The Optimization of E-Commerce Digital Marketing Management Strategy Under the Background of Big Data

Haibei Feng*

Weifang Engineering Vocational College Weifang city Shandong Province 262500 China

*1007175847@qq.com

Abstract. This paper discusses the optimization of e-commerce digital marketing management strategy under big data, emphasizing the role of data-driven strategy in improving marketing effect and market competitiveness. The process of consumer data collection, processing and analysis is analyzed, the purchase intention is predicted by logistic regression and random forest model, and the feature importance analysis is carried out. The study evaluated marketing effectiveness before and after the implementation of a data-driven strategy, confirming its effectiveness in increasing sales and market response.

Keywords: Big data; Electronic commerce; Data-driven marketing; Prediction model

1 Introduction

In the digital age, e-commerce competition is intensifying. Companies need innovative digital marketing strategies to excel. This paper discusses optimizing e-commerce marketing in a big data environment. By analyzing consumer data and leveraging machine learning, firms can target markets more effectively and boost campaign efficiency [1]. The study evaluates the impact of data-driven strategies and demonstrates using big data to refine marketing approaches, offering practical advice for e-commerce businesses.

2 Marketing Strategy

2.1 Data-Driven Marketing Decisions

Data-driven marketing strategies form marketing decisions through systematic collection and analysis of consumer data. The process includes data collection, processing, and behavior analysis[2]. The collection stage collects users' website behavior and social interaction data, and the data processing stage cleans and classifies data to ensure quality. Finally, statistics and machine learning are used for in-depth analysis to identify consumer behavior patterns and support marketing strategy development[3].

E. P. H. Lau et al. (eds.), *Proceedings of the 2024 3rd International Conference on Information Economy, Data Modelling and Cloud Computing (ICIDC 2024)*, Advances in Computer Science Research 114,
https://doi.org/10.2991/978-94-6463-504-1_15

2.2 Customer Segmentation

In digital marketing, customer segmentation is achieved through cluster analysis, As shown in Figure 1. grouping consumers according to behaviors and preferences in order to implement targeted marketing strategies[4]. Cluster analysis is unsupervised learning that groups data points by analyzing their patterns. For example, using K-means clustering algorithms, businesses can identify different consumer types, such as value customers, occasional customers, and window customers, and tailor marketing strategies to each group's needs, such as loyalty programs or limited-time coupons, to increase conversion rates [5].

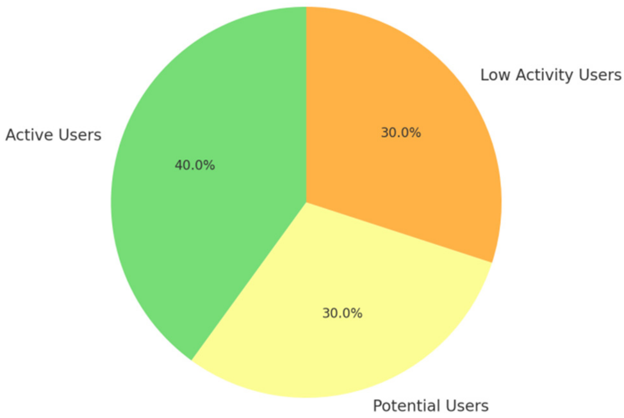

Fig. 1. Customer segmentation results diagram

2.3 Personalized Recommendation

Personalized recommendation systems use machine learning techniques and association rules to analyze consumer data to provide product recommendations tailored to individual user preferences. The purpose of such a system is to enhance the user experience and increase the purchase rate, reduce information overload through precision marketing, and make it easier for consumers to find the goods they are interested in. Association rule mining is a method to discover interesting relationships between data items and is often used in market basket analysis. This approach is based on the rule that if A user buys good A (such as bread), they are also likely to buy good B (such as butter). The association rule formula is:

$$\mathrm{Confidence}(A \to B) = \frac{\mathrm{Support}(A,B)}{\mathrm{Support}(A)} \tag{1}$$

Where, Support(A, B) represents the proportion of transactions that purchase both goods A and B in the total transactions, and Support(A) is the proportion of transactions that purchase goods A in the total transactions.[6].

