AEM JOURNAL.ORG

Research on E-commerce User Profile Construction and Personalized Service Strategies Based on Multi-Source Heterogeneous Data

Junfeng Ren
Zhengzhou University of Economics and Business, ZhengZhou 450000, Henan, China
Corresponding author: Junfeng Ren (e-mail: rcy120224@163.com)

ABSTRACT As e-commerce enters an era of intensive competition, data-driven refined operation has become essential for improving service accuracy and user conversion. However, user data are often scattered across different platforms and systems, producing data silos, distorted user profiles, and ineffective recommendations. This study proposes a user profile construction framework based on multi-source heterogeneous data fusion. E-commerce transaction data, behavioral logs, social media content, search records, product information, and customer service records are integrated through data cleaning, entity alignment, and feature fusion. A multidimensional tag system is then constructed from five dimensions: basic attributes, consumption power, interest preference, social relationship, and lifecycle stage. Based on the resulting profiles, dynamic personalized service strategies are designed through user segmentation and recommendation mechanisms, and their effectiveness is verified through A/B testing. The results show that user profiles constructed from multi-source data achieve significantly higher coverage and accuracy than profiles based on a single data source. The study provides an operable framework for overcoming data barriers and upgrading intelligent e-commerce services.

INDEX TERMS multi-source heterogeneous data, user profiling, personalized recommendation, data fusion, e-commerce

I. INTRODUCTION

THE extensive penetration of digital technology is fundamentally changing the operational logic of e-commerce by synchronizing consumer data with the technical precision of smart weaving and high-performance fiber production. This digital shift ensures that market demands directly inform the innovative engineering of advanced functional fibers, creating a seamless link between digital analytics and modern manufacturing. After enjoying significant growth fueled by traffic dividends the e-commerce industry is fast progressing towards the user-centric, intensive model. User profiling as a digital tool to characterize user characteristics and understand user needs, from auxiliary analytical methods, has become an important engine for precision marketing and intelligent services.

However, the major challenge in constructing user profiles is the complexity of data sources. Users leave traces of their transactions on shopping platforms, post product reviews on social media, enter keywords into search engines, and give feedback to a company on their user experience through C.S. systems - these behavioral trajectories are left across heterogeneous systems, with different data formats, update frequencies, and semantic structures. Traditional methods often use one particular source of data: creating a consumer capability profile just from transaction data; creating an interest preference profile just from browsing logs. Users are often viewed in one-dimensional ways. An even more serious challenge is that the same user could be represented in different identities on different platforms, and cross device & channel behavioral breakpoints make it difficult to completely reconstruct the user journey.

This research seeks to concentrate on the above mentioned issues and proposes to discuss a systematic solution. At the methodological level, a comprehensive framework of user profile construction including data collection, data cleaning and alignment, feature fusion, and tag generation is proposed and an improved clustering algorithm is proposed to realize dynamic user segmentation. At the application level, it designs personalized recommendation, coupon distribution and content push strategy according to the segmentation of user profiles, and verify whether it is effective through quantitative experiments. The theoretical contribution of this research is to extend the scope of multi-source data fusion technology in e-commerce user modeling, while the practical contribution provides a full-path reference from data integration to strategy implementation for enterprises specialized in high-performance fiber commerce and smart distribution. This framework enables a seamless transition from raw data analytics to the technical marketing of advanced functional fabrics and innovative weaving materials.

II. RELATED WORK

With the deep integration of digital technology and e-commerce, the evolution of marketing strategies has shown a significant transformation from experience-driven to data intelligence-driven. Semenda et al . [1] explored how to apply game theory and matrix analysis to the social media marketing strategies of ecommerce companies. Putri and Setiawan [2] analyzed the strategic contributions of influencers to the effectiveness of e-commerce marketing to understand how they help brands build trust, expand influence and ultimately promote sales. Wang and Guo [3] revealed that consumer behavior is significantly influenced by their validation psychology in live marketing of travel e-commerce. Sun and Wang [4] confirmed the powerful capabilities of AI technology in analyzing customer engagement sequences and optimizing conversion funnels. Through analysis, companies can gain a deeper understanding of the complete path from initial contact to final conversion, and identify key touchpoints and drop-off links. Against the backdrop of the booming development of big data and artificial intelligence technologies, Wang [5] explored how to use the power of big data and artificial intelligence to promote the innovation of e-commerce marketing models to adapt to the new technological environment and market demands.

