups

We adopt the Amazon Reviews DataSet, which is provided by Amazon and is a dataset of millions of product reviews on Amazon.com. Includes a large amount of data on Amazon that provides in-depth analysis and insights for a wide range of product categories. Additionally, we compare our model with existing collaborative filtering (Colf) and comment rating-based (CR) methods with recommendation accuracy.

### 4.2   Experimental Analysis

In this simulated experimental analysis, we compare three e-commerce product recommendation methods. The evaluation metrics used include accuracy, precision, and recall for comparison. Initially, we demonstrates the recommendation accuracy in following Figure 2.
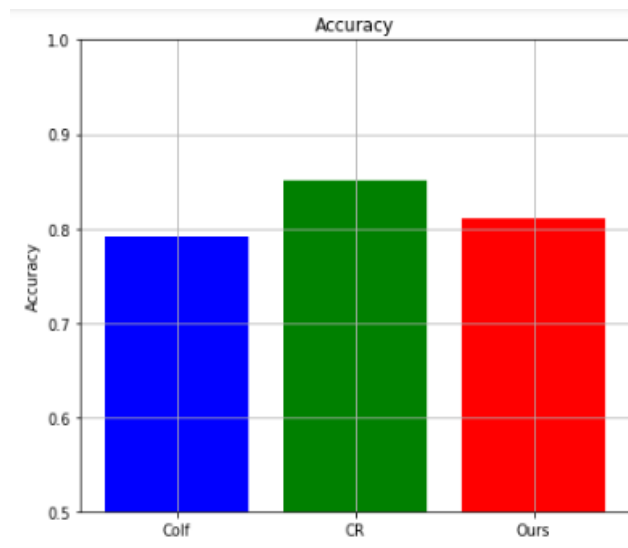
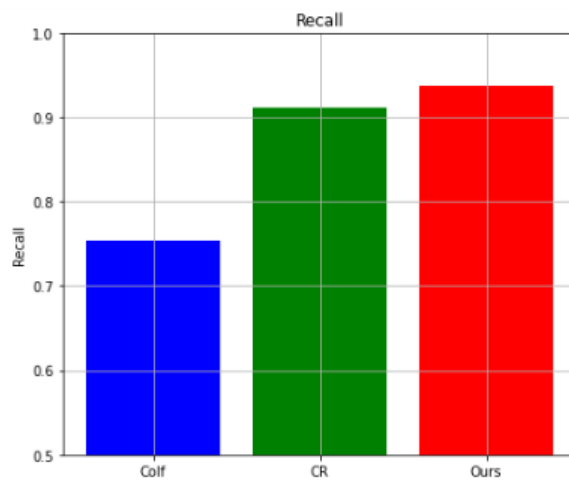**Figure 2** Recommendation accuracy comparison results.

Additionally, The Precision evaluation index of the product recommendation system is an important indicator to measure the performance of the recommendation system. It is mainly used to evaluate how many products recommended by the recommend system are actually of

interest to users. Precision evaluates the effectiveness of recommend systems. A high Precision value means that most of the recommended products are of interest to the user, which can increase user satisfaction and system usage. Following Figure 3 compares our model with existing methods with the metric of precision in identical environment.
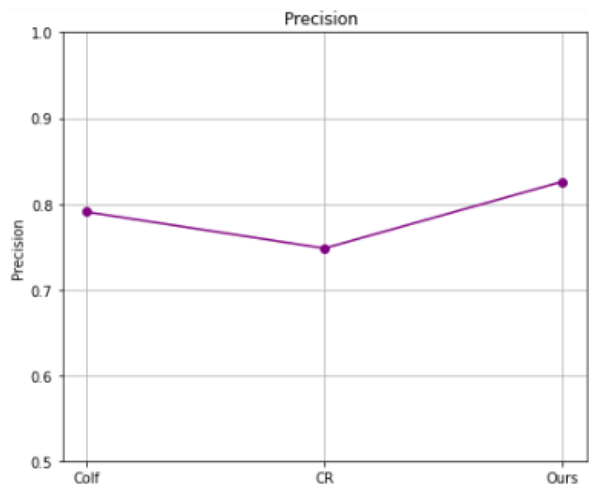
Figure 3 Recommendation precision comparison results.

