Individualized Marketing and Consumer Behavior Analysis Driven by Big Data Algorithm

Jiao Wang[1,2], Bo Ma[3] and Linzhi Nan[4]

[1] Hao Jing College of Shaanxi University of Science and Technology, Xi'an, Shaanxi 712046, China, WJ15229718361@163.com
[2]Bangkok Thonburi University, Bangkok 10170, Thailand, WJ15229718361@163.com
[3]Xi'an Shendian Electric Co., Ltd., Xi'an, Shaanxi 712046, China, Maxwell_8813@163.com
[4]Universiti Utara Malaysia, Kedah Darul Aman 06010, Malaysia, nan_linzhi@cob.uum.edu.my

Corresponding author: Linzhi Nan, nan_linzhi@cob.uum.edu.my

Abstract. Driven by big data (BD), in order to meet people's consumption needs, enterprises must further develop and formulate new marketing strategies to meet consumers' needs. This requires enterprises to use data to analyze consumers and study consumer behavior. This article puts forward a computer-aided individualized marketing and consumer behavior analysis model based on BD algorithm, which helps enterprises to formulate precise marketing strategies based on consumer behavior characteristics on the basis of in-depth understanding of consumer needs. The experimental results show that the recall rate of this algorithm is above 86%, which is more than 20% higher than that of decision tree (DT) algorithm. The accuracy rate is over 95%, which is more than 10% higher than the DT algorithm. These data show that this algorithm can predict consumers' consumption tendency more comprehensively and accurately, and provide more accurate and comprehensive support for individualized marketing decisions. Through in-depth study and application of this model, enterprises can better understand consumer demand and market trends, and improve marketing effect and market competitiveness. At the same time, the government and enterprises should also pay attention to protecting consumers' rights and fair competition in the market, so as to promote the healthy development of the marketing field.

Keywords: Big Data; Individualized Marketing; Consumer Behavior; CAD
DOI: https://doi.org/10.14733/cadaps.2024.S13.165-178

1 INTRODUCTION

The emergence of BD has invisibly changed people's way of thinking and lifestyle, and the difficulty of data processing will also greatly increase. In the process of development, BD has various characteristics such as capacity, variety, speed, variability, authenticity, and is developed in the form of structured, semi-structured, and unstructured data structures, which is beneficial for enterprises to conduct more in-depth research on people's consumption behavior in society and has a profound impact on the development of consumer behavior and precision marketing. Provide customized product choices for consumers based on their needs and health standards. For example, launching food or daily necessities that meet specific nutritional needs. Establish a membership system to provide members with exclusive discounts, point redemption and other benefits to increase customer stickiness. Meanwhile, Chandra et al. [1] conducted membership data analysis to gain a deeper understanding of consumers' shopping habits and needs, providing strong support for product development and marketing strategies. Promote the concept of healthy shopping through online advertising, social media promotion, and other means, guiding consumers to pay attention to information on product quality, nutritional value, and environmental protection. In addition, regular health life lectures or cooking activities can be held to increase consumers' awareness and interest in healthy food. Ensure that the website interface is concise and clear, and product information is clear and easy to understand. Provide convenient shopping processes and diverse payment methods to meet the different needs of consumers. In this context, deep mining and analysis through BD can help enterprises study consumer behavior, provide reliable and accurate information for subsequent precision marketing, and find the best marketing methods and means.

With the continuous development of internet technology, the application of omnichannel marketing systems in various industries is becoming increasingly widespread. How to achieve precision marketing is a core issue in this system. The shared kitchen platform, as an emerging model in the catering industry in recent years, aims to integrate catering resources through the concept of shared economy, and provide consumers with higher quality and convenient food services. Chiu and Chuang [2] take the shared kitchen platform as an example to explore the solutions and effects of applying transfer learning to achieve precision marketing in omnichannel systems. Transfer learning is a machine learning method that applies knowledge learned in one task or domain to another different task or domain. In omnichannel marketing systems, transfer learning can be applied to multiple aspects such as user behavior analysis, product recommendation, and price prediction. More and more scholars are paying attention to the application of transfer learning in omnichannel marketing and have achieved certain research results. In the era of BD, data has become the core competitiveness of enterprises and an important capital. With the support of data, enterprises can truly benefit from an invincible position. Multi data research will also bring about changes in corporate culture, organizational models, and even business operations. In order to meet people's consumption needs, enterprises must further develop and formulate new marketing strategies to meet consumer needs. Darsono et al. [3] explore the strategic policies of small and medium-sized enterprises for marketing through e-commerce. Small and medium-sized enterprises need to keep up with the trend of the times, constantly learn and innovate to adapt to the changes and development needs of the e-commerce market. By utilizing e-commerce for marketing, small and medium-sized enterprises can better enhance brand awareness, increase sales, and reduce operating costs. Driven by both policy support and market opportunities, small and medium-sized enterprises need to formulate appropriate strategic policies in areas such as clear brand positioning, optimized website design, reasonable channel selection, and strengthened customer service to adapt to the development and changes of the e-commerce market. This article is based on consumer behavior research and explores how enterprises can reasonably utilize modern information technology in the current era of BD, re-examine consumer behavior characteristics, and develop more efficient and precise marketing strategies.

