

Application of Big Data and Statistical Modeling to Enable Consumer Behavior Analysis in the Digital Economy Landscape

Haoran Shen

Nanjing University of Finance and Economics, Nanjing 210000, Jiangsu, China

Abstract: The rapid development of digital economy has changed the consumer behavior and market operation logic, and the traditional research methods can no longer cope with the diverse data sources and behavioral dynamics. This paper discusses the application of big data and statistical modeling in consumer behavior analysis, firstly, analyzing the change of consumer behavior in digital economy, including the evolution of information acquisition, decision path and platform interaction mechanism; secondly, sorting out the practical application of big data in user profiling, sentiment analysis and behavior prediction; finally, combining multiple regression, time series and machine learning methods, mining the data laws behind consumer behavior. The study finds that data-driven and intelligent modeling is reconstructing the paradigm of consumption research, providing strong support for enterprise marketing optimization and policy making.

Keywords: digital economy; consumer behavior; big data analysis; statistical modeling; user profiling

1. Introduction

With the rapid development of digital technology, the consumer behavior mode has undergone a profound change, which is no longer a linear and passive decision-making process, but a dynamic and personalized behavioral path deeply intervened by data, platforms, and algorithms[1]. At the same time, the integration and application of big data technology and statistical modeling methods provide a new methodological support for the in-depth excavation of consumer psychology, the prediction of market trends, and the optimization of marketing strategies[2]. Traditional consumer behavior research focuses on subjective methods such as questionnaires and interviews, and suffers from the problems of underrepresentation and poor dynamics. The emerging digital technology, especially the breakthroughs in data collection, behavioral analysis, model prediction, etc., opens up the possibility of constructing a more realistic and accurate consumer behavior research system[3]. Therefore, this paper will focus on the application of big data and statistical modeling technology in the analysis of consumer behavior in the era of digital economy, and try to explore the mechanism and realization path of its power from the theoretical and technical level.

2. New features of consumer behavior in the context of digital economy

2.1 The development trend and connotation evolution of digital economy

Digital economy not only refers to the new economic form formed by relying on digital technologies such as the Internet, cloud computing, big data, artificial intelligence, etc., but also an economic paradigm that profoundly reshapes the way of resource allocation and the logic of value creation. In China, the digital economy has been incorporated into the national strategic development plan, and its proportion in GDP continues to rise. Especially in retail, e-commerce, finance, entertainment and other consumer sectors, digitalization has shifted from “auxiliary tool” to “infrastructure”, profoundly affecting consumers' behavioral choices and paths. During the evolution of digital economy, new features such as platform-type enterprises, data factor market and algorithm-driven mechanism are constantly emerging, injecting brand-new theoretical and technological impetus into the study of consumer behavior.

2.2 Digital Transformation of Consumer Behavior Models

Traditional consumer behavior models are mostly based on offline environments, emphasizing the logical processes of information acquisition, motivation formation, comparative evaluation and purchase decision. However, in the context of digital economy, this linear model is gradually replaced by fragmented, instantaneous and social behavior paths[4]. Consumers are no longer passive individuals receiving information, but “digital nodes” actively participating, interacting and spreading. For example, new consumption scenarios such as short video seeding, e-commerce live broadcasting, and social platform sun ordering have reconfigured the allocation of consumers' attention and the basis for decision-making.

At the same time, data feedback and algorithmic recommendations continue to shape user preferences, making consumer behavior more irrational and emotional. Understanding this transformation trend is an important prerequisite for constructing a new theoretical model of consumer behavior.

2.3 The rise of personalized consumption driven by digital platforms

The rapid rise of digital platforms, especially the popularity of e-commerce companies such as Taobao, Jingdong, Pinduoduo, and Douyin, has made the information environment faced by consumers highly personalized and complex. Based on big data analysis and user profiling technology, platforms are able to realize product recommendation and precise marketing for thousands of people. This personalized experience greatly improves consumer satisfaction, but also inadvertently exacerbates consumer stickiness and decision-making dependence. More importantly, every browse, click and purchase of consumers on the platform will be recorded by the system and used to optimize its subsequent content delivery, realizing the transformation from “products looking for people” to “people looking for products”. The algorithmic mechanism of digital platforms not only improves transaction efficiency, but also quietly influences consumers' cognitive judgments and values, and it is worthwhile to explore its behavioral consequences in depth[5].

