

Research on predicting the impact of promotional activities on consumer behavior in omnichannel retailing

Shiduo Wang

Joint Institute of the Dalian University of Technology and Belarusian State University

bxwangshiduo@mail.dlut.edu.cn

Abstract. In recent years, the omnichannel sales model has rapidly emerged and is widely used in a context in which consumer buying behavior is influenced by a variety of factors that produce complex changes. Among these factors, the impact of omnichannel promotions on consumer behavior is particularly critical, and thus this research area has received much attention. Predicting the shopping feedback of different types of consumers with different numbers of promotions can help companies develop targeted promotional strategies and improve the consumer shopping experience. In this paper, we classify consumers into different types by clustering them in three dimensions: basic personal information, shopping preference, and shopping channel preference. Based on this, a machine learning prediction model based on the random forest algorithm is built in each of the three dimensions, and the model's performance is evaluated by plotting ROC curves. In order to improve the model's performance, this paper uses the up-and-down sampling method to balance the data. The prediction model based on consumers' "basic personal information", "shopping preference," and "shopping channel preference" has been successfully developed, and the prediction results are excellent. This study provides guidance for companies to develop targeted promotion strategies and offers new ideas and methods for marketing and data analysis. In the future, we can continue to dig deeper into consumer behavior data, classify consumers in more dimensions, build more accurate prediction models, and continue to improve the scientific and accurate promotion decisions of enterprises.

Keywords: omnichannel, promotion, consumer behavior, Random forest, K-means.

1. Introduction

Omni-channel is an integrated channel model derived and developed from single channel and multi-channel. Its emergence is facilitated by people's changing consumption concepts, growing demand, booming technology and the continuous progress of the channel itself. Since 2013, omnichannel has integrated e-commerce model and traditional sales model into one, becoming an important tool for corporate investment and strategic layout.

The concept of omnichannel retailing was first systematically articulated by Darrell Rigby in 2011 and then further refined by Peter C. Verhoef (2015). Omnichannel refers to the sales behavior of a company based on the combination and integration (cross-channel) of as many retail channels as possible [1]. Omnichannel retailing is characterized by "overall, comprehensive, and full line", i.e., the company follows the entire purchase process in real time, not only provides consumers with products, but also interacts with customers instantly with feedback, and is able to develop and integrate multiple types of channels [2]. The development and formation of omnichannel retailing is inextricably linked to technological change on the one hand, and is also influenced by various factors such as changes in shopper behavior, increased demand for experience, comprehensive upgrading of information technology, rapid development of logistics technology, and requirements for management efficiency improvement [3].

According to a study in Asia (2019), about 90% of consumers in Asia prefer to shop online using different channels (e.g., mobile apps, brick-and-mortar, or desktop). Recent market statistics also show that retailers can increase their annual growth rate by 30% and retain 89% of their existing consumers by adopting an omnichannel marketing approach. The impact of omnichannel on consumer behavior is enormous: First, omnichannel marketing can create a better consumer experience through the combination of online and offline brick-and-mortar stores. The seamless

interaction between brick-and-mortar stores and online shopping platforms can meet the different shopping needs of different consumers. Second, omnichannel marketing provides consumers with more and more flexible shopping choices by offering multiple channel options. Third, omnichannel marketing leads to a shift in consumer focus on loyalty relative to a single channel by providing a wider range of consumer choices and more flexible purchasing options. Research shows that consumers' loyalty tends to decline as they become more autonomous and flexible in their choices. This effect has created a need for companies to creatively provide targeted personalized experiences, such as customized choices and experiences, to build long-term, stable customer relationships.

Consumer buying behavior is influenced by many factors. Commodity factors are among the most significant influences on consumer buying behavior. The characteristics of the product itself, such as brand reputation, product quality, and degree of innovation, have a significant impact on consumers' purchase decisions. (Agmeka 2019) [4]. To some extent, the retail environment can influence consumer buying behavior. Retail environment factors have different degrees of influence on consumers' buying behavior; for example, price changes and displays during promotions can influence consumers' buying behavior by affecting elements such as value perceptions, price and quality perceptions, and understanding risks and benefits. Details such as music, decorations, and colors of the retail environment may have a positive or negative impact on consumers' emotions and experiences (Luisa M. Martinez, 2021). [5] In addition to this, consumers' own circumstances largely influence their purchase behavior. For example, the consumer's family composition, income level, etc. (Stella Nordhagen, 2023) [6]. Also, influencing factors such as the consumer's education level and personal status also affect the consumer's buying behavior (Amir Noori, 2023). [7].