In today's market competition, consumer demand is changing faster and faster, and companies must constantly update their marketing strategies to adapt to consumer needs. Traditional marketing strategies are often based on experience and intuition, making it difficult to achieve precise marketing. The computer-assisted individualized marketing and consumer behavior analysis model based on BD algorithms can help enterprises analyze massive amounts of data, mine consumer behavior characteristics and potential needs, and develop precise marketing strategies based on consumer behavior characteristics. In the era of social media, we live in a world full of sympathy and compassion. On social media, people often evoke sympathy from others by sharing their own difficulties. However, this approach often only brings short-term attention and assistance. Once attention decreases, people often forget their initial commitments, resulting in a lot of effort being wasted. In contrast, personalized career related marketing that goes beyond compassion refers to providing personalized solutions to audiences by understanding their needs and interests. This marketing approach not only focuses on the pain of the audience, but also on how to help them overcome difficulties. By providing practical help and advice, this marketing approach can better stimulate people's action and encourage them to actively participate in the process of solving problems [4]. This article constructs a computer-assisted individualized marketing and consumer behavior analysis model based on BD algorithms. Data mining and analysis techniques are used to deeply mine and analyze massive data, extract consumer behavior characteristics and potential needs, and perform multi-dimensional analysis and visual display of consumer data. Through this approach, it can effectively enhance the influence of the enterprise, enhance consumer loyalty, and play an important role in enhancing the core competitiveness of the enterprise.

By analyzing massive amounts of data, enterprises can accurately understand consumers' needs and behavioral characteristics, thereby formulating more accurate marketing strategies. By analyzing consumer behavior, enterprises can understand consumers' needs for products and services, thereby optimizing the design and functionality of products and services, and improving consumer satisfaction. A computer-assisted individualized marketing and consumer behavior analysis model based on BD algorithms can help enterprises achieve automated marketing, save labor and material costs, and improve marketing efficiency. Through precise marketing and optimization of products and services, enterprises can better meet the needs of consumers, improve market share and competitiveness. In today's market competition, enterprises must constantly update their marketing strategies to meet the needs of consumers. The computer-aided individualized marketing and consumer behavior analysis model based on BD algorithms can help enterprises better understand consumer needs, formulate precise marketing strategies, and improve market competitiveness. The main innovations of the study are as follows:

⊖ This article explores the research methods and application scenarios of computer-aided individualized marketing and consumer behavior analysis models from the perspective of BD algorithms.

⊜ This article constructs a computer-assisted individualized marketing and consumer behavior analysis model based on BD algorithms, helping enterprises analyze massive amounts of data, mine consumer behavior characteristics and potential needs, and develop precise marketing strategies based on consumer behavior characteristics.

⊛ By integrating and analyzing data from different sources, it is possible to more accurately understand consumers' needs and behavioral characteristics, and improve the targeted and effective marketing strategies.

The chapters of this article are arranged as follows:

The first section is an introduction, introducing the concept and importance of individualized marketing, and proposing the research purpose and issues of this article; The second section is a review of relevant research, introducing relevant research in the field of individualized marketing, and proposing the research methods and innovative points of this article; The third section constructs the marketing analysis model of this article; The fourth section is algorithm and test analysis, presenting experimental results, including recall and accuracy indicators, and analyzing and discussing the results; The fifth section is the conclusion, summarizing the experimental results and main contributions, and providing prospects for the future development of individualized marketing.

2 RELATED WORKS

With the rapid development of e-commerce, intelligent digital marketing has become the key for enterprises to gain competitive advantages. Emotional analysis and business intelligence frameworks play a crucial role in intelligent digital marketing. Kyaw et al. [5] provide a detailed introduction to the composition, application scenarios, and advantages of sentiment analysis business intelligence frameworks, and look forward to their future development trends. The sentiment analysis business intelligence framework is an intelligent marketing tool that integrates data collection, data cleaning, sentiment analysis, and business intelligence display. This framework captures, integrates, and analyzes text data of consumers on social media, comment websites, e-commerce platforms, and other channels. Explore the true needs, attitudes, and behavioral patterns of consumers, and provide precise marketing strategy recommendations for enterprises. With the continuous development of internet technology, e-commerce enterprises have gradually become an important force in the global trade market. However, behind the rapid development, e-commerce enterprise marketing management is facing many practical difficulties. Li [6] explores these dilemmas, main motivating factors, and solutions, aiming to provide reference for the healthy development of e-commerce enterprises. The consumer behavior in the e-commerce market is becoming increasingly complex and variable, influenced by various factors, such as economic conditions, social culture, technological progress, etc. This makes it difficult for e-commerce enterprises to adjust their marketing strategies in a timely manner to meet the needs of consumers. The entry threshold for the e-commerce market is relatively low, and more and more enterprises are entering this field. The marketing costs of e-commerce enterprises are constantly increasing, including expenses for advertising placement, promotional activities, search engine optimization, and other aspects.

In today's highly informationized business environment, algorithms have become an important tool for enterprise operation and management. Li and Li [7] explored the application of victory algorithms in consumer manipulation, personalized pricing, big data management, and manufacturing and service operation management, and analyzed their impact on enterprise competitiveness. Consumer manipulation mainly refers to enhancing consumer experience and conversion rate through algorithm design. For example, many e-commerce platforms use recommendation algorithms to recommend related products to users based on their shopping history, browsing history, and other personal information, in order to improve purchase conversion rates. In addition, algorithm applications such as optimizing website design and simplifying shopping processes can further enhance the consumer experience. Data has become the core resource of business competition. In the context of big data, machine learning technology has become a powerful tool in many fields due to its powerful data processing and analysis capabilities. Especially in the field of user behavior analysis and investigation, machine learning technology can help enterprises better understand consumer needs, optimize products and services, and improve market competitiveness. Martín et al. [8] explore the current models and applications of user behavior analysis surveys based on machine learning technology. User behavior analysis survey is a method of collecting and analyzing user data to understand user needs, behavioral preferences, and purchasing habits. Its advantage lies in helping enterprises better understand consumer demand, predict market trends, optimize products and services, and improve market competitiveness. At the same time, through user behavior analysis and investigation, enterprises can also timely detect market changes and competitive trends, providing strong support for decision-making.