Zhuk and Yatskyi[6] emphasized the key role of artificial intelligence and machine learning in improving the efficiency and effectiveness of e-commerce marketing. Smerichevskyi et al. [7] proposed a strategic framework for the integrated development of marketing and logistics in the e-commerce innovation ecosystem , emphasizing that in order to gain a sustainable competitive advantage in a dynamic market, enterprises must plan and manage marketing and logistics strategies in a unified manner. Sakinah and Heruwasto[8] confirmed the key role of e-commerce marketing incentives in building consumer trust and satisfaction, and thus affecting their loyalty. Kalogiannidis et al. [9] concluded that strategic marketing management has a significant positive impact on the success of e-commerce promotions. Sanbella et al. [10] investigated, identified and optimized online marketing strategies to increase sales and support the growth of the e-commerce industry . The above results mostly focus on a single data source or a specific marketing link, and the integration and mining of cross-platform, multimodal and heterogeneous data is still insufficient . This paper aims to overcome the limitations of existing research by integrating multi-source heterogeneous data from e-commerce platforms, social media, search logs, and behavioral trajectories to construct a high-precision and evolvable user profile system. Based on this system, dynamic personalized service strategies are designed to provide theoretical support and practical pathways for the refined operation and intelligent service of e-commerce.

III. METHODS

A. ACQUISITION AND PREPROCESSING OF MULTI-SOURCE HETEROGENEOUS DATA

This study's data collection contained six data sources from e-commerce platform numbers in the middle of the size, including user basic information tables, transaction order tables, behavior log tables, product information tables, user reviews tables, and customer service conversation records. The total worth of raw data was about 4.5TB, comprising some 850,000 registered users and some 3.2 million valid transaction records. The main problem to be overcome during the data collection phase was the unified identification of the user identities in the different channels. The same user can login to PC and mobile terminals with different accounts, or have anonymous browsing behaviors on different terminals. To solve this problem, this research has built a unified user identification system based on the combination of device fingerprints and account binding. For logged-in users, the registered mobile phone number was used as the primary key for association; and for unlogged-in users, the co-occurrence relationships of Cookie ID, Device ID, IP address and browser fingerprint were analyzed and, using a graph connectivity algorithm, the multi-device behaviors belonging to the same individual were merged[11,12].

The data cleaning phase deals with the characteristics of the different data sources, separately. Transaction order data is filled with duplicate records and outliers. Records with logical conflicts such as negative order amounts or payment times earlier than order placement times, are cleaned out. For multiple occurrences for the same order number, the first record is retained and is flagged as duplicate. Behavioral log data is the biggest, with around 8 million records per day, containing 6 types of events: page views, searches, clicks, favorites, add to cart and order. During cleaning the crawler and non-human behaviour is filtered out. Accesses from a single IP greater than 20 requests per second is flagged as anomalous and eliminated, and missing key fields in records are deleted. Social media's comment data has a lot of unstructured text in it. Regular expressions can be employed to get rid of the html tags, emojis, and advertising links. Jieba word segmentation for Chinese word segmentation, custom e-commerce dictionary loaded, the accuracy of proper noun recognition is better. Customer service conversation records: Customer service conversation records include private information. Sensitive fields such as user names and contact information are anonymized in the data import phase, and only structured features remain, such as how long conversations with a customer last, whether the sentiment of the information is positive or negative and what kind of questions are asked[13,14].