3 Model Building

3.1 Predictive Models

Logistic regression is a statistical model widely used for binary classification problems. It predicts by estimating the likelihood of an event's occurrence[7]. In the context of building a model to predict purchase intentions, logistic regression takes various features (such as user age, purchase history, and page viewing time) as inputs, and outputs the probability of a user purchasing a specific product. The formula for the model is expressed as follows:

$$P\left(y = 1 \mid x\right) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\cdots+\beta_n x_n)}} \tag{2}$$

Here, x represents the input features, β represents the model parameters, and P(y=1|x) is the predicted probability of a user making a purchase given the features x.

3.2 Feature Engineering

Feature engineering is crucial in e-commerce predictive models, improving model accuracy and gaining insight into marketing strategies by analyzing user website behavior data such as page views and time spent on the site[8]. Temporal characteristics (visit time, weekend or weekday) and social network interactions (likes, comments) reflect user interest in a product, enhancing the potential for purchase prediction, As shown in Figure 2.

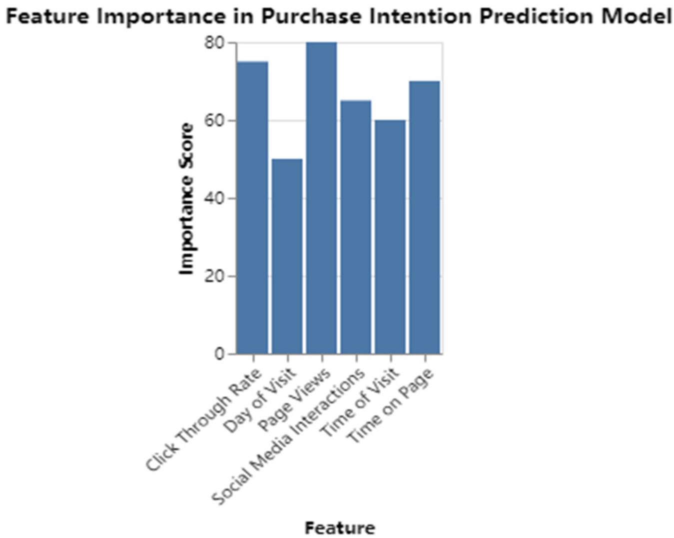

Fig. 2. Feature importance ranking chart

3.3 Model Training and Verification

Model training and validation are key steps to ensure the accuracy and generalization ability of the predictive model. In the field of e-commerce, the establishment of purchase intention prediction model involves data set segmentation, cross-validation and model optimization[9]. The data is divided into training sets, validation sets, and test sets, which are used to build, adjust, and evaluate model performance respectively. Cross-validation improves the stability of the evaluation by repeating training and validation on different data subsets. Model tuning optimizes hyperparameters through methods such as grid search.[10]. The training convergence curve shown in Figure 3 shows that the error rate decreases with training, indicating the optimization performance of the model, and there is no sign of overfitting, ensuring the reliability of the prediction.

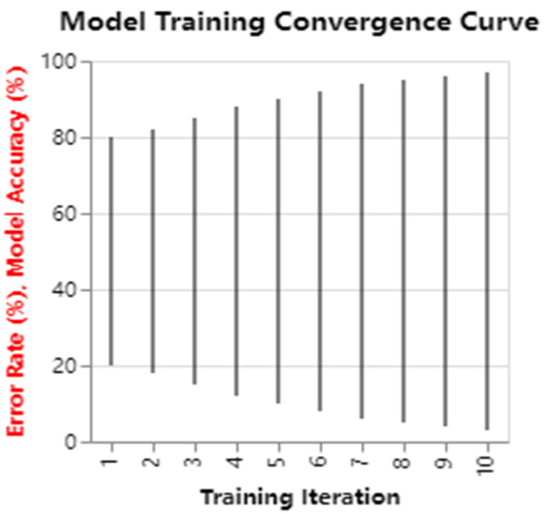

Fig. 3. Convergence curve of model training

4 Effect Analysis

4.1 Model Effect

Recall measures the model's ability to identify positive instances. For purchase intention predictions, the recall rate represents how many of all users that the model correctly identified actually purchased. The ROC curve shows the relationship between the model's true case rate (TPR) and false positive case rate (FPR) at different threshold Settings. The area under ROC curve (AUC) is an important index to measure the overall performance of the model. The closer the AUC value is to 1, the better the performance of the model and the higher the accuracy of distinguishing between positive and negative classes. The area under the ROC curve (AUC) in Figure 4 is 0.83, indicating that the model performs better than a random guess.