Due to the complexity of consumer buying behavior, the influence of promotional behavior on consumer buying behavior has important research significance. Promotion is a strategy to make consumers purchase a product or service within the expected time through a series of marketing activities. The form of promotion includes discounts, gifts, coupons, etc. The frequency of promotion refers to the number of promotional activities and the length of the time period; usually, the more frequent the promotion, the higher the sales. However, excessive and frequent promotions may also affect consumers' brand perceptions and value perceptions of the product, causing them to no longer feel the appeal of the promotion (P.G.F. Nixon, 1989) [8]. Also, the scope of a promotion can have an impact on its effectiveness, and usually promotions with wider coverage result in greater sales growth. According to studies, promotions can have a positive impact on consumer buying behavior, especially during economic recessions. [9] In summary, promotions are an important marketing strategy that can be implemented in a variety of forms and frequencies, and the impact of scope on the effectiveness of promotions needs to be considered. In this study, the main focus was on the prediction of consumer behavior under different types of promotions.

In summary, the rise and successful application of the omnichannel sales model have not only broadened corporate sales channels but also improved consumers' shopping experiences and choice flexibility. At the same time, the factors influencing consumers' purchasing behavior are diverse and complex, among which the impact of promotions on consumers' purchasing behavior is another important research area that has received much attention. By predicting consumer behavior under different numbers of promotions in the context of omnichannel, it will help companies make more targeted promotional strategies to meet consumers' needs. Therefore, this paper constructs a prediction model for the behavior of different types of consumer responses to promotions in the omnichannel context, which is good for bridging the research gap in the current direction and providing a reference for enterprises' strategy formulation.

2. Literature Review

With the intensification of market competition, companies are paying increasing attention to the study of promotional strategies and their effects. After years of research, a large number of literature reviews have been developed that provide a comprehensive summary of the form,

frequency, and scope of promotional strategies, promotional methods, and the mechanisms of the effects of price promotions on consumer behavior. (Eric T. Anderson, 2019) [10] for consumer behavior studies covering factors of consumer decisions and behavior, behavioral motivations and attitudes, emotional states, etc. (Kem Z.K. Zhang, 2016). [11] However, among the many studies, there is little further research on the analysis of consumer types to predict the effectiveness of promotions in an omnichannel context, which is the next step in the research direction.

In a study on promotional strategies, Xi Han (2021) [12] found that online rating-based promotional strategies have a positive impact on consumers' brand awareness as well as purchase intention in a digital marketing environment. Also, Hisashi Kurata (2007) [13] showed that the frequency and depth of promotional strategies can be customized according to product type and market demand. As for the study of consumer behavior, conventional studies have been analyzed in terms of consumer psychological and social factors, consumer attitudes and motivations, emotional states and behaviors, etc. (Augusto Bargoni, 2023) [14]

In the study of consumer behavior under promotional activities, knowing the needs and perceived states of different types of consumers is the key to building consumer analysis models (David Juárez-Varón, 2023) [15]. For example, Monica Kukar-Kinney (2015) [16] focuses on the perceived strengths of consumers in online marketing environments for different discount strengths in different advertising and promotional environments. And Lili Zheng (2023) [17] conducted a categorization study of consumers based on data from social media in order to provide guidance for companies to launch differentiated promotional strategies. However, few academic studies have been conducted to analyze different types of consumers under omnichannel to predict the promotional effects of multiple promotions.

Therefore, predictive modeling of consecutive multiple promotions based on consumer types in an omnichannel context is an important direction of current research. Similar real-time data that can be classified based on consumers' income, demographic characteristics and behavior types can predict the promotion effect and improve enterprises' insight into market trends. This research direction will provide new insights into business operations, and the existing research findings provide a foundation for further inquiry. Since the current academic community lacks predictions of the effects of promotions on different types of consumers under omnichannel, this paper aims to establish prediction models for predicting different consumers' feedback on promotions under different numbers of promotions in the context of omnichannel, starting from consumer shopping channels, consumer characteristics, and the number of promotions, to address the current research gap and provide references for the formulation of enterprise promotion strategies.

3. Model Construction

The purpose of this study is to address the problem of predicting the specific effects of promotions in an omnichannel context for different types of consumers with different numbers of promotions. This prediction problem can be characterized in three dimensions: consumers' basic personal information, consumers' item purchase preference, and consumers' shopping channel preference. Consumers' basic personal information includes their year of birth, education level, marital status, annual household income, number of children in the household, number of teenagers, etc. Consumers' shopping preferences are characterized by investigating the amount of different items such as meat, desserts, and alcoholic beverages that consumers have purchased in the past two years. Consumer purchase channel preferences were characterized by the amount of purchases made by consumers in this company on the website, mobile, and offline. Thus, we separated the independent variables of the original dataset according to these three dimensions to form three sub-datasets, after which we used the k-means algorithm to cluster the consumers from the three dimensions to classify them into different types. After obtaining the clustering results, we add the classification results to the three sub-datasets, taking the consumer responses in the five promotions as the dependent variables and upsampling and downsampling the three to obtain three new datasets

for machine learning. After that, a random forest algorithm was used to construct three prediction models based on "consumer's personal profile", "consumer's item purchase preference", and "consumer's shopping channel preference" by supervised learning based on the three new datasets, and the prediction effectiveness of the models is evaluated by plotting ROC curves.