B. USER FEATURE EXTRACTION AND PROFILE TAGGING SYSTEM CONSTRUCTION

Based on the preprocessed fused data, this study extracts user features from five dimensions to create a multi-layered user profile tagging system. Gender, age, region, membership level, and registration channel are some of the basic attribute dimensions. Gender and age is obtained directly from registration information and missing values are filled in based on semantic inference from purchased product categories and browsed content.

Spending power dimension includes extended pixels based on the RFM model. In addition to traditional metrics such as the purchase time, purchase frequency, spending amount, three kinds of derived metrics are introduced, including average order value, coupon using rate and high-value item proportion. High-value products are defined as products (SKUs) whose prices on the platform are more than 50% higher than the median price of similar products. Spending statistics are used to separate promotional periods from regular sales periods in order to identify the price sensitive and the essential needs users. Statistics show that around 23.5% of the sample users are active users with a monthly purchase frequency of over 3 times with the average order value concentrated in the range of 150 and 350 yuan[15].

Interest preference dimensions that are pulled from behavior logs and search records. Statistics are compiled on the categories of products browsed, searched and favorite of users in order to calculate the preference weights for major categories (for example, women's clothing, digital products, beauty products) and subcategories (for example, dresses, mobile phones, lipsticks). To avoid duplicate calculations of pre-conversion and conversion behaviors within the same user-product pair, this study only considers the behavior with the highest weight for multiple behavioral interactions related to the same product within the same session when calculating the overall preference score. Behavior weights are defined as follows: order placement (3.0), adding to cart (2.0), favorites (1.5), browsing (1.0), and click (0.5). The overall preference score for each category is calculated based on this. Search keywords are segmented and their frequencies are statistically analyzed; high-frequency words are extracted to obtain candidate interest tags. For text data, the TF-IDF algorithm is used to extract feature words from reviews and customer service records to construct user attention tags.

The social relationship dimension is obtained mainly from the comment interactions, sharing behavior, and registration invitation records. The number of likes and replies received after a user will post a comment is statistically analyzed to calculate the user's index of influence. Follow and @mention relationships between users are identified to construct a user social network graph, and PageRank algorithm is used to compute the centrality of users in the social graph. For registration invitation behavior, the binding relationship between the inviter and the invitor is recorded and forms invitation link data. Analysis shows that user groups who have strong social relationships have category similarity and temporal synchronization of their purchasing behavior, providing a basis for social fission strategies.

The user lifecycle stages are separated according to user registrations duration, activity changes and purchase intervals. An improved survival analysis model is applied to calculate the user churn probability. Users are divided into five stages: new user stage (registered for less than 30 days), growth stage (registered for more than 30 days and made more than 2 purchases per month), mature stage (registered for more than 90 days and made more than 3 purchases per month), decline stage (purchase frequency in the past 30 days decreased by more than 50% compared to the previous three months), and churn stage (no purchases made for 90 consecutive days). Marketing strategies should be differentiated between users in different stages: the focus is on guiding repeat purchases in the new customer stage the focus is on reactivation intervention in the decline stage.

C. USER DYNAMIC GROUPING BASED ON IMPROVED CLUSTERING ALGORITHM

Based on the tagging system, this study uses an improved clustering algorithm to dynamically segment users into clusters for refined operations. Traditional k-means algorithms are sensitive to initial cluster centers and questionable to noise interference. This study introduces an initial selection of centers using density peaks, that is, SAPK-means algorithm. First, the local density and high-density distance of each sample point is calculated. Points with high local density and large distance from other points with high densities are selected as initial cluster centers, which will effective avoid the interference of isolated points on the clustering results. The construction of the clustering feature vector integrates the five dimensions of the tags mentioned above. Continuous variables are normalized, and categorical variables are one-hot encoded.

