

A Study on Marketing Strategies for Precise Analysis of E-commerce Customer Data in the Context of the Digital Economy

Minyue Liang

SHEIN Distribution Corporation, Los Angeles, California, 90015, USA

Keywords: Digital Economy; E-Commerce; Customer Data; Precision Marketing

Abstract: Against the backdrop of the deepening development of the digital economy, the competitive dynamics of e-commerce are undergoing a fundamental shift from the expansion of traffic volume to the in-depth exploitation of data value. Based on a statistical analysis of user behaviour data from e-commerce platforms, this paper systematically deconstructs the value creation mechanism of precise customer data analysis. The research findings reveal that user consumption behaviour exhibits significant geographical concentration, with users from economically developed provinces accounting for a prominent proportion; personalised recommendations based on user profiling can increase the order conversion rate by more than 15 per cent; and conversion funnel analysis has uncovered structural attrition bottlenecks at each stage from ‘selection’ to ‘payment’. However, in practice, enterprises commonly face structural challenges such as data silos, algorithmic bias and privacy compliance. Against this backdrop, this paper proposes the construction of a four-pronged precision marketing system comprising a ‘data foundation, intelligent middle platform, strategy front-end and evaluation closed loop’, providing a theoretical framework and practical guidance for the digital marketing transformation of e-commerce enterprises.

1. Introduction

The digital economy has become the core driver of high-quality economic development in China. In 2025, national online retail sales reached 15.97 trillion yuan, maintaining China’s position as the world’s largest online retail market for the 13th consecutive year. However, the continued expansion of the market has not been accompanied by a proportional increase in corporate profits. The waning of the traffic dividend and rising customer acquisition costs have forced the e-commerce sector to shift from extensive growth to value-driven operations. Against this backdrop, high hopes are pinned on the precise analysis of customer data. In theory, by conducting in-depth analysis of multi-source data-such as user behaviour patterns and transaction records-enterprises can construct detailed customer profiles, implement personalised recommendations and dynamic pricing, thereby significantly boosting marketing efficiency. However, the high costs of data collection have not been effectively translated into actionable marketing initiatives. Drawing on user behaviour data from e-commerce platforms and combining this with empirical industry research, this paper

systematically explores the value creation mechanisms, practical challenges and strategic optimisation pathways associated with the precise analysis of customer data^[1].

2. Application Scenarios and Value Creation through Precise Analysis of E-commerce Customer Data

2.1 Statistical Analysis of User Consumption Behaviour

Statistical analysis of user transaction data from e-commerce platforms can reveal the fundamental characteristics and distribution patterns of user consumption behaviour. Table 1 presents the descriptive statistics based on 20,183 transaction records from a particular e-commerce platform.

Table 1: Descriptive Statistics of E-commerce User Transaction Data

Indicator	Sample Size	Min	Max	Mean	Standard Deviation
Purchase Quantity	20,183	1	82	1.51	1.395
Actual Payment Amount (CNY)	20,183	-300	1,028	117.01	83.735
Shipping Fee (CNY)	20,183	0	18	0.24	1.379

Data source: User transaction data from e-commerce platforms

The above data shows that the average number of items purchased per transaction is approximately 1.51, indicating that most consumers tend to buy single items, whilst bulk purchases are relatively rare; the mean amount actually paid is 117.01 yuan, but the standard deviation is as high as 83.74 yuan, reflecting significant variation in consumers' spending power; the mean postage cost is only 0.24 yuan, suggesting that free delivery has become the norm in e-commerce transactions.

2.2 Characteristics of User Geographical Distribution and Their Implications for Marketing

The geographical distribution of users holds significant decision-making value for targeted marketing on e-commerce platforms. Based on regional distribution reports provided by e-commerce platform user data analysis tools, operators can analyse the distribution of buyers' delivery locations, thereby assessing market capacity in different regions and formulating differentiated product selection and operational strategies. E-commerce big data indicates that the top five provinces by online retail turnover in 2022 were, in order, Guangdong, Zhejiang, Beijing, Shanghai and Jiangsu. The top ten provinces collectively accounted for 86.91 per cent of the nation's online retail turnover, reflecting the high concentration of e-commerce consumption in leading provinces. Guangdong, Jiangsu, Beijing and Shanghai all feature in the top five, constituting the core regions for e-commerce consumption. From a regional perspective, the eastern region accounts for 61.58 per cent of the nation's total number of online shops, 86.88 per cent of sales volume and 89.89 per cent of sales revenue, forming a polarised pattern characterised by 'high density and high output'. In terms of purchasing power, Guangdong, Jiangsu, Beijing and Shanghai all featured among the top ten provinces by purchasing power on JD.com in the first half of 2024, further corroborating the concentration of consumption in these regions. It is worth noting that although the number of internet users in Shanghai and Beijing is only about half that of Jiangsu or Zhejiang, their total online spending ranks among the highest in the country, reflecting their more prominent per capita online spending capacity^[2].