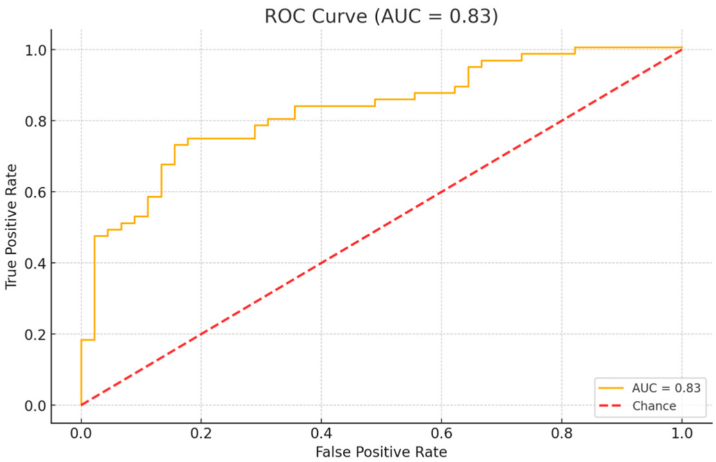

Fig. 4. ROC curve of prediction model

4.2 Marketing Effectiveness

Figure 5 visually shows the change in average sales before and after the implementation of a data-driven marketing strategy. As can be seen in the chart, the average sales after the implementation of the strategy increased from about $120,000 before the implementation to about $140,000. This shows that data-driven marketing strategies are effective in improving sales results. The graph provides a direct economic incentive for companies to implement relevant strategies, while also confirming the important role of data analytics in optimizing sales strategies.

Fig. 5. Marketing campaign effect comparison chart

5 Conclusion

This study deeply analyzes the optimization of e-commerce digital marketing management strategy under the background of big data. The research shows that the application of logistic regression and random forest prediction models can effectively predict the purchase intention of consumers, so as to guide marketing decisions. The implementation of data-driven marketing strategies can significantly improve the accuracy and sales effectiveness of marketing campaigns. By comparing the sales data before and after the implementation of the strategy, the effectiveness and benefits of data-driven strategy in practical application are confirmed.

Reference

1. Li-hui M. Research on countermeasures of digital economy enabling high-quality development of feed industry under Chinese-style modernization[J]. Feed Research, 2024, 47(1).
2. Ruoan R, Han C. Research on digital transformation and upgrading path of garment enterprises under the 14th Five-Year Plan[J]. Wool Textile Journal, 2022, 50(4).
3. De Zuani A. Digital Marketing in China: An analysis of the social media strategy of Tod's Group[J]. 2017.
4. Teng Z. The analyses of the marketing strategies of internet electronic commerce with retail trade as a sample: The thoughts based on Creative Communication Management[C]//2011 International Conference on E-Business and E-Government (ICEE). IEEE, 2011: 1-3.
5. Huang W F. Network marketing promotion sand table simulation experiment teaching[C]//2018 5th International Conference on Education, Management, Arts, Economics and Social Science (ICEMAESS 2018). Atlantis Press, 2018: 1221-1225.
6. Guercini S, Bernal P M, Prentice C. New marketing in fashion e-commerce[J]. Journal of global fashion marketing, 2018, 9(1): 1-8.
7. Sijia N I U, Lei S. Strategic analysis of marketing channels of clothing brand under digital transformation[J]. Wool Textile Journal, 2020, 48(4).
8. Cai ShuFang C S F. Experience in informatization of foreign cooperatives and its inspiration to China[J]. 2014.
9. Fu Y. Live streaming commerce: a review and prospects[C]//2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021). Atlantis Press, 2021: 2546-2552.
10. Li J. The Study of E-Business Model Transformation on China Lottery[J]. 2019.