Statistical model is a prediction method based on historical data to establish mathematical models. Sharma and Dalip [9] predicted future consumer behavior by analyzing historical sales data, demographic information, and more. Market research is a prediction method that uses surveys, interviews, and other methods to understand consumer demand, purchase intention, and other information. Able to obtain first-hand information from consumers and suitable for qualitative analysis. Due to the significant impact of sample selection and survey methods, there may be deviations in the results. By training a large amount of data, machine learning algorithms can automatically identify consumer behavior patterns and make predictions. Statistical models and market research have high accuracy in predicting consumer behavior, but require significant investment in data collection and processing. Machine learning algorithms and natural language processing technologies can automatically process and analyze large amounts of data, and provide real-time consumer behavior prediction, but require sufficient data for training and parameter adjustment. In today's digital age, e-commerce personalization has become a shopping experience pursued by more and more consumers. For privacy conscious consumers, how to protect personal privacy while enjoying personalized recommendations has become an important issue of concern.

Song et al. [10] conducted a survey on the personalization of e-commerce among privacy conscious consumers, exploring its current status, advantages, challenges, and future development trends. E-commerce personalization refers to providing personalized product recommendations and services to consumers by collecting and analyzing their shopping history, browsing history, interest preferences, and other information. This personalized recommendation can help consumers find their desired products more quickly based on multiple dimensions such as shopping behavior, preferences, and needs. At the same time, personalized e-commerce can also help improve merchants' sales and customer satisfaction.

With the advent of the digital age, consumers' demands and expectations for brand marketing communication are also constantly upgrading. In order to stand out in the fierce market competition, brands need to use advanced technological means to achieve personalized marketing communication. Timokhovich and Bulycheva [11] introduced a technology that utilizes artificial intelligence to achieve personalized brand marketing communication, helping businesses better establish connections with target audiences. The importance of personalized brand marketing communication is increasingly prominent. Modern consumers pay more attention to personalized experiences and needs, and they hope to receive customized content tailored to their interests and needs in brand marketing communication. The rapid development of artificial intelligence technology has provided strong support for personalized brand marketing communication. When posting content on social media, it is important to ensure that the content is interesting, valuable, and can attract the attention of the target audience. You can publish game screenshots, videos, guides, player feedback, and other content. At the same time, it is important to pay attention to interacting with the target audience, replying to comments and private messages, and paying attention to their dynamics in order to better understand their needs and feedback. Collaborating with KOL (opinion leaders) and internet celebrities is another effective way to promote computer games. These people have a certain influence and fan base on social media, and Wawrowski et al. [12] collaborate to increase the exposure and popularity of the game. Invite game anchors to conduct live streaming trials, or collaborate with internet celebrities to create game guides and share insights. Social media marketing has become an important means of promoting computer games. To do a good job in social media marketing, it is first necessary to understand the characteristics and needs of the target audience, and choose appropriate social media platforms and tools. Finally, it is important to be adept at utilizing KOL and internet celebrities for collaborative promotion to expand the influence and audience of the game. With the advent of the big data era, significant progress has been made in real estate valuation research based on feature price models. This model is increasingly widely used in the field of valuation, providing strong support for evaluating the value of real estate. Wei et al. [13] will explore the significance, previous research, research methods, discussion of results, and conclusions and prospects of real estate valuation based on feature price models in the era of big data. The arrival of the big data era has provided massive data support for real estate valuation research. The research on real estate valuation based on feature price models aims to analyze the feature price data of real estate, explore the rules and trends within it, and provide a basis for accurate valuation of real estate. By utilizing big data technology and machine learning algorithms, researchers can more comprehensively consider the impact of various factors on the value of real estate, thereby improving the accuracy and reliability of valuation.

Wei et al. [14] analyzed consumer purchasing behavior and found that consumers browse and compare a large amount of products before making a purchase. Therefore, enterprises need to optimize their product recommendation algorithms and search functions in order to better meet the needs of consumers. In addition, consumers are more inclined to choose reputable merchants and products with good reviews when making purchases. Through blockchain security technology, the enterprise has discovered that its trading platform has some security risks, such as account theft and payment information leakage. In response to these issues, enterprises need to strengthen account security protection and payment security protection, and improve consumer trust and satisfaction. By utilizing blockchain security technology, the enterprise has achieved anonymization and encryption processing of consumer personal information. While protecting consumer privacy, it enhances the credibility and security of transactions. In today's market environment, consumer behavior has an increasingly significant impact on market structure. With the development of technology, the emergence of intelligent models for identifying market structures has provided new perspectives and methods for monitoring product competition. Zhan et al. [15] introduced how to use consumer behavior analysis and intelligent model recognition technology to monitor product competition, aiming to provide reference for enterprises to optimize product strategies and enhance competitiveness. Consumer behavior refers to the reactions and attitudes exhibited by consumers in the process of acquiring, purchasing, and using products or services. The formation of consumer behavior is influenced by various factors. Such as personal characteristics, economic status, social culture, etc. Through in-depth analysis of consumer behavior, we can understand consumers' needs, preferences, and behavioral patterns, thereby providing a basis for the formulation of product strategies. The VIKOR (SSC VIKOR) method has broad application value in the study of automotive consumer purchasing behavior. By analyzing the impact of after-sales service on consumer purchasing decisions, Zhou et al. [16] can better understand consumer needs and expectations, and develop corresponding improvement measures. At the same time, improving the quality of after-sales service is one of the key factors in enhancing the competitiveness of automotive brands, and it is also a focus of consumer attention. Therefore, automobile manufacturers should attach importance to the quality and satisfaction of after-sales service, continuously optimize service processes, improve service levels, in order to attract more consumers to purchase and maintain brand loyalty.