3.1 K-means algorithm

The K-means algorithm is a commonly used cluster analysis algorithm whose main idea is to divide the dataset into K clusters, each containing a portion of the data in the dataset, with the largest possible spacing between clusters. Essentially, it is an unsupervised learning clustering process that allows for the classification and grouping of data without prior knowledge.

Specifically, the K-means algorithm is based on the following process:

First, k points are randomly generated as centroids, i.e., "centers of mass", which are randomly generated or should be initialized based on a priori knowledge so that the expected number of clusters in the clustering is as close as possible to the target number. Calculate the distance from each data object to each center of mass, and assign each data object to the cluster with the closest center of mass. Recalculate the centroids of the data objects in each cluster and use them as the new center of mass. Repeat the previous step until all data objects are assigned to a cluster or a preset training count or error threshold is reached.

In a practical implementation of the algorithm, the following mathematical expressions can be used to solve each step of K-means:

a. randomly generate k prime centers, which can be represented by the matrix X and the number

$$C = [c_1 \ \cdots \ c_K]$$

of clusters K as follows:

b. For any data object x, the distance to each center of mass is calculated, which can be expressed by the following Euclidean distance formula:

$$d(x_i, c_j) = \sqrt{\sum_{d=1}^D (x_i^d - c_j^d)^2}$$

c. For each data object, find the cluster in which the nearest prime is located:

$$l_i = \arg \min_j d(x_i, c_j)$$

d. For each cluster j, recalculate its center of mass:

$$c_j = \frac{1}{|S_j|} \sum_{x_k \in S_j} x_k$$

e. Repeat steps 2-4 until the center of mass no longer changes or the preset stopping condition is reached.

3.2 Random Forest Model

Random Forest (RF) is an ensemble learning (EL) method that uses a self-service resampling strategy to increase the robustness of the model based on the use of decision trees by the base learner while using a random subset of features for training to improve the generalization ability of the model. RF performs well in solving high-dimensional, nonlinear, and other problems, and therefore it has been widely used in regression problems.

The core idea of random forest is to integrate by building multiple decision trees, using the bagging technique (bootstrap aggregating) to self-resample the samples, randomly selecting one sample set of the same size from the original dataset at a time for training, then splitting using random feature subsets at each node, and finally averaging or voting the prediction results of multiple decision trees.

Specifically, for a dataset D containing N samples and M independent variables, the training process of random forest can be divided into the following steps:

A. Random sampling to obtain the dataset:

For T decision trees, the dataset D is resampled T times by self-service to obtain T different training sets D1, D2, ..., DT.

Using self-service resampling: For the original dataset D containing N samples, n samples are randomly selected with put-back from D to generate the i-th training set Di, i.e., randomized with put-back sampling using self-service sampling method.

$$P(x_i \in D) = \frac{1}{N}$$

B. Feature random selection:

Random forests rank the contributions of each split point based on global features, keep the top k features to form a random feature subset, and find the optimal separation only on that subset. Feature random selection is an important technique in random forests to randomly select a portion of features for splitting at each node to increase the diversity and robustness of the model and avoid overfitting. In random forests, feature random selection is specifically implemented by randomly selecting a subset of features and then selecting the best feature for splitting in that subset.

Specifically, suppose a subset of features containing m features is randomly selected, for each node, the random forest selects a best feature from the subset K for splitting. Here, the random forest uses a Gini Index-based metric to find the best feature for splitting.

$$Gini(p) = \sum_{k=1}^K p_k(1 - p_k)$$

The Gini coefficient is calculated as:

In decision tree node splitting, the smaller the Gini coefficient is, the higher the purity of the node is, so the smaller the Gini coefficient is, the more the features meet the splitting requirements.

C. Decision tree building:

Suppose a training sample set D containing m samples and k independent variables has been obtained, and the decision tree starts from the root node and divides the sample set D into left and right sub-nodes, making the sample points inside each sub-node more and more pure until some stopping condition is satisfied.

$$f(x) = \sum_{m=1}^M a^m I(x \in R_m)$$

For each node n, the best feature and threshold in the subset is selected for splitting, i.e., minimizing the impurity (e.g., variance) of the samples within the node to obtain the left and right child nodes. Repeat steps 2 and 3 up to the leaf nodes.

d. Final prediction

In the random forest classification problem, the voting method used for final prediction is a simple and effective way to integrate. Suppose there are T decision trees in the random forest, and each tree classifies the input samples and outputs the labels of a certain category. For a given test sample, the prediction result of the random forest is obtained by voting on the predictions of all decision trees according to certain rules.