Finally, the module recall is also important in recommendation results, which calculates the number of products that a user likes in a public recommendation list divided by the total number of products that the user is actually interested in. The high recall rate indicates that the recommend system is able to effectively identify items that may be of interest to users from a large number of products. This is important to ensure that users don't miss out on any products that might be of interest, especially in content-rich areas. Following Figure 4 demonstrates the comparison of recalls for our proposed model.

Figure 4 Recommendation recall comparison results.

5   Conclusions

In the field of cross-border e-commerce, the use of big data to analyze consumer behavior can achieve accurate product recommendations and the extraction of relevant features. The model reveals the diversity of consumer preferences across different regions and cultures, prompting platforms to adapt their products and marketing strategies to local tastes and preferences. Personalized experiences enabled by big data are key, including personalized product recommendations, targeted marketing campaigns, and customized shopping experiences. In addition, big data plays an important role in predictive analysis and demand forecasting. For future improvement strategies, big data provides e-commerce platforms with a critical way to understand and meet diverse consumer needs, optimize operations, and remain adaptable in an ever-changing digital marketplace.

References

[1]   Zhou, Fuli, and Yijie Liu. "Blockchain-enabled cross-border e-commerce supply chain management: A bibliometric systematic review." Sustainability 14.23 (2022): 15918.

[2]   Hazarika, Bidyut B., and Reza Mousavi. "Review of cross-border E-commerce and directions for future research." Journal of Global Information Management (JGIM) 30.2 (2021): 1-18.

[3]   He, Yugang, Renhong Wu, and Yong-Jae Choi. "International logistics and cross-border E-commerce trade: who matters whom?." Sustainability 13.4 (2021): 1745.

[4]   Zhu, Wenlong, Jian Mou, and Morad Benyoucef. "Exploring purchase intention in cross-border E-commerce: A three stage model." Journal of Retailing and Consumer Services 51 (2019): 320-330.

[5]   Zhang, Xiaheng, and Sukun Liu. "Action mechanism and model of cross-border e-commerce green supply chain based on customer behavior." Mathematical Problems in Engineering 2021 (2021): 1-11.

[6]   He, Xijin, Shuxia Meng, and Juanjuan Liang. "Analysis of cross-border E-Commerce logistics model based on embedded system and genetic algorithm." Microprocessors and Microsystems 82 (2021): 103827.

[7]   Li, Bing, Jiahua Li, and Xijun Ou. "Hybrid recommendation algorithm of cross-border e-commerce items based on artificial intelligence and multiview collaborative fusion." Neural Computing and Applications 34.9 (2022): 6753-6762.

[8]   Zhang, Zhongqiang. "An optimization model for logistics distribution network of cross-border e-commerce based on personalized recommendation algorithm." Security and Communication Networks 2021 (2021): 1-11.

[9]   Chen, Jin, and W. U. Chunqiong. "The design of cross-border E-commerce recommendation system based on big data technology." 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP). IEEE, 2021.

[10]   Li, Yuanqiao, and Yuanying Deng. "Analysis of the impact of cross-border e-commerce on consumer behavior anxiety from the perspective of consumer psychology." CNS Spectrums 28.S1 (2023): S31-S32.

[11]   Tikhomirova, Anna, et al. "How culture and trustworthiness interact in different e-commerce contexts: A comparative analysis of consumers' intention to purchase on platforms of different origins." Frontiers in psychology 12 (2021): 746467.

[12]   Kokkodis, Marios, and Theodoros Lappas. "Your hometown matters: Popularity-difference bias in online reputation platforms." Information Systems Research 31.2 (2020): 412-430.

[13]   Ghasemi, Negin, and Saeedeh Momtazi. "Neural text similarity of user reviews for improving collaborative filtering recommender systems." Electronic Commerce Research and Applications 45 (2021): 101019