Most of the above studies are limited to a single data source, such as social media and e-commerce websites, and it is difficult to fully understand consumers' behaviors and needs. This article adopts multi-channel and multi-type data sources, including social media and e-commerce websites, so as to obtain more comprehensive and accurate consumer data. By integrating and analyzing data from different sources, we can more accurately understand consumers' needs and behavioral characteristics. In this article, a computer-aided individualized marketing and consumer behavior analysis model based on BD algorithm is constructed, aiming at providing a universal and scalable model and method for enterprises.

3 COMPUTER ASSISTED INDIVIDUALIZED MARKETING AND CONSUMER BEHAVIOR ANALYSIS MODEL

With the continuous development of BD technology, more and more enterprises begin to use BD algorithms to assist individualized marketing and consumer behavior analysis. In order to better integrate these methods, this article constructs a computer-aided individualized marketing and consumer behavior analysis model based on BD. Collect consumer data through multi-channel and multi-type data sources, including social media data and e-commerce website data. The collected data are preprocessed, such as cleaning, de-duplication, labeling, etc., and the original data are converted into an analyzable data format. Use BD algorithm to mine and analyze the preprocessed data, extract consumers' behavior characteristics and potential needs, and understand consumers' interests, preferences, buying habits and other information. According to the results of consumer behavior analysis, formulate individualized marketing strategies, including product recommendation, advertising and promotion activities. Evaluate and optimize the effect of the model to continuously improve the accuracy and reliability of the model.

Deep learning (DL) technology is used to extract features from the preprocessed data. The DL model is trained by using the extracted features and corresponding consumer behavior tags (such as purchase, click, etc.). Based on the optimized model, it can be used to predict and make decisions on new consumer data. For example, we can predict consumers' purchase intention, recommend related products and services, and formulate individualized marketing strategies. According to the actual results of prediction and decision-making, the model is fed back and adjusted. This may include adjustment of model parameters, re-labeling of data sets, etc. Through this method, enterprises can better understand consumer demand and market trends, improve the pertinence and effectiveness of marketing strategies, and then improve market share and profitability. The DL architecture of consumer behavior analysis is shown in Figure 1.

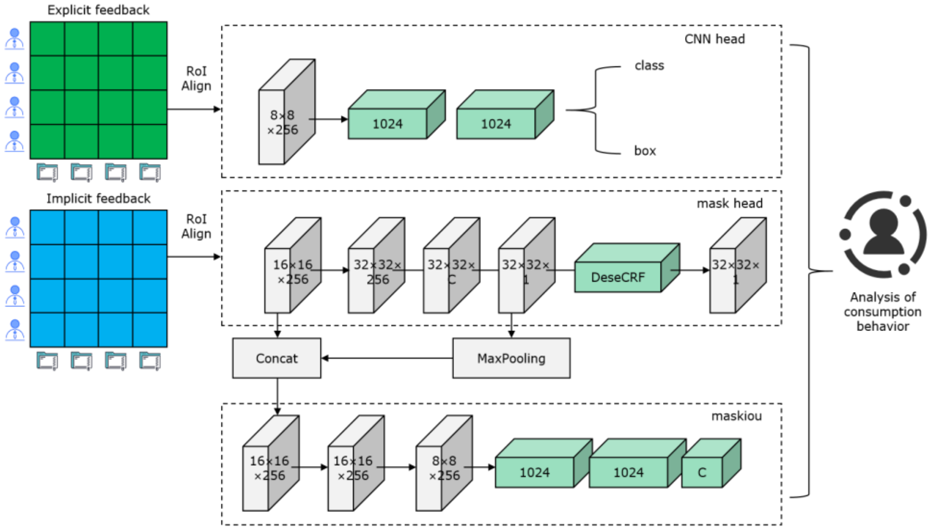

Figure 1: DL architecture of consumer behavior analysis.

Describe the weight w_{ij} word frequency TF of the product attribute j in the product i, and f_{ij} represents the number of times the element d_j appears in the document i. The more times, the greater the importance (competitiveness) of the element. The frequency of element d_j in document i:

$$TF_{ij} = \frac{f_{ij}}{\max f_{ij}} \tag{1}$$

The inverse frequency IDF_j of the element d_j in the document set is:

$$IDF_j = \log_2 \frac{N}{n_j} \tag{2}$$

Calculate the weight of the feature attribute to the whole product, and define the feature degree w_{ij} of the attribute a_j in the product A_i:

$$w_{ij} = TF_{ij} * IDF_j \tag{3}$$

Scoring similarity $PSS(i, j)$ between items i and j can be defined as:

$$PSS(i, j) = \sum_{u \in U_{ij}} PSS_u(R_{ui}, R_{uj}) \tag{4}$$

Where U_{ij} represents the set of users who have evaluated both items i and j. $PSS_u(R_{ui}, R_{uj})$ represents the similarity of user u's rating between items i and j.