The basic principle of the voting method is to count the results of each decision tree and select the category with the most occurrences or use the probability average for voting, and the final prediction result is the plurality or probability average of the prediction results of all trees

Specifically, suppose there are T decision trees in the random forest, and the classification results of each sample on these T trees are y1, y2, ..., yT. For the classification problem, the final prediction for a given sample x can be calculated using either Hard voting method or Soft voting method:

Hard voting method: the category with the most occurrences among all decision tree voting results is taken as the final classification result. If the ith category in T tree is selected for a number of times, the prediction result of sample x is

$$\arg \max_i \{c_i\}$$

Soft voting method: for each sample, its predicted probability over all decision trees is averaged to obtain the probability of occurrence of all categories, and the category with the highest probability is selected as the result in the final prediction:

$$\hat{y}_x = \arg \max_i \left\{ \frac{1}{T} \sum_{t=1}^T p_{it} \right\}$$

In this study, we use the hard voting method.

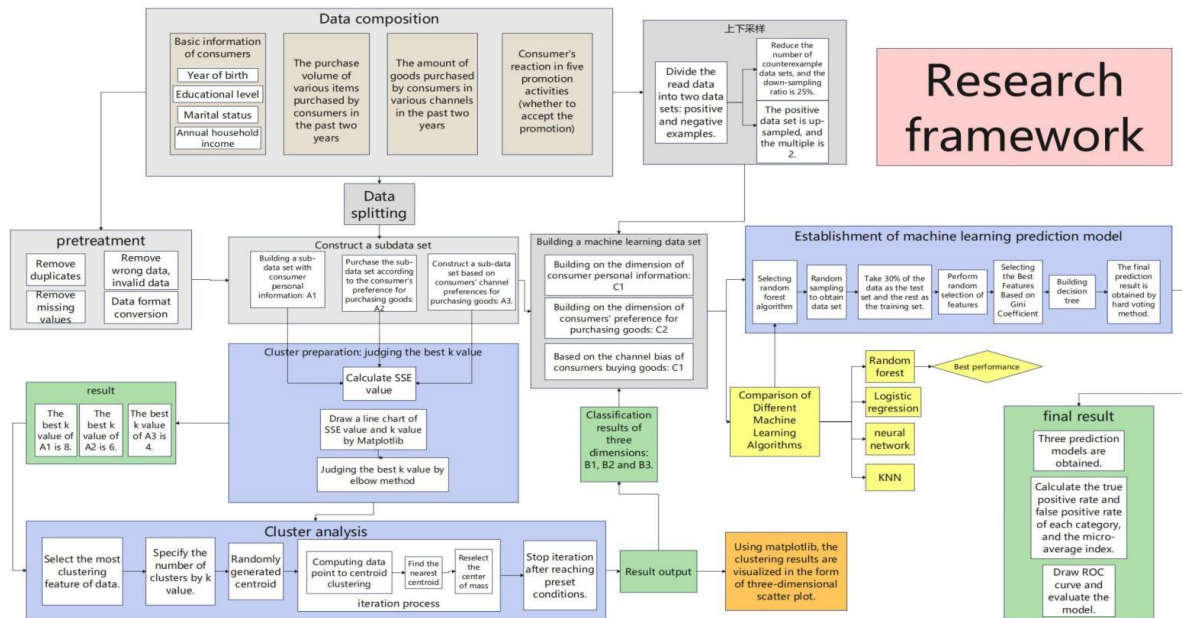


Fig.1. The research framework

4. Research framework

4.1 Data Acquisition and Description

In order to build a prediction model of the impact of promotional campaigns on different types of consumers in an omnichannel context through machine learning algorithms, this paper uses the dataset "Easy Analysis of a Company's Ideal Customers: Customer Personality Analysis in a Marketing Campaign" from the Kaggle website as the original data for training the machine learning model. Analysis in a Marketing Campaign" from the Kaggle website is used as the original data for training the machine learning model. The dataset was provided by "Aman Chauhan." The dataset summarizes the personal information and purchases of 2217 consumers in four dimensions: personal information of consumers, number of different products purchased by consumers, number of purchases made by consumers in different channels, and effectiveness of the campaign. The personal information of consumers includes 10 aspects such as "age", "income level" and "children's status", and the number of different products purchased by consumers includes The number of different products purchased by consumers includes six dimensions, such as "alcohol", "fruit" and "meat", and the shopping channels include "offline stores" and "websites". " The effect of the promotion is indicated by whether consumers accept the promotion in five promotions, with 1 being acceptance and 0 being rejection.