The number of clusters were obtained combining elbow-rule and silhouette-method. The sum of squared error of each cluster and the silhouette coefficient were obtained for varying k values. The k value that had the maximum silhouette coefficient and a gradually decreasing sum of the squared errors was selected. The calculation results show that the silhouette coefficient is 0.37 when k=5. Although this value is below the conventional statistical threshold of 0.5, it indicates a certain degree of behavioral overlap among user groups, which is common in real e-commerce behavior data—because user preferences are multifaceted and consumption behaviors can migrate across categories (e.g., the same user may simultaneously exhibit characteristics of "price sensitivity" and "high value"). To verify the robustness of the clustering structure, we also calculated the Davies-Bouldin index (DBI=1.24) and the Calinski-Harabasz index (CHI=3127.6). Both indices support that the clustering partitioning with k=5 is superior to other k values (DBI=1.41, CHI=2789 when k=4; DBI=1.35, CHI=2981 when k=6). Furthermore, through ten random initializations and repeated clustering, the consistency of sample allocation among clusters reached 89.2%, indicating that the clustering results have good stability. More importantly, the differentiated service strategy based on this clustering result showed significant improvements in business metrics in subsequent A/B testing (see Table 3), validating the effectiveness of user segmentation from an application perspective. Therefore, despite a silhouette coefficient of 0.37, considering multiple internal and external validation criteria, the k=5 clustering is considered to have acceptable effectiveness and interpretability in business scenarios. Table 1 shows the characteristics and size of the five user groups.

High-value and loyal users which are 13.24% are the key source of profit. This group has the characteristics of registration time generally exceeding one year, frequent purchase times are 4.3 times, average order value is 528 yuan, and most like to choose high-priced categories. They are active in comment interactions, influencing the purchasing decisions of other users and are less susceptible to promotions of the platform, more bound

to brand loyalty and shopping habits to repeat purchases. Potential growth users - the 21.91%, key players in the future of this type of value growth for the platform This group has mainly registered between 3-6 months ago and they have recently increased their frequency of purchase as well as their active look for golf club, however their average order value is not yet high. They favor categories of women clothing and maternity products, pay attention to new products and trendy products. Mid-tier, stable users, comprising 36.95%, are the platform's core user base. This group has regular consumption behaviour, with an average time to purchase of 2 weeks, and an average order value. They are somewhat sensitive to promotional activities but are not blindly following the trends. They favor buying high frequency products, household goods or food, and they pay a lot attention to value for money of products. Price-sensitive users: They were equal to 18.43%, with highly increased activity during major promotional periods, with low daily conversion rates. This group frequently browsed and added things to their carts, but had long purchase decision cycles and were influenced a lot by price fluctuations. The dormant user group accounts for 9.46%. This group has an extremely low marketing response rate and is difficult to repurchase and reactivate. Therefore, the platform needs to pay attention to and make efforts to retain this group in order to evaluate its return on investment.

IV. RESULTS AND DISCUSSION

A. THE EFFECT OF MULTI-SOURCE DATA FUSION ON IMPROVING PROFILE QUALITY

In order to check the effect of heterogeneous data fusion of multiple data sources on the improvement of user profile data quality, this paper designed a comparative experiment. User profiles were built based on single data source (transaction data only) and multi-source fused data, respectively, and evaluated on three dimensions i.e., coverage, accuracy, and timeliness. The evaluation index of the coverage was the proportion of tagged users over all active users; accuracy was checked by manual labeling of 1000 randomly chosen users; and timeliness was discussed by the profile update frequency and the response speed to the change in tags. Experimental results show that after integrating multi-source data the tag coverage of user profiles increased from 64.7% with a single data source to 91.3%. Transaction data only covers users that made purchases, but once behavioral logs can be integrated, interest preference tags can also be generated for browsing users who did not make purchases. After incorporating the information from the social media data, the social relationship and influence of the users were characterized. 34.2% of users got social tags. After integrating search logs, the accuracy of user intent recognition has been significantly improved, especially in the "browsing without purchasing" situation, in which the search keywords were direct reflection of user needs, and the interest tag hit rate was increased by 42.6%.