If a source sends out a set of information such as $X_1, X_2, \cdots, X_n$, then the self-information content of X_i is the uncertainty before the receiver receives the information X_i or how much information the receiver gets after receiving X_i. Here, $I(X_i)$ is used to represent the information amount of X_i:

$$I(X_i) = \log_2 \frac{1}{P(X_i)} = -\log_2 P(X_i) \tag{5}$$

The Apriori algorithm is improved by introducing related methods, assuming that the data set D is divided into blocks $D_1, D_2, \cdots, D_n$, and the global minimum support number is $minsup\_count$. The local minimum support number of a data block D_i is expressed as $minsup\_count_i \ (i = 1,2,\cdots,n)$:

$$minsup\_count_i = minsup\_count * \frac{\|D_i\|}{D} \tag{6}$$

Select the characteristics related to individualized marketing decision, as well as various possible combinations and transformations. These features will be used to build a clustering model. Select an appropriate clustering algorithm, such as K-means or hierarchical clustering, and build a clustering model according to the results of feature selection. The model will divide consumers into different groups according to their characteristics. The evaluation index is used to evaluate the clustering results to determine the clustering effect. If the effect is not ideal, you can adjust the clustering algorithm or re-select features. According to the actual results of prediction and decision-making, the model is fed back and adjusted. This may include adjustment of model parameters, re-labeling of data sets, etc. At the same time, it is also necessary to continuously pay attention to new market trends and consumer demand in order to continuously optimize models and strategies. The clustering analysis process of individualized marketing decision based on BD is shown in Figure 2.

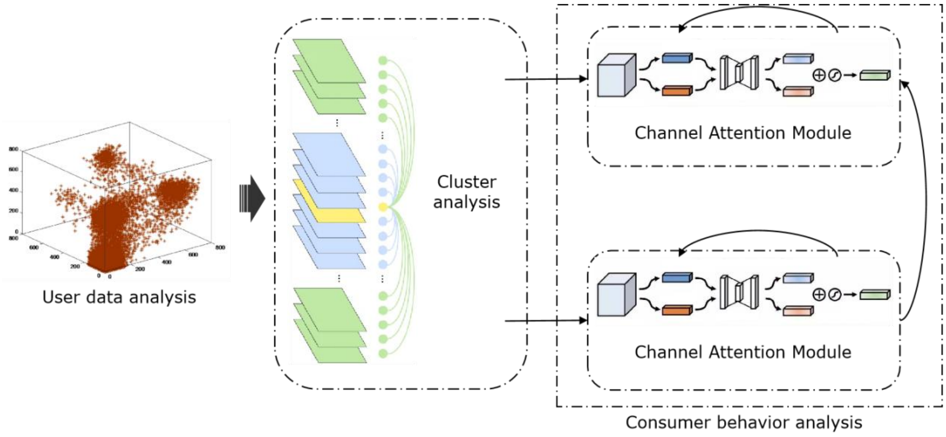

Figure 2: Cluster analysis of marketing decisions.

Analyze the extracted user interest characteristics, and mine the potential needs and behavior patterns of users. According to the actual application effect and feedback, the DL model is optimized and adjusted. This may include adjusting the model parameters, increasing or decreasing the number of layers, changing the activation function, etc. Due to the powerful expressive ability and fitting ability of DL technology, the algorithm can better deal with complex consumer behavior patterns and potential needs, thus more accurately predicting consumer behavior and recommending suitable products or services. The weight of each intermediate node word in the interest spanning tree is:

$$Node(p_j)\cdot w_j = \sum_{i=1}^{k} w_i \tag{7}$$

The freshness of each intermediate node word is:

$$Node(p_j)\cdot x_j = \sum_{i=1}^{k}\left(\frac{w_i x_i}{w_j}\right) \tag{8}$$

Among them, w_j is the weight of intermediate node p_j, x_j is the freshness of entries of intermediate node p_j, k is the number of children of node p_j, w_i is the weight of children's interesting entries p_i, and x_i is the freshness of children's interesting entries p_i.

In the product marketing decision-making, fishbein model is used to predict the attitude of consumers towards this specific object in the whole process:

$$A_0 = \sum_{i=1}^{N} b_i e_i \tag{9}$$

Support degree is the probability that a certain set (X,Y) appears in the total set I:

$$\text{support}\,(X \to Y) = \frac{P(X,Y)}{P(I)} = \frac{P(X \cup Y)}{P(I)} \tag{10}$$

Confidence indicates the probability of deducing Y according to association rules under the condition of X event:

$$confidence(X \to Y) = P(Y \mid X) = \frac{P(X,Y)}{P(X)} = \frac{P(X \cup Y)}{P(X)} \tag{11}$$

The degree of promotion indicates the ratio of the probability of containing Y under the condition of containing X to the probability of containing Y without X:

$$lift(X \to Y) = \frac{P(Y \mid X)}{P(Y)} \tag{12}$$

According to the model prediction results and consumer information, individualized marketing decision-making opinions are generated. Specifically, according to consumers' purchasing history and preferences, we can recommend suitable products or services and formulate corresponding marketing strategies and programs. According to the actual marketing effect and consumer feedback, the generated marketing decision opinions are adjusted and optimized. During the implementation of marketing activities, it is necessary to continuously monitor and evaluate the marketing effect, so as to provide reference and improvement direction for subsequent marketing activities. To generate marketing decision-making opinions through this algorithm, it is necessary to combine consumer data, model prediction results and actual market conditions, formulate targeted marketing strategies and programs, and constantly adjust and optimize them to improve marketing effects and user satisfaction.