4.2 Experimental environment

For the experiments in this paper, the operating system is Mac OS 12.5, the CPU is an M1 (8-core CPU), the memory size is 16GB, and the graphics card driver is an NVIDIA GeForce MX350.

The development environment is Python 3.7.6, and the development tool is Jupyter Notebook.

4.3 Data processing

(1) Raw processing

First, the data are initially cleaned, including removing duplicate items, missing values, abnormal values, incorrect data, and invalid data, while solving the problems of inconsistency in data formats and data conversion. For classification problems, one-hot coding is used to transform them into numerical information.

(2) Dividing the data set

In order to train the machine learning model in a more targeted way, the original dataset is divided into three dimensions: "personal information of consumers", "number of different products purchased by consumers" and "number of consumers shopping in different channels." In this study, the original dataset was divided into three dimensions: "consumers' personal information," "consumers' purchase of different products," and "consumers' shopping volume in different channels," and the data contained in each dimension was used as the independent variable, and "consumers' acceptance of the five promotions" was used as the dependent variable. The three sub-datasets were constructed: "consumers," "items", and "channels".

(3) Up-and-down sampling method

We analyzed the data on "whether consumers accepted the promotion in the five promotions", and found that the number of "consumers accepted the promotion (data expressed as 1)" was much smaller than the number of "consumers did not accept the promotion (data expressed as 0)". The number of "consumers accepting the promotion (represented as 1)" is much smaller than the number of "consumers not accepting the promotion (represented as 0)". This indicates that the data set is severely unbalanced, so this study used up-and-down sampling to improve the data balance: the read data were divided into two data sets based on positive and negative cases. Next, the number of counterexample datasets was reduced using the down-sampling method with a 25% down-sampling ratio. The positive example dataset is upsampled with a multiplicity of 2. Finally, the upsampled positive example dataset and the downsampled negative example dataset are combined into a new dataset, and the order is randomly disrupted.

4.4 Cluster analysis

In order to better predict the response of different consumers to the promotion behavior under the omni-channel background, this study divides the three dimensions of "consumer's personal situation", "consumer's shopping preference" and "consumer's preference for different shopping channels", and uses K-means algorithm to cluster consumers under each dimension.

4.4.1 Dimension 1: Consumers' personal information:

The clustering analysis was performed with consumers' personal information as the characteristic variable.

(1) Determining the best K-value:

First, the data was read and saved in a Pandas DataFrame object. Next, the data are clustered using the K-Means algorithm, the sum of squared errors (SSE) of the samples within each cluster is calculated, and the SSE values are saved in the list `sse_list`. A line plot of SSE values versus K values is plotted using Matplotlib, and the best K value is selected by the elbow method. Meanwhile, the K-Means algorithm is used to perform clustering operations on the data while calculating the average contour coefficient of each cluster and saving the average contour coefficient value in the list `score_list`. Then, the code plots the line graph of contour coefficients versus K values using Matplotlib, and the best K values are selected by the contour coefficient method. The best K value obtained by combining the above two methods is 8.

(2) Performing cluster analysis:

five columns of data were extracted as features for clustering, and the K-Means algorithm was called to perform cluster analysis. We specify the number of clusters as 8, i.e., the consumers in the dataset are divided into 8 categories. The K-Means algorithm is based on prototype clustering, and

the data are divided into a specified number of clusters by an iterative algorithm, and some clustering centers are found as representatives of each cluster. The algorithm stops iterating when the sample categories within the clusters do not change or the number of iterations has been reached, and the clustering results are imported into "Finally, we use Matplotlib to visualize the clustering results in the form of a 3D scatter plot.

The color of each data point indicates the label of the cluster to which it belongs, and the center of the cluster is indicated by a black asterisk, as shown in the Fig.2:

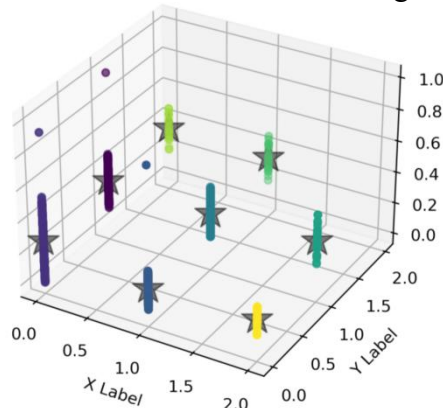


Fig.2

4.4.2 Dimension 2: Consumer shopping category preference:

The number of various products purchased by consumers in the past two years is used as the characteristic variable for cluster analysis. the same method as dimension 1 is used, the best K value of 6 is judged comprehensively, the K-Means algorithm is used, the consumers are clustered, the clustering results are imported into the ""items"" dataset, and the clustering results are visualized in the form of a three-dimensional scatter plot using matplotlib, as shown in Fig 3.