2.3 User Profiling and Quantification of the Effectiveness of Personalised Recommendations

User profiling serves as the technical cornerstone of precision marketing. By constructing user profiling models based on information such as user attributes, preferences and behaviour, and assigning tags to users, important insights can be provided for personalised recommendations and precision marketing. E-commerce platforms can gain an in-depth understanding of user needs by extracting effective features-such as transaction amounts, transaction frequency, time of the most recent transaction and product category preferences-through feature engineering. For a detailed analysis, see the table 2.

Table 2: Quantification of the Effectiveness of Personalised Recommendation Systems

Application Scenario	Performance Indicator	Improvement Margin	Data Source
Taobao API User Portrait Recommendation	User Order Conversion Rate	$\geq 15\%$	Alibaba Cloud Developer Community
Personalized Rearrangement of Homepage Feed	Click-Through Rate (CTR)	50%	Practical Cases of Dagan Data
Personalized Rearrangement of Homepage Feed	GMV Contribution	25%	Practical Cases of Dagan Data
Precise Recommendation for Long-Tail Goods	Exposure Volume	180%	Vertical Footwear E-Commerce Case
Remarketing of Long-Tail Goods	Secondary Reach Conversion Rate	50%	Outdoor Goods E-Commerce Case
Cycle Prediction of High-Frequency Repurchase Products	Repurchase Rate	40%	Daily Necessities E-Commerce Case

It is worth noting that following the launch of RecGPT-Taobao's in-house developed large-scale recommendation model with 10 billion parameters-the recommendation feed powered by this model has achieved double-digit growth in user clicks, whilst both the number of items added to users' wishlists and average session duration have increased by more than 5 per cent. The model is capable of summarising and conducting in-depth analysis of users' consumption behaviour on Taobao spanning over a decade, enabling precise engagement by 'anticipating and providing recommendations even before users are aware of their needs'.

3. Challenges Facing Accurate Analysis of Customer Data

3.1 The Data Silo Effect and Data Quality Dilemmas

Data silos are the primary obstacle hindering the effectiveness of precision marketing. User data for e-commerce enterprises is often scattered across multiple independent platforms, such as transaction systems, customer service systems, logistics systems and social media networks. Inconsistent data standards and incompatible interfaces between these systems make it difficult to piece together a complete picture of user behaviour. A typical example is that data on users' interests and preferences on social media, complaint records in customer service systems, and purchase histories in transaction systems are often stored in different databases, making it difficult to form a unified view of the user. This not only results in user profiles that tend to be one-dimensional but also leads to a lack of coordination in cross-channel marketing engagement. At the same time, data quality varies considerably-issues such as duplicate user information, missing fields and noisy behavioural records are widespread, directly affecting the reliability of user

profiling and the accuracy of recommendation models. More seriously, as much as 86.91 per cent of online retail turnover is concentrated in the top ten provinces. This implies that user data density is extremely high in leading markets, whilst user profiles in long-tail markets are difficult to construct accurately due to sparse data, creating a ‘Matthew effect’ where ‘the rich get richer and the poor get poorer’.

3.2 Algorithmic Bias and the ‘Information Bubble’ Effect

Algorithms are not neutral technological tools; their decision-making logic is highly dependent on the quality of training data and the soundness of feature engineering. Research has found that the fairness of artificial intelligence systems depends on training data, which often reflects and amplifies historical or societal biases. In e-commerce contexts, algorithmic bias manifests in various ways: advertising systems may display differentiated prices based on users’ postcodes, gender or browsing history, leaving users in low-income communities facing fewer choices and higher prices; personalised recommendations may, through over-reliance on historical behaviour, lock users into a limited range of product categories, creating an ‘information bubble’. More subtly, research published in “Nature” in 2025 indicated that AI systems are not only biased in their output but also in the way they evaluate content; content generated by AI is prioritised by other AI systems, creating a feedback loop of ‘AI bias’. This “discriminatory feedback loop” means that once an algorithm develops a bias, it will continuously reinforce itself over time, ultimately leading to a severe misallocation of marketing resources. In e-commerce practice, this implies that if a particular user group is “overlooked” or “misjudged” by the algorithm due to insufficient historical data, this bias will continue to amplify in subsequent recommendation iterations, resulting in an irreversible “algorithmic lock-in”.