Accuracy was assessed using a confusion matrix and F1 score. For a sample of 1000 users, three operations staff independently labeled them with accurate tags based on complete user behavior records, with the majority vote used as the benchmark. The F1 scores for each tag dimension were calculated for both single-source and fused data profiles, and the results are shown in Table 2.

As shown in Table 2, the multi-source fusion data was substantially better than single data sources in all of the comparable dimensions. The dimension where the biggest improvement was registered was the brand preference that improved from 0.47 to 0.73 as the F1 score. Transaction data only reflects brands purchased by users, and the fusion of search and browsing data also found that users followed brands but did not purchase them, which provides a more complete data of users' brand choice. For example, the F1 score for category preference changed from 0.61 to 0.84 and was mainly driven by collecting more granular information in the browsing and click behavior logs, which depicted the distribution of the attention of the users in the different categories. The enhanced accuracy in determining the gender and age was facilitated by cross validation of multi-source information; when information was lacking in the registration, it could be surmised by the categories of purchased goods and content preferences. The social influence feature was not possible to build with a single data source, but combined the way that comments interact and sharing data successfully identified users with opinion leader potential.

In regard to the timeliness of data, the update cycle for the integrated data profiles has been reduced from T+7 to T+1. Transaction data can often have a settlement lag issue but behavioral logs are available in real

time. When a user creates some new browsing activity or search activity, the preference tags can be updated in hours.

TABLE 1. User grouping results and feature descriptions based on the improved clustering algorithm

User group types	Percentage (%)	Core Feature Description	Average order value (RMB)	Average monthly purchase frequency	Main preference categories
High-value loyal type	13.24	Long registration period, high repurchase rate, high average order value, and active comments and interactions.	528	4.3	Digital products, large home appliances, and beauty products
Potential growth type	21.91	Registration time is moderate, purchase frequency is increasing rapidly, and search activity is high.	276	2.8	Women's clothing, shoes and bags, maternity and baby products
Mid-to-Steady	36.95	Largest proportion, regular consumption behavior, moderate price sensitivity	189	1.9	Home, food, personal care
Price sensitive	18.43	High activity during promotional periods, high coupon usage rate, more browsing but fewer orders.	132	1.2	Department stores, fresh produce, and clothing
Sleeping Loss Type	9.46	No purchases in the past 90 days, low frequency of historical purchases, and very little interaction.	95	0.3	No obvious preference

TABLE 2. Accuracy comparison between single-source data source profiles and multi-source fusion profiles (F1 score)

Tag Dimension	Single data source (transactions only)	Multi-source fusion data	Increase
Gender recognition	0.72	0.86	+19.4%
Age groups	0.58	0.79	+36.2%
spending power	0.83	0.88	+6.0%
Category preferences	0.61	0.84	+37.7%
Brand Preference	0.47	0.73	+55.3%
Price sensitivity	0.53	0.71	+34.0%
social influence	-	0.69	not applicable
Average F1 score	0.62	0.79	+27.4%

B. EFFECTIVENESS OF PERSONALIZED SERVICE STRATEGIES BASED ON USER SEGMENTATION

Based on the above-mentioned five user groups, this study designed and implemented differentiated personalized service strategies, and verified the effectiveness of the differentiated personalized service strategies through three months A/B test. The experimental group used a user-segment-based personalized strategy, while the control group used the medium-term uniform conventional strategy of the platform. The experimental period ran from the first quarter of 2025, with a total of 150,000 users being involved. Users were assigned randomly into the experimental group and control group based on user groups to ensure that there was a balanced distribution of users into each group.