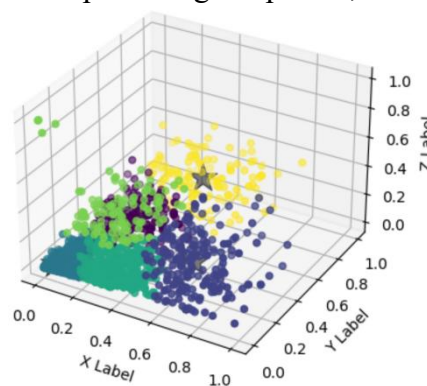


Fig.3

4.4.3 Dimension 3: consumers' preference for different shopping channels:

The number of products purchased by consumers in various channels in the past two years is used as the characteristic variable for clustering analysis. The same method as dimension 1 is used to determine the best K value of 4, and the K-Means algorithm is used to cluster consumers. The clustering results are visualized in the form of 3D scatter plots using Matplotlib. The clustering results are visualized as a 3D scatter plot using Matplotlib, and the clustering results are placed in the sub-dataset "channels". The results are shown in the Fig.4 :

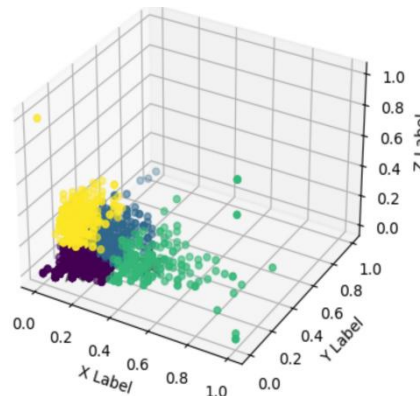


Fig4

4.5 Regression analysis

(1)Data organization:

After the process of cluster analysis, we imported the classification results into the "consumers," "items", and "channels" data. The final data was obtained.

(2)Model selection:

By comparing the ROC images and prediction accuracy of different machine learning algorithms (random forest, KNN, logistic regression analysis, and neural network analysis), the random forest algorithm has the highest comprehensive prediction level, and finally, the random forest algorithm is selected to build the machine learning model.

(3)Analysis process:

First, the data is read using the Pandas library, where the first row is used as the table header and the rest as data values. The data set is divided into a training set and a test set, with 30% of the data taken as the test set and the rest as the training set to test the ability of the model to perform on the new data. A random forest regression model is defined using Random Forest Regressor from the Scikit-Learn library, which contains parameters such as the number of decision trees, n estimators, max depth, and random seeds (random state). In the model training phase, the model is generated by fitting the data through the fit method. The prediction test data are predicted by the predict method for the samples.

5. Analysis of results

5.1 ROC curve plotting:

The random forest model is a very popular machine learning model, which is widely used and can cope with a variety of complex scenarios. When evaluating the random forest model, a common method is to plot the ROC curve to measure the performance of the model by the area size. To assess the performance of the model, prediction accuracy was calculated as an indicator of the model's performance in predicting test data. True positive and false positive rates were calculated for each category and used to plot the ROC curve. An additional "micro-average" metric is also used to aggregate the results across multiple categories. The micro-average true and false positive rates are obtained by summing the results from multiple categories, and the micro-average ROC curves can be plotted, and both curves are shown in the figure to help compare the performance of the prediction results across all categories. Finally, the ROC curves are visualized through the Matplotlib library to make the curves more clearly demonstrate how well our model performs on data classification. The following are the results of the three cluster analysis models: Fig. 5 represents the model based on consumers' personal information; Fig. 6 represents the model based on consumer shopping category preference; and Fig. 7 represents the model based on consumers' preferences for different shopping channels.