3.3 The Compliance Tension Between Privacy Protection and ‘Big Data Price Discrimination’

Privacy protection has become an insurmountable red line for targeted marketing. With the ongoing implementation of the Personal Information Protection Law and the introduction of the EU AI Act, e-commerce platforms face increasingly stringent compliance requirements when collecting, storing and using user data. In 2024, the US Federal Trade Commission launched its ‘AI Compliance Initiative’, focusing on cracking down on companies that mishandle consumer data. Data shows that in 2025 alone, there were nearly 67,000 phishing attacks targeting e-commerce platforms, whilst the number of individual users in the retail and e-commerce sectors affected by ransomware attacks increased by 152 per cent compared to two years prior. The rising risk of data breaches has left businesses facing a difficult balancing act between data utilisation and security protection^[3].

Particularly problematic is the issue of ‘big data price discrimination’, whereby platforms use consumers’ stored personal information and behavioural data to create user profiles, identify their ‘reservation price’, and implement differential pricing accordingly. In practice, such behaviour manifests in various forms, including ‘higher prices for regular customers’, ‘prices rising the more you browse’, and ‘price increases for repeat purchases’, and has been characterised in judicial practice as an infringement constituting ‘abuse of algorithmic power and infringement of consumers’ rights to personal information’. In 2025, iTutorGroup paid a settlement of US\$365,000 after its AI recruitment software was found to have automatically rejected female job applicants aged 55 and over and male applicants aged 60 and over. This case serves as a warning that the legal risks of algorithmic discrimination have moved from theoretical speculation to actual judicial rulings. For e-commerce companies, striking a balance between the commercial efficiency of personalised recommendations and the compliance requirements for fair trading has become a strategic challenge

that must be addressed head-on.

4. Exploring Pathways for Optimising Marketing Strategies

4.1 Breaking Down Silos and Improving Data Quality

A data foundation is the cornerstone of precision marketing. Organisations should establish unified standards for data collection and governance, integrate data channels from multiple sources-including transactions, logistics, customer service and social media-and construct a comprehensive view of user behaviour. At the technical level, a Customer Data Platform (CDP) can be introduced as a unified data hub to standardise and consolidate user data scattered across various business systems, thereby forming a data asset system centred on a 'unified user ID'. For long-tail markets characterised by sparse data, technologies such as third-party data supplementation and federated learning can be utilised to achieve collaborative cross-domain modelling whilst safeguarding user privacy. At the same time, a routine data quality monitoring mechanism should be established to automatically detect and cleanse missing fields, duplicate records and outliers (such as the negative amount records in Table 1), ensuring the reliability of the 'raw materials' fed into profiling models.

4.2 Empowering Algorithmic Fairness and Explainability

The intelligent middle platform serves as the 'capability layer' connecting underlying data with front-end marketing strategies; its core task is to transform raw data into actionable marketing insights. However, the construction of the intelligent middle platform must not merely pursue the optimisation of quantitative metrics such as recommendation accuracy and conversion efficiency; it must also incorporate algorithmic fairness and explainability as equally important design principles.

Enterprises should build an algorithmic middleware platform equipped with 'explainable AI' capabilities, ensuring that recommendation logic and pricing decisions possess technical characteristics that are auditable and traceable. Explainability not only meets compliance requirements but also holds direct commercial value: when users can understand the logic behind recommendations, their trust in and acceptance of the recommended content increase significantly. To address the issue of algorithmic bias, a mechanism for algorithmic compliance auditing should be introduced. Algorithm compliance audits differ from traditional technical testing; their core task is not merely to verify the technical performance of algorithms, but also to assess their potential impact on the rights and interests of various stakeholders (particularly vulnerable groups), and to identify the algorithmic and training data characteristics that give rise to these impacts. Specifically, e-commerce platforms should regularly conduct fairness testing on user profiling algorithms, personalised recommendation algorithms and dynamic pricing algorithms to detect whether there are systematic discriminatory outputs based on sensitive attributes such as geographical location, gender, age and spending power. For example, 'counterfactual fairness tests' could be designed-where all other conditions remain constant whilst only one sensitive attribute of a user is altered-to observe whether there are significant differences in the algorithm's output. If the differences exceed a reasonable threshold, this indicates that the algorithm may be biased, necessitating further scrutiny of the representativeness of the training data and the neutrality of feature engineering.