4 ALGORITHM TESTING AND ANALYSIS

Extract consumer consumption data from actual marketing data, including consumer's personal information, purchase history, evaluation feedback and other information. After cleaning and preprocessing, the data are constructed into training set and test set. The training set is used to train the model and DT algorithm in this article to learn the consumption behavior pattern of consumers. The test set is used to evaluate the performance of the two models and compare their recall and accuracy. Java+Redis framework and Kafka+Spark Streaming+HDFS framework are quite different in design and implementation. Java+Redis framework may pay more attention to the processing speed of memory data, while Kafka+Spark Streaming+HDFS framework may pay more attention to the streaming processing and distributed storage of large-scale data. In these two frameworks, different data volumes are processed and the data processing speed is tested, as shown in Figure 3.

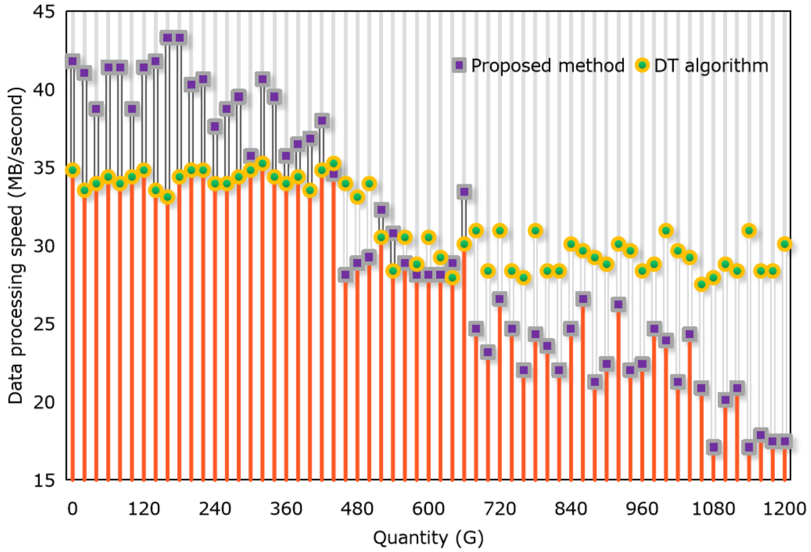

Figure 3: Data processing speed.

When dealing with small-scale data, the data processing speed of Java+Redis framework may be faster. This may be because the memory storage of Redis and the fast data processing ability of Java can be fully exerted at this scale. When dealing with large-scale data, the data processing speed of Kafka+Spark Streaming+HDFS framework may be faster. This may be because Spark Streaming's distributed processing ability and HDFS's distributed storage ability can be brought into full play under large-scale data.

Before interval discretization, data cleaning and preprocessing are very important steps to improve the quality and reliability of data. In the dataset, there may be some fields that are not filled in or data is lost. These missing data may affect the accuracy of subsequent analysis. Methods to deal with missing values can include filling in missing values or deleting data records containing missing values. For some data with wide range of values of features, it is necessary to standardize them. Standardization usually refers to subtracting the mean value and then dividing it by the standard deviation, so that the mean value of data can be 0 and the standard deviation can be 1, so as to make the distribution of data more balanced. When dealing with consumer feature data, it may be necessary to select a large number of features to avoid over-fitting and improve model performance. After data cleaning and preprocessing, interval discretization can be carried out. This process mainly converts continuous consumer characteristic data into discrete interval data according to the results of data distribution analysis. The result of discretization is shown in Figure 4.

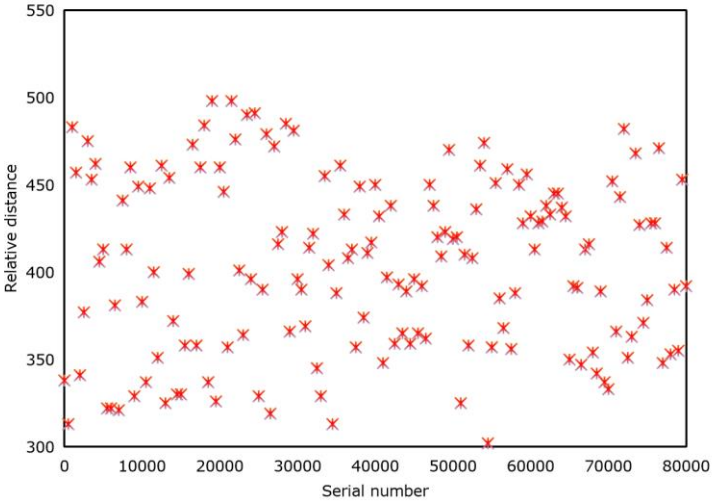

Figure 4: Data De-outlier processing.

According to the results of Figure 5, we can see that the results of individualized marketing decision model learning show a convergence trend. This means that with the increase of training rounds, the model output data gradually approaches the real consumption behavior data. This convergence trend shows that the model is gradually learning and understanding consumer behavior patterns, and can accurately predict consumers' interests and behaviors to some extent.