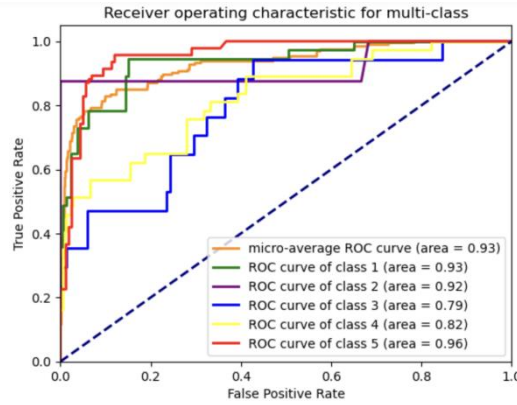


Fig.5. consumers

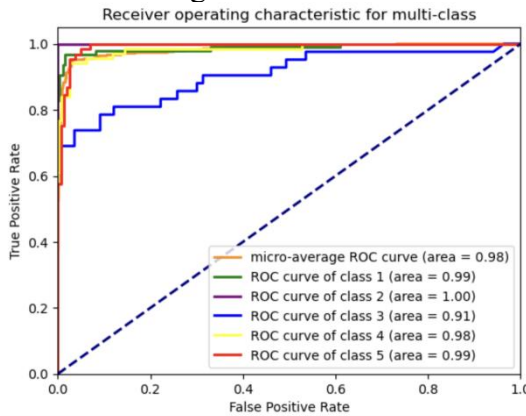


Fig.6. items

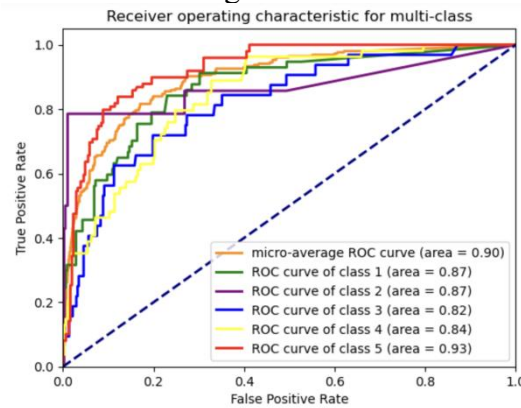


Fig.7. channels

5.2 ROC curve analysis

Three random forest models, "consumers", "items" and "channels", were used to compare and analyze these models. The ROC curve areas of the micro-average curve for the "consumers", "items," and "channels" models are 0.93, 0.98, and 0.90, respectively. In the figure, classes 1 – 5 indicate the effect of promotions 1 – 5, respectively.

First of all, for evaluating the model's performance, appropriate metrics should be selected for evaluation according to the actual application requirements. In most cases, the ROC curve is usually the preferred metric for evaluating model performance in dichotomous problems, and it can be extended to multiclassification problems by calculating multiclassification ROC curves in both micro-average and macro-average ways. For the "consumers" model, we found that the area of the ROC curves for each dependent variable varied greatly, with the first and fifth dependent variables having relatively high areas, while the second, third, and fourth dependent variables had smaller areas, indicating that the model predicted better the first and fifth promotion results.

For the "items" model, we found that its overall performance was the best, with a micro-averaged ROC curve area of 0.98. For the five dependent variables, the second dependent variable had the largest ROC curve area of 1, indicating that this dependent variable had a greater influence on the prediction results of the model. The ROC curve of the second dependent variable is the most accurate.

Finally, for the "channels" model, its overall performance is weak, with a micro-averaged ROC curve area of 0.9. For the five dependent variables, we find that the first and second dependent variables have the largest ROC curve areas, while the third and fourth dependent variables have smaller ROC curve areas. The area is smaller for the third and fourth dependent variables, indicating that the prediction of the effect of the previous and second promotions is more accurate.

6. Conclusion

This paper provides three prediction models for the impact of the number of promotions on different categories of consumers in an omnichannel context: "consumers' personal situation", "consumers' shopping category preferences," and "consumers' preferences for different Three models are provided to predict the impact of the number of promotions on different categories of consumers under the conditions of "consumer's personal situation", "consumer's shopping category preference" and "consumer's preference for different shopping channels". In this study, the K-means algorithm was used to classify the consumer groups in three dimensions: 8 categories from the perspective of "consumer's personal situation", 6 categories from the perspective of "consumer's shopping preference", and 6 categories from the perspective of "consumer's preference for different shopping channels". From the perspective of "consumer preference for different shopping channels", consumers were classified into 4 categories, and this classification can help companies better develop promotion strategies for different categories of consumers.

In addition, based on the random forest algorithm, this study established a machine learning prediction model with the original data of consumers in the above three dimensions and the classification results as the independent variables and the consumer responses in the five promotions as the dependent variables and achieved good prediction results. In this study, the models were evaluated by plotting ROC curves. The average ROC curve area of all three models is greater than 0.9, which indicates that the models are effective. The model built from the perspective of "consumer shopping preference" has an average curve area of 0.98, which is the best prediction effect. The model built from the perspective of "consumer's personal situation" has the best prediction for the first and fifth promotions, and the model built from the perspective of "consumer's shopping preference" has good prediction for all promotions except for the third promotion, which is not so good. The model from the perspective of "consumers' preference for different shopping channels" predicts the best results for the fifth promotion.

In order to improve the prediction accuracy of the model, we analyzed the data characteristics of the dependent variable and found that the number of consumers accepting the promotions was much smaller than the number of consumers rejecting the promotions. In order to improve the data balance and thus the validity of the prediction model, we performed up-and-down sampling of the data, which effectively improved the prediction accuracy.

In summary, this study explores the impact of different numbers of promotions on consumers under omnichannel. A machine learning prediction model based on the random forest algorithm was built by classifying consumer groups in three dimensions, including personal situation, shopping category preference, and preference for different shopping channels. The performance of the model was evaluated by plotting ROC curves, and it was found that the model had good prediction results, with the best prediction results from the perspective of "consumer shopping category preference". In addition, the prediction accuracy of the model was successfully improved by up-and-down sampling to improve the data balance. The findings of this study provide guidance for companies to

effectively develop targeted promotion strategies and offer new ideas and methods for research in the field of promotional marketing and data analysis.