The strategies designed for the different user groups are as follows: High-value, loyal users' core needs are exclusive experience and priority accessibility to new products. Strategies include: pushing new products of limited edition, exclusive customized products, free shipping and free express refund participation, inviting them to beta testing of brands and member salon. The core needs of potential growth users is to find interests and to increase conversion. Strategies include: cross-category suggestion based on the most recent search and browsing behavior, pushing related product combinations and scenario combinations, limited-time category coupons to guide choice, etc. One of the constant needs of stable users in the middle range of the market is efficient shopping and stable repeat shopping. Some of the strategies include: Pushing the remind to re-stock and discounts on frequently purchased items, proactive outreach on the basis of previous buying cycle to anticipate the timings of purchase, recommending similar alternative products to enrich the choice. The foundation needs of price sensitive users are getting discounts and decision making on price comparison. Strategies include: pushing pre-sale and flash sale announcement to prompt bundling purchase interest, issuing tiered discount coupons to promote bundled purchases, showing price trends and historical low price indicator. The overriding strategy with dormant and churned users is inexpensive reactivation and value reassessment. After the strategy was implemented, the changes in the core indicators of each group were statistically analyzed, and the results are shown in Table 3.

All-in-all, the personalized strategy made the click-through rate (CTR) for the recommended products rise by 27.3%, the rate at which the user made the order increased by 18.6%, and the rate at which the user returned a purchase went down by 12.4%, but as well the number of users that were active in the month increased by 16.9%. The effects were different for different user groups. Potential growth users were most responsive to the personalized strategy (with a 34.7% increase in CTR, and 24.5% increase in order conversion rate), indicated that accurate interest matching has a potent influence on their decision-making. This group also experienced 12.3% increase in average order value and decreased return rate by 9.4%, reflecting the fact that the better match between recommended content and actual needs had reduced impulsive purchases. Price-sensitive user experienced 29.3% CTR increase and 21.6% conversion rate increase with significant 15.8% reduction in return rate but 4.2% reduction in average order value. This result is in line with expectations: the approach for this group was about getting people to convert but not about increasing average order value. Purchases that were based on coupons and flash sales were mostly low priced items but the large reduction in return rate suggests that the recommended products best satisfied their actual needs, which reduced their level of regret following their impulsive purchase. The 23.7% rise in activity shows that in the case when people were precisely targeted, the price sensitive users had a greater willingness to participate.

TABLE 3. Evaluation of the effectiveness of personalized strategies based on user segmentation (experimental group vs. control group)

User Group	Recommended click-through rate improvement	Improved order conversion rate	Changes in average order value	Changes in return rate	Monthly Activity Changes
High-value loyal type	+21.4%	+13.2%	+8.7%	-3.2%	+5.8%
Potential growth type	+34.7%	+24.5%	+12.3%	-9.4%	+18.3%
Mid-to-Steady	+26.8%	+17.9%	+5.6%	-6.1%	+11.2%
Price sensitive	+29.3%	+21.6%	-4.2%	-15.8%	+23.7%
Sleeping Loss Type	+18.2%	+11.5%	-2.1%	-8.5%	+32.4%
Overall average	+27.3%	+18.6%	+4.8%	-12.4%	+16.9%

To facilitate the evaluation of the actual effectiveness of the improvements, Table 4 lists the baseline return rate (i.e., the original return rate of the control group) for each user group and the change in the absolute return rate of the experimental group.

TABLE 4. Changes in Baseline Return Rate and Experimental Group Return Rate for Each User Group

User Group	Control Group Baseline Return Rate (%)	Experimental Group Return Rate Change (percentage points)	Experimental Group Final Return Rate (%)
High-value Loyal	6.2	-3.2	3.0
Potential Growth	18.5	-9.4	9.1
Medium Stable	12.3	-6.1	6.2
Price-sensitive	24.7	-15.8	8.9
Dormant	28.4	-8.5	19.9
Overall Average	16.8	-12.4	4.4 (weighted average)

As shown in Table 4, the baseline return rate for dormant users (28.4%) was significantly higher than that of other groups. Although the return rate in the experimental group decreased by 8.5 percentage points, its final return rate (19.9%) was still higher than that of other groups. This phenomenon is mainly due to "expectation discrepancies": dormant users have not interacted with the platform for a long time, and their purchasing behavior after reactivation is highly uncertain. A discrepancy easily arises between their expected matching degree of recommended products and their actual experience, leading to higher-than-average return behavior. This suggests that strategies for dormant users need further optimization, such as through more cautious interest detection or the introduction of trial mechanisms to reduce expectation discrepancies. All improvement indicators underwent rigorous A/B testing, and the statistical tests used a two-sample t-test. The p-values of all core indicators were less than 0.01, and the 95% confidence intervals did not include zero. Therefore, the percentage changes reported in Table 3 represent genuine logical improvements, rather than statistical errors.