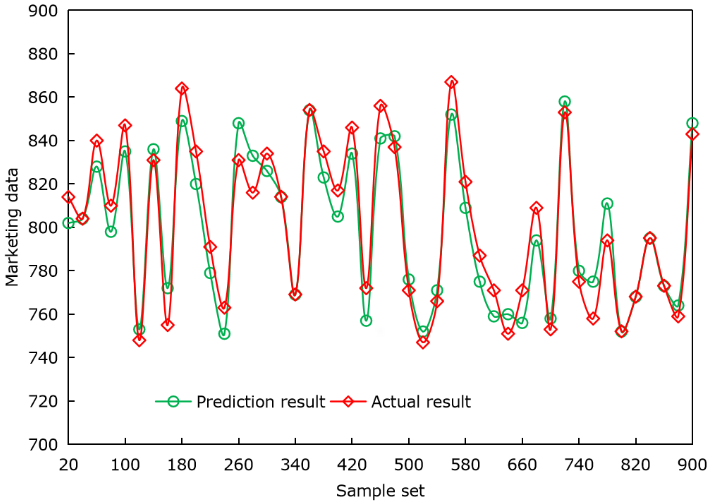

Figure 5: Learning results of individualized marketing decision model.

At the initial stage of training, there is a big difference between the model output data and the real consumption behavior data, because the model has not had enough time to learn the patterns in the data. However, with the increase of training rounds, we can see that the output data of the model gradually approaches the real consumption behavior data. This shows that the model has the ability of learning and adapting, and the model can fit the data well when the training time is long enough. By comparing the model output data with the real consumption behavior data, the accuracy of the model can be evaluated. If the model output data is very close to the real consumption behavior data, then the model can be considered accurate. With the increase of training rounds, the gap between the model output data and the real consumption behavior data gradually decreases, which shows that the accuracy of the model is gradually improved.

From the comparison results of Figure 6 and Figure 7, the recall and accuracy of the algorithm in this article are better than that of DT algorithm in the prediction of consumer propensity in individualized marketing.

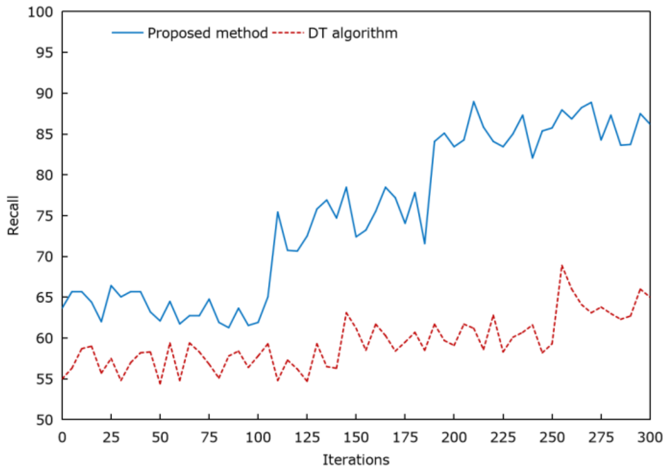

Figure 6: Recall rate of consumption tendency prediction.

The recall rate of DT algorithm is low, and the highest is only about 70%. This means that DT algorithm can only successfully identify a part of all consumers with a specific consumption tendency, but can't comprehensively find out all consumers with this consumption tendency. In contrast, the recall rate of this algorithm is over 86%, which is significantly higher than that of DT algorithm. This means that even among some consumers, the algorithm in this article can accurately identify their consumption tendency, thus covering the consumer groups with this consumption tendency more comprehensively.

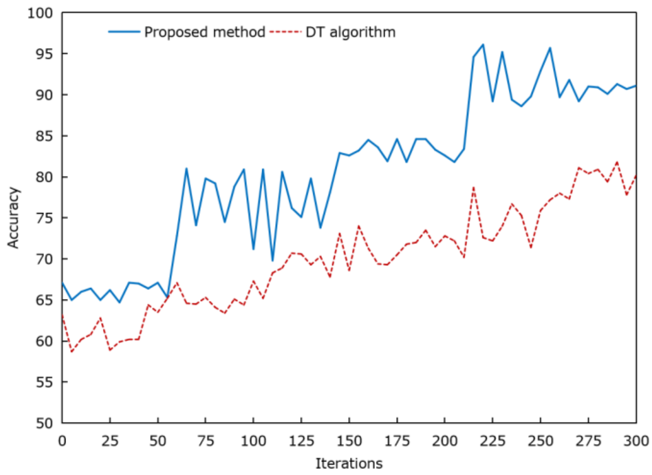

Figure 7: Accuracy of consumption tendency prediction.

The accuracy of DT algorithm is relatively low, and the highest accuracy is only about 80%. In contrast, the accuracy of this algorithm has reached more than 95%, which is greatly improved compared with DT algorithm. This algorithm can accurately judge the real consumption tendency of consumers. To sum up, the recall and accuracy of this algorithm in predicting consumer propensity of individualized marketing are better than DT algorithm. This shows that this algorithm can predict consumers' consumption tendency more comprehensively and accurately, thus playing a greater role in individualized marketing. The reason for this result may be that this algorithm adopts advanced technologies such as DL, which can better process and analyze consumer data, thus improving the accuracy and comprehensiveness of prediction. Through the application of this model, enterprises can better understand consumer demand and market trends, formulate more accurate marketing strategies, improve market competitiveness and obtain greater commercial value.

5 CONCLUSION

In the era of BD, data has become the core competitiveness of enterprises and an important capital of enterprises. This article constructs a computer-aided individualized marketing and consumer behavior analysis model based on BD algorithm, and uses data mining and analysis technology to deeply mine and analyze massive data to extract consumer behavior characteristics and potential needs. The results show that the algorithm in this article can cover consumers with specific consumption tendency more comprehensively, and the prediction accuracy is higher. This shows that the algorithm in this article can better support individualized marketing decision-making and improve marketing effect. DL technology can better process and analyze consumer data and explore potential patterns in the data, thus improving the accuracy and comprehensiveness of forecasting. By comparing the performance of DT algorithm and this algorithm in the prediction of consumer propensity in individualized marketing, it can be concluded that this algorithm is an effective individualized marketing decision support tool. The algorithm can improve the marketing effect, reduce the marketing cost, and provide more accurate and comprehensive prediction results for marketing decision makers. Therefore, in practical application, this algorithm has great potential and can provide strong support for the development of individualized marketing.