2.4 Reshaping Consumer Decision-making Paths by Emerging Technologies

Supported by technologies such as big data, artificial intelligence and the Internet of Things, the decision-making path of consumers is undergoing unprecedented reconstruction. Take the intelligent recommendation system as an example, it can realize personalized product push based on multi-dimensional information such as user's historical behavior, interest and preference and even geographic location, so that consumers can get the products they need without actively searching for them. In addition, voice assistant, virtual dress mirror, AR shopping guide and other technologies are gradually penetrating the daily consumption scene, changing the traditional “cognition-comparison-decision-making” mode. The starting point of decision-making no longer begins with the subjective needs of consumers, but with the potential interest “triggered” by the system.

3. The application of big data technology in the analysis of consumer behavior in practice

3.1 The value of big data collection and management in consumer research

The arrival of the big data era has provided unprecedentedly rich data resources for consumer behavior analysis. From online transaction records and social platform interactions to offline behavioral trajectories and member consumption history, various types of data present high frequency, high capacity and diverse characteristics. Compared with the traditional questionnaire or interview method, such objective and real-time data records can reflect consumer behavior more realistically[6]. Scientific data collection not only relies on technical tools, such as data crawling, sensor technology and API interfaces, but also must consider the legality and ethical boundaries. At the same time, data management becomes a key aspect that affects the quality of research, including data cleaning, integration, standardization and storage. Only under a good data governance framework can consumer behavior research have a reliable foundation, avoiding the dilemma of “data-rich but information-poor”.

3.2 User Image Construction and Behavioral Characteristics Mining Technology

User portrait is one of the core tools of big data consumption analysis, the essence of which is to transform massive scattered data into a structured, labeled collection of user information. Through statistical methods such as cluster analysis, association rule mining, principal component analysis, etc., a multi-level portrait can be constructed from the dimensions of consumers' gender, age, interest and preference, purchasing power, and behavioral trajectory. This not only helps precision marketing, but also can be used to segment user groups and optimize product positioning and service strategies. In commercial practice, portrait models are often used in conjunction with recommender systems to achieve personalized recommendations and customized services. Researchers, on the other hand, continuously optimize the model structure and algorithmic parameters to improve the fit and prediction of real user behavior.

3.3 Social Media and Sentiment Analysis to Assist Consumption Insight

In the context of increasingly socialized digital consumption scenes, social media has become an important data source for insight into consumer psychology. User comments, likes, retweets and pop-ups on Weibo, WeChat, Xiaohongshu, Douyin, and other platforms contain a wealth of emotional information. Sentiment analysis technology is a natural language processing tool developed on this basis, which is able to recognize positive, negative and neutral emotions in text content,

providing powerful support for brand evaluation, product feedback, user satisfaction and other research. For example, a hybrid model that combines sentiment lexicon and machine learning methods can realize automated classification and sentiment polarity judgment of user reviews. By monitoring the emotional trend, enterprises can timely adjust product strategies and marketing tactics, thus enhancing consumer trust and loyalty.

3.4 Real-time data-driven dynamic consumption prediction

Compared with traditional statistical analysis, real-time data provides a more timely and dynamic perspective for consumption prediction. By capturing consumers' immediate behavioral changes, such as the amount of hot topic searches, the participation rate of spike activities, and online access paths, short-cycle, high-frequency prediction models can be constructed. In terms of technical implementation, streaming data processing frameworks (e.g., Apache Kafka, Spark Streaming) make it possible to support real-time decision-making with virtually no delay from data collection to modeling. At the same time, algorithm selection also tends to be lightweight and iterative, facilitating quick response to fluctuations in the consumer market. For example, a hybrid approach combining ARIMA and LSTM models can balance short-term volatility and long-term trends to improve forecast accuracy. This data-driven forecasting mechanism is becoming an important means for companies to improve operational sensitivity and resilience.

4. Statistical Modeling Methods for Accurate Analysis of Consumer Behavior

4.1 Application of Multiple Regression and Classification Models in Consumption Decision Research

As the basic tools of statistical analysis, multiple regression and classification models play a fundamental role in consumer behavior research. By constructing regression models with willingness to consume, purchase amount, brand loyalty, etc. as dependent variables, researchers can quantify the degree of influence of explanatory variables (e.g., price sensitivity, promotional responsiveness, advertisement contact frequency) on behavioral outcomes. Meanwhile, categorical models such as Logistic regression, discriminant analysis, and decision trees are used to predict binary or multi-category decision-making problems such as whether a consumer will purchase or not, or whether he/she will repurchase or not. In practical research, the modeling needs to fully consider the multicollinearity, heteroskedasticity and data normality of the variables to improve the goodness-of-fit and inferential power. In addition, cross-validation and robustness testing are also key steps to ensure the reliability of the model, so that the analysis results are more convincing and practical guidance.