Further studies can explore the effects of promotions under different channels, such as the comparison of online promotions with offline promotions and the response rates of different e-commerce platforms to promotions, and these studies can help companies better select promotional channels and promotional methods to improve the effectiveness of promotions. Meanwhile, the use of other machine learning algorithms such as neural networks and support vector machines, combined with more consumer features and richer information on promotional activities, can also further improve the accuracy and interpretability of predictions. In conclusion, this study provides useful explorations and insights for practical applications and theoretical research. In addition, the models in this study can be considered to be applied to a wider range of industries, such as finance and healthcare, to make predictions and decisions for different problems and objectives. In the process of applying the model, more factors, such as consumer psychology and social network information, as well as more flexible promotional strategy design, can also be considered to make the application scenarios more diverse and personalized. In summary, this study provides an important reference for the development and refined management of promotional marketing strategies and also demonstrates the practical application of machine learning in the marketing field, which has far-reaching significance and value.

Reference

- [1] Peter C. Verhoef , P.K. Kannan , J. Jeffrey Inman. From Multi-Channel Retailing to Omni-Channel Retailing: Introduction to the Special Issue on Multi-Channel Retailing. *Journal of Retailing*.2015,91(2) :174-181
- [2] Catherine Watson , Jeff McCarthy , Jennifer Rowley. Consumer attitudes towards mobile marketing in the smart phone era *International Journal of Information Management* 2013,33(5): 840-849
- [3] Guillaume Do Vale, Isabelle Collin-Lachaud, Xavier Lecocq. Micro-level practices of bricolage during business model innovation process: The case of digital transformation towards omni-channel retailing. *Scandinavian Journal of Management* 2021,37(2): 101154
- [4] Fanni Agmeka, Ruhmaya Nida Wathoni, Adhi Setyo Santoso. The Influence of Discount Framing towards Brand Reputation and Brand Image on Purchase Intention and Actual Behaviour in e-commerce *Procedia Computer Science* 2019,161:851-858
- [5] Luisa M. Martinez , Belén Rando, Luisa Agante , Ana Maria Abreu. True colors: Consumers' packaging choices depend on the color of retail environment. *Journal of Retailing and Consumer Services* 2021,59:102372
- [6] Stella Nordhagen , Kathrin M. Demmler. How do food companies try to reach lower-income consumers, and do they succeed? Insights from a systematic review. *Global Food Security* 2023,37:100699
- [7] Amir Noori Hossein Bonakdari et al. A novel Multiple Attribute Decision-making approach for assessing the effectiveness of advertising to a target audience on drinking water consumers' behavior considering age and education level *Habitat International*. *Habitat International*, 2023,133: 102749
- [8] P.G.F. Nixon, Jenny C. King. OVERPROMOTION OF THE DIETARY THEORY OF CORONARY DISORDERS. *The Lancet*, 1989,334(8675):1334-1335
- [9] Chih-Chin Liang, Yu-Wen Lin. Online promotion effects under time limitation - A study of survey and physiological signals. *Decision Support Systems*, 2023,170: 113963
- [10] Eric T. Anderson a, Edward J. Fox. How price promotions work: A review of practice and theory. *Handbook of the Economics of Marketing*, 2019,1: 497-552
- [11] Kem Z.K. Zhang, Morad Benyoucef. Consumer behavior in social commerce: A literature review. *Decision Support Systems*, 2016,86:95-108
- [12] Xi Han, Jia Tina Du. How online ratings and trust influence health consumers' physician selection intentions: An experimental study. *Telematics and Informatics*, 2021,62: 101631

- [13] Hisashi Kurata , John J. Liu. Optimal promotion planning—depth and frequency—for a two-stage supply chain under Markov switching demand. *European Journal of Operational Research*, 2007,177(2):1026-1043
- [14] Augusto Bargoni, Ilan Alon, Alberto Ferraris. A systematic review of family business and consumer behaviour. *Journal of Business Research*, 2023,158:113698
- [15] David Juárez-Varón, Ana Mengual-Recuerda. Footwear consumer behavior: The influence of stimuli on emotions and decision making. *Journal of Business Research*, 2023,164; 114016
- [16] Monika Kukar-Kinney, Jeffrey R. Carlson. A fresh look at consumers' discounting of discounts in online and bricks-and-mortar shopping *International Journal of Research in Marketing*, 2015,32(4), 442-444
- [17] Lili Zheng. The classification of online consumer reviews: A systematic literature review and integrative framework.*Journal of Business Research*, 2021,135: 226-251