The increase in various metrics for loyal users with a high value was comparatively moderate, but the 8.7% increase in average order value suggests that the new products and personalized recommendations succeeded in upgrading consumption. The return rate for this group was already low (with very little room to further lower, 3.2%). The 5.8% increase in activity is due to the fact that the core users are already highly active and it is hard for them to be more so. Mid-tier, stable users performed steadily with moderate increments of all the metrics, indicating that regular recommendations and restocking reminders were effective to meeting cyclical user needs. Dormant, churned users showed the greatest activity increase of 32.4%, suggesting how effective targeting strategies may be to reactivate users. However, the increase in conversion rate of this group was 11.5% which was less than the increase in activity, leading us to conclude that there are still conversion frictions between follow-up visits and repeat purchases. The return rate was decreased by 8.5%, and this number is still higher than other groups, which might have been related to expectation discrepancies due to long-term inactivity.

V. CONCLUSION

This study addresses the issues of insufficient multi-source data fusion and crude personalized service strategies in e-commerce user profiling, proposing a systematic solution to enhance the technical marketing of high-performance fibers and smart weaving materials. By refining these digital models, the research enables enterprises to bridge the gap between complex consumer requirements and the innovative production of advanced functional fibers. At the methodological level, a full-process framework covering data collection, cleaning and alignment, feature extraction, and tag generation is constructed, solving key technical challenges such as cross-channel user identification and heterogeneous data alignment. An improved SAPK-means clustering algorithm is introduced to classify users into five groups: high-value loyal users, potential growth users, stable and reliable users, price-sensitive users, and dormant/churned users. At the application level, a differentiated personalized service strategy based on user segmentation is designed and implemented, and the effectiveness of the strategy is verified through A/B testing.

This study has certain limitations. First, the data source is limited to a single e-commerce platform, and the generalizability of the conclusions needs to be verified on more platforms. Second, the research period was six months, failing to examine the evolution of user profiles and segmentation over a longer period. Third, the personalized strategy design is mainly based on historical behavior, and the capture of real-time intent remains limited. Future research could expand the range of data sources, introducing external social platforms and offline behavioral data to construct more complete cross-domain user profiles. Simultaneously, dynamic decision-making methods such as reinforcement learning could be explored to achieve real-time adaptive optimization of strategies. With the advancement of large-scale model technology, combining knowledge graphs with large language models to achieve automatic generation of profile tags and enhanced reasoning is a promising direction for refining the technical classification of high-performance fibers and smart weaving materials. This integration facilitates more sophisticated decision-making in the digital marketing of advanced functional fibers, ensuring that complex material properties are accurately mapped to evolving market demands.

AUTHOR CONTRIBUTIONS

Junfeng Ren designed, collected and analyzed the data, and drafted the manuscript. Junfeng Ren conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Junfeng Ren participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

The research was supposed by .
Zhengzhou University of Economics and Business 2026 Annual Education and Teaching Reform Research and Practice Project: Research on the Construction of Collaborative Talent Cultivation Ecosystem and Long-term Operation Mechanism under the Integration of Industry and Education (Project No. jg26006).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] O. Semenda, Y. Sokolova, O. Korovina, et al., "Using social media analysis to improve E-commerce marketing strategies," International Review of Management and Marketing, vol. 14, no. 4, pp. 61-71, 2024, doi: 10.32479/irmm.16196.