4.2 Consumption Trend Analysis with Time Series and Panel Data Modeling

Consumption behavior has obvious time dynamics and regional heterogeneity, so time series models and panel data models have become important tools for portraying consumption trends. Time series analysis, such as ARIMA, SARIMA, etc., is suitable for capturing cyclical, seasonal and sudden fluctuations in consumption data, and is used to predict consumption behavior under specific events, such as holiday sales and traffic during promotional periods. Panel data models, on the other hand, integrate cross-sectional and time-series dimensions, can control for both individual differences and temporal variations, and are widely used in cross-city, cross-channel, or cross-category consumption studies[7]. Through the fixed-effects and random-effects models, we can analyze the differences in consumption preferences and responses among different groups, providing a strong basis for market segmentation and policy formulation.

4.3 Machine Learning Algorithms and Deep Models for Consumption Pattern Recognition

With the explosive growth of data volume, traditional statistical models face many limitations when dealing with non-linear, high dimensional and unstructured data. For this reason, machine learning and deep learning models have gradually become an important tool for consumer behavior pattern recognition. Algorithms such as Random Forest, Support Vector Machine (SVM), and K Nearest Neighbor Algorithm (KNN) can achieve automatic classification and behavior prediction of consumers. Deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM) can handle complex data types such as images, text and time series, and are especially suitable for personalized recommendation systems and context-aware consumer scenarios. Although such models have certain "black box" problems, their transparency and practicality are gradually improving through interpretable techniques (e.g., SHAP value analysis).

4.4 Visualization of Modeling Results and Business Decision Transformation

The ultimate goal of statistical modeling is not only to analyze, but also to provide actionable insights for business decisions. Therefore, the visualization of results and strategy transformation are crucial[8]. With visualization tools such as Tableau, Power BI, and ECharts, researchers can present complex model results in interactive charts, heat maps, path

diagrams, etc., making it easy for non-specialists to understand and participate in decision-making. Meanwhile, the design of dashboards and visualization reports should focus on user experience, highlighting key indicators and trend changes to help management react quickly. In addition, the behavioral grouping, scene setting and marketing trigger mechanism based on the modeling results should be connected with the CRM system and marketing automation platform to form a closed-loop system of “insight-decision-making-execution-feedback”, so as to truly realize data-driven and refined operation. This will truly realize data-driven and refined operation.

5. Conclusion

In summary, in today's rapidly developing digital economy, big data not only enriches the data source of consumer behavior research, but also promotes the development of real-time, personalized and predictive analysis methods. In addition, the combination of statistical modeling and artificial intelligence has continuously improved the accuracy and explanatory power of consumer behavior research. In the future, research should pay more attention to data ethics, model interpretability, and interdisciplinary integration in order to drive consumer behavior research toward more scientific, transparent, and intelligent development.

References

- [1] Sun Q. Research on consumer behavior analysis and prediction model based on artificial intelligence[J]. Market Outlook,2024,(23):34-36.
- [2] HONG Yongmiao, WANG Shouyang. Big data, machine learning and statistics: challenges and opportunities[J]. Journal of Econometrics,2021,1(01):17-35.
- [3] Feng ZY. Analysis of Consumer Behavior Reinvented by Big Data-Driven Precision Marketing Strategy[J]. Business Culture,2024,(20):116-118.
- [4] Liu Y. Analysis of consumers' consumption behavior of food in the context of digital economy - based on e-commerce model[J/OL]. Frontiers of Engineering Management Science and Technology,1-12[2025-04-10].
- [5] Zhang Dian, Chen Dan. Consumer Behavior Analysis and Personalized Marketing Practice Driven by Big Data[J]. Modern Business,2024,(21):51-54.
- [6] Lamberton C, Stephen A T. A thematic exploration of digital, social media, and mobile marketing: research evolution from 2000 to 2015 and an agenda for future inquiry[J]. Journal of marketing, 2016, 80(6): 146-172.
- [7] Wedel M, Kannan P K. Marketing analytics for data-rich environments[J]. Journal of marketing, 2016, 80(6): 97-121.
- [8] Davenport T, Guha A, Grewal D, et al. How artificial intelligence will change the future of marketing[J]. Journal of the Academy of Marketing Science, 2020, 48: 24-42.

Author Bio

Haoran Shen (2004.8), male, the Han nationality, Born in Suqian, Jiangsu province, Education Background: Bachelor, Research Interest: statistic model.