4.3 Optimising the Strategy Front End

The strategy front end is the final layer through which marketing value is realised; its core task is

to translate data insights into actionable marketing initiatives. Building on a robust data foundation and an intelligent middle platform, enterprises can achieve three-tiered, progressive strategy optimisation. The first tier is customer segmentation. Similar customers are categorised into distinct groups based on purchase history, demographic and behavioural data, and promotional messages are tailored accordingly. For example, users can be categorised into different segments such as high-value repeat buyers, price-sensitive customers, new product explorers and dormant or at-risk customers, each matched with a differentiated marketing strategy. The second layer is personalised recommendations. Machine learning algorithms are used to identify purchasing patterns, pushing recommendations for related products when users browse specific items to facilitate cross-selling and upselling. The third layer is dynamic pricing and promotional optimisation. By analysing competitive pricing, demand elasticity and stock levels in real time, pricing strategies are automatically adjusted to ensure the right users are reached at the right time and at the right price. Empirical evidence shows that dynamic advertising strategies based on first-party data and purchase intent can increase ROAS (Return on Advertising Spend) by over 38 per cent.

4.4 Establishing a Closed-Loop Evaluation System

A closed-loop evaluation system is essential for the continuous evolution of a precision marketing framework. Organisations should establish an evaluation framework covering three dimensions: ‘efficiency’, ‘effectiveness’ and ‘compliance’. The efficiency dimension focuses on process metrics such as marketing response rates, click-through rates and order conversion rates, which are used to monitor the quality of strategy execution in real time; the effectiveness dimension measures outcome metrics such as customer lifetime value, repeat purchase rates and GMV contribution, which are used to assess the long-term commercial value of strategies; the compliance dimension tracks risk indicators such as algorithmic fairness scores, user complaint rates and the number of data breach incidents, ensuring that marketing campaigns do not cross compliance red lines.

Companies can continuously validate the effectiveness of their strategies through A/B testing, feeding the test results back into the data foundation and algorithmic models to form a complete closed-loop comprising ‘data, insights, strategy, evaluation and optimisation’. Within this closed loop, evaluation should not be viewed as the end point, but rather as the starting point for a new round of data collection and strategy optimisation. User feedback data generated from every marketing engagement is fed back into the data foundation, driving the dynamic updating of user profiles and the continuous iteration of recommendation models. Only in this way can a precision marketing system break free from the limitations of ‘one-off development and static operation’, and continually approach optimal marketing decisions through continuous closed-loop feedback.

5. Conclusion

Against the backdrop of the digital economy’s development, the competitive logic of e-commerce marketing is shifting from the expansion of traffic scale to the in-depth mining of data value. Based on statistical analysis of e-commerce user behaviour data, this paper systematically deconstructs the value creation mechanisms, practical challenges and optimisation pathways of precise customer data analysis. Research indicates that the geographical concentration of user consumption behaviour and the differentiation in purchasing power constitute fundamental constraints on precision marketing, whilst personalised recommendations based on user profiling have demonstrated significant efficiency gains across multiple scenarios. However, if structural challenges such as data silos, algorithmic bias and privacy compliance are not effectively resolved, they will severely constrain the realisation of data value. To this end, the four-pronged precision

marketing system proposed in this paper-comprising a ‘data foundation, intelligent middle platform, strategy front end and evaluation closed loop’-aims to provide an actionable theoretical framework for e-commerce enterprises to achieve digital marketing transformation whilst navigating the dual constraints of efficiency and compliance.

References

- [1] Shaban A A, Zeebaree S R M .*A comprehensive review of digital marketing strategies for e-commerce success*[J]. *International Journal of Scientific World*, 2025, 11(1):66-82.DOI:10.14419/th8ey858.
- [2] Rolando B, Mulyono H .*E-Commerce as a Catalyst for Digital Economy Development: A Study of Marketing Strategies and Their Impact*[J].*Journal of Distribution Science*, 2025, 23(4):61-79.DOI:10.15722/jds.23.04.202504.61.
- [3] Waseem, Abbasi M U, Hamid F .*Impact of Digital Marketing Strategies on Marketing Performance: Role of Social Media Performance and E-commerce*[J].*International Journal of Management Research & Emerging Science (IJMRES)*, 2024, 14(3).DOI:10.56536/ijmres.v14i3.656.