[2] B. Putri L W and T. Setiawan B L, "Analyzing the strategic contribution of social media influencers to e-commerce marketing effectiveness," JUMDER: Jurnal Bisnis Digital dan Ekonomi Kreatif, vol. 1, no. 2, pp. 12-21, 2025, doi: 10.1234/jumder.v1i2.21.

[3] Y. Wang and R. Guo, "Tourism e-commerce marketing following live-streaming: consumer behavior and verification psychology," Tourism Review, vol. 80, no. 4, pp. 914-927, 2025, doi: 10.1108/TR-10-2023-0738.

[4] M. Sun and H. Wang, "AI-Powered Customer Engagement Sequence Analysis and Conversion Funnel Optimization in Multi-Channel E-Commerce Marketing," Journal of Industrial Engineering and Applied Science, vol. 3, no. 5, pp. 7-20, 2025, doi: 10.70393/6a69656173.333138.

[5] J. Wang, "Innovation of e-commerce marketing model under the background of big data and artificial intelligence," Journal of computational methods in sciences and engineering, vol. 22, no. 5, pp. 1721-1727, 2022, doi: 10.3233/JCM-226152.

[6] A. Zhuk and O. Yatskyi, "The use of artificial intelligence and machine learning in e-commerce marketing," Technology audit and production reserves, vol. 3, no. 4, pp. 33-38, 2024, doi: 10.15587/2706-5448.2024.305280.

[7] S. Smerichevskyi, A. Kovalchuk, T. Obydiennova, et al., "Development strategies for marketing and logistics in the innovative ecosystem of e-commerce," Revista Gestão & Tecnologia, vol. 25, no. 2, pp. 90-107, 2025, doi: 10.20397/2177-6652/2025.v25i2.3162.

[8] M. Sakinah and I. Heruwasto, "Pengaruh e-commerce marketing stimuli terhadap trust, customer satisfaction, dan customer loyalty," Jurnal Samudra Ekonomi dan Bisnis, vol. 15, no. 2, pp. 258-271, 2024, doi: 10.33059/jseb.v15i2.8946.

[9] S. Kalogiannidis, F. A. Kagioglou, and O. Papaevangelou, "Strategic marketing management influence on making e-commerce promotions," International Journal of Scientific Research and Management, vol. 10, no. 1, pp. 2949-2955, 2022, doi: 10.18535/ijsrm/v10i1.em11.

[10] L. Sanbella, I. Van Versie, and S. Audiah, "Online marketing strategy optimization to increase sales and e-commerce development: An integrated approach in the digital age," Startupreneur Business Digital (SABDA Journal), vol. 3, no. 1, pp. 54-66, 2024, doi: 10.33050/sabda.v3i1.492.

[11] J. Purnomo Y, "Digital marketing strategy to increase sales conversion on e-commerce platforms," Journal of Contemporary Administration and Management (ADMAN), vol. 1, no. 2, pp. 54-62, 2023, doi: 10.61100/adman.v1i2.23.

[12] Q. Liu, H. Tang, L. Wu, et al., "Deep learning-driven e-commerce marketing communication for recommending shopping system and optimizing user experience," Journal of Organizational and End User Computing (JOEUC), vol. 36, no. 1, pp. 1-28, 2024, doi: 10.4018/JOEUC.343258.

[13] B. Rolando and H. Mulyono, "E-commerce as a catalyst for digital economy development: a study of marketing strategies and their impact," Journal of Distribution Science, vol. 23, no. 4, pp. 61-79, 2025, doi: 10.15722/jds.23.04.202504.61.

[14] R. Hasan, "Enhancing market competitiveness through AI-powered SEO and digital marketing strategies in e-commerce," ASRC Procedia: Global Perspectives in Science and Scholarship, vol. 1, no. 01, pp. 465-500, 2025, doi: 10.63125/31tpjc54.

[15] M. Yiming, "Research on innovation of cross-border e-commerce marketing channels," Industrial Engineering and Innovation Management, vol. 6, no. 7, pp. 40-44, 2023, doi: 10.23977/ieim.2023.060706.