In the future research, more advanced DL model can be adopted, which combines various data sources and features to improve the accuracy and generalization ability of prediction. Moreover, by introducing reinforcement learning and other technologies, automatic decision-making and adjustment can be realized, and the efficiency and response speed of marketing decision-making can be improved. By further deepening the development of consumer behavior research, optimizing algorithm model and expanding application scenarios, the accuracy and efficiency of individualized marketing decisions can be improved and more value can be brought to consumers and enterprises.

Jiao Wang, https://orcid.org/0009-0001-6451-1198
Bo Ma, https://orcid.org/0009-0004-2052-2735
Linzhi Nan, https://orcid.org/0009-0009-2263-389X

REFERENCES

[1] Chandra, S.; Verma, S.; Lim, W.-M.; Kumar, S.; Donthu, N.: Personalization in personalized marketing: Trends and ways forward, Psychology & Marketing, 39(8), 2022, 1529-1562. https://doi.org/10.1002/mar.21670

[2] Chiu, M.-C.; Chuang, K.-H.: Applying transfer learning to achieve precision marketing in an omni-channel system–a case study of a sharing kitchen platform, International Journal of Production Research, 59(24), 2021, 7594-7609. https://doi.org/10.1080/00207543.2020.1868595

[3] Darsono, J.-T.; Susana, E.; Prihantono, E.-Y.: Strategic policies for small and medium businesses in marketing through e-commerce, Journal of Entrepreneurship and Sustainability Issues, 7(2), 2019, 1230-1245. https://doi.org/10.9770/jesi.2019.7.2(30)
[4] Kim, J.; Kim, M.: Using personalization for cause-related marketing beyond compassion fade on social media, Journal of Research in Interactive Marketing, 17(2), 2023, 299-316. https://doi.org/10.1108/JRIM-10-2021-0242
[5] Kyaw, K.-S.; Tepsongkroh, P.; Thongkamkaew, C.; Sasha, F.: Business intelligent framework using sentiment analysis for smart digital marketing in the e-commerce era, Asia Social Issues, 16(3), 2023, e252965. https://doi.org/10.48048/asi.2023.252965
[6] Li, C.: E-commerce enterprise marketing management: realistic dilemma, main incentives and resolution strategies, Open Journal of Social Sciences, 08(7), 2020, 245-251. https://doi.org/10.4236/jss.2020.87020
[7] Li, X.; Li, K.-J.: Beating the algorithm: Consumer manipulation, personalized pricing, and big data management, Manufacturing & Service Operations Management, 25(1), 2023, 36-49. https://doi.org/10.1287/msom.2022.1153
[8] Martín, A.; Fernández, I.-A.; Martín, D.-I.; Beltrán, M.: A survey for user behavior analysis based on machine learning techniques: current models and applications, Applied Intelligence, 51(8), 2021, 6029-6055. https://doi.org/10.1007/s10489-020-02160-x
[9] Sharma, S.; Dalip, R.-M.: Comparative analysis of various tools to predict consumer behaviour, Journal of Computational and Theoretical Nanoscience, 16(9), 2019, 3860-3866(7). https://doi.org/10.1166/jctn.2019.8262
[10] Song, Y.-W.-G.; Lim, H.-S.; Oh, J.: "We think you may like this": An investigation of electronic commerce personalization for privacy-conscious consumers, Psychology and Marketing, 38(10), 2021, 1723-1740. https://doi.org/10.1002/mar.21501
[11] Timokhovich, A.-N.; Bulycheva, O.-S.: Technologies for personalization of brand marketing communications using artificial intelligence, Digital Sociology, 3(4), 2021, 19-24. https://doi.org/10.26425/2658-347X-2020-3-4-19-24
[12] Wawrowski, B.; Otola, I.: Social media marketing in creative industries: how to use social media marketing to promote computer games?, Information (Switzerland), 11(5), 2020, 242. https://doi.org/10.3390/info11050242
[13] Wei, C.; Fu, M.; Wang, L.; Yang, H.; Tang, F.; Xiong, Y.: The research development of hedonic price model-based real estate appraisal in the era of big data, Land, 11(3), 2022, 334. https://doi.org/10.3390/land11030334
[14] Wei, W.; Sivaparthipan, C.-B.; Kumar, P.-M.: Online shopping behavior analysis for smart business using big data analytics and blockchain security, International Journal of Modeling, Simulation, and Scientific Computing, 13(04), 2022, 2250053. https://doi.org/10.1142/S1793962322500532
[15] Zhan, M.; Gao, H.; Liu, H.: Identifying market structure to monitor product competition using a consumer-behavior-based intelligence model, Asia Pacific Journal of Marketing and Logistics, 33(1), 2021, 99-123. https://doi.org/10.1108/APJML-08-2019-0497
[16] Zhou, F.; Lim, M.-K.; He, Y.; Pratap, S.: What attracts vehicle consumers' buying: A Saaty scale-based VIKOR (SSC-VIKOR) approach from after-sales textual perspective?, Industrial Management & Data Systems, 120(1), 2020, 57-78. https://doi.org/10.1108/IMDS-01-2019-0034