Analysis of Internet Sports Ads Creation Based on Consumption Behavior and Consumption Psychology

Zhuo Bian[1]*

Abstract
The willingness of users to consume sports changes depending on the types of sports advertisements created. The development of Internet sports advertisements interacts with multiple facets of user sports consumption, including role identification, consumption willingness, and consumer satisfaction. This research examines the production of Internet sports advertisements in light of user psychology and consumption behavior. After extracting the characteristics of Internet-based sports consumption, the authors explained the significance of each characteristic dimension of sports consumption. Then, a stacking model was developed to anticipate consumer psychology, and the modeling procedure was described in detail. In addition, the approach for constructing Internet sports advertisements was described, and customer consuming psychology was grouped using the Gaussian mixture model. The effectiveness of the proposed model was demonstrated through experiments. Based on user behavior and the psychology of sports consumption, this study highlights the design ideas for Internet sports advertisements.
Keywords: consumption behavior; consumption psychology; Internet sports ads; ads creation

1. Introduction

Users' willingness to consume sports changes according to the types of sports advertisements is created (Santos & Amado, 2016; Xiao et al., 2009; Yi-fan & Chang, 2012; Yi, Lin, & Fengyun, 2010; Zhiqiang, 2010). In the era of the Internet, Internet sports advertisements, as a popular advertising medium, are widely used to promote sports goods and services (Beusker, Stoy, & Pollalis, 2012; Huang, 2014; Ibrahim, 2014; Lemas et al., 2021; Sanchez et al., 2012; Yang & Ming, 2021). Internet sports ad creation in the context of user consumption behavior and psychology has attracted the attention of advertising and design professionals both domestically and internationally (Bi, 2012; Fu & Miller, 2022; Hora et al., 2022; Khan et al., 2019; Khoironi, Anggoro, & Sudarno, 2019; Lilje & Mosler, 2017; Mäntymäki & Salo, 2015; Singh & Yassine, 2017; Toma et al., 2019). The development of Internet sports advertisements interacts with multiple facets of user sports consumption, including role identification, consumption willingness, and consumer satisfaction. It is anticipated that further investigation and research will reveal the impact of user consumption behavior and psychology on the design of Internet sports advertisements.

The rising diversity of online platform designs and sports equipment usage patterns characterizes the era of big data. Consumer data gathering and consumer behavior analysis encompass many data sources. Song (2021) obtained population data, sports equipment behavior, and consumption intentions through traditional questionnaire surveys and focus group interviews and analyzed the data on sports consumption behavior, revealing that sports consumption awareness influences 28.5% of sports consumption behavior. Sánchez-Zambrano, Zárate, and Torán (2020) planned to measure the brand loyalty of Colombian consumers to several sports companies. The relationship between variables (i.e., satisfaction, brand value, trust, quality, and commitment) was measured after a random sampling of 300 individuals, statistical testing of the null hypothesis via classic correlation analysis, and measurement of the relationship between variables (i.e., brand value, trust, quality, and loyalty) was conducted. The survey and analysis results help sports brands to make decisions that strengthen customer loyalty. Riantini et al. (2019) sought to examine the impact of the electronic marketing mix on the purchasing decisions of online retail consumers of Indonesian sporting goods wholesalers. Descriptive analysis and multiple linear regression were performed on the questionnaire survey data.

Due to the recent expansion of the sports gear sector, sportswear is currently worn by people of all ages. Existing research has gathered and analyzed user experiences, reviews, and comments from e-commerce

[1] Art Academy of Northeast Agriculture University, Harbin 150001, China
*Corresponding author's email: bianzhuo1219@126.com

websites, blogs, and social media platforms. Singhvi and Srivastava (2020) compared the favorable and negative remarks on four criteria of each sports brand using statistical methods. Comparative and evaluative research was conducted on consumer knowledge in the selection of sports brands. Data was extracted from online communities using network graphics and text-mining techniques. Lüthje (2004) evaluated the innovation activities and characteristics of 153 users of outdoor sports goods and noted their high level of innovation. Their research also demonstrates that innovative consumers may be reliably distinguished from non-innovative users based on their anticipated benefits and levels of expertise in utilizing innovative items.

The preceding literature has elaborated on the general consumer buying psychology, providing a theoretical foundation for our work. For instance, predecessors have surveyed and studied the general public's understanding of ad creation within consumer psychology and quantified the consumption behavior and consumption psychology of users of various ages. In addition, the creation of appropriate advertisements has also been quantitatively studied. Based on user consuming behavior and psychology, the user portrait serves as a guide for designing Internet sports ad development. It enables the provision of precise services, targeted marketing, and specific recommendations for users.

There are many papers on general user consumption behavior and consumer psychology in the United States and overseas. Still, there are few systematic studies of Internet consumption behavior and consumer psychological characteristics. Virtually no academic has studied the creation of Internet sports advertisements.

The primary sections are as follows: The second section extracts the characteristics of Internet sports consumption behavior and defines each characteristic dimension of sports consumption. The third section creates a stacking model to forecast consumer psychology and describes the modeling method in detail. Section 4 describes the development method for the Internet sports advertisement creation strategy and groups consumer consuming psychology using the Gaussian mixture model. The effectiveness of the proposed model was demonstrated through experiments. Based on user behavior and the psychology of sports consumption, this study highlights the design ideas for Internet sports advertisements.

2. Feature Extraction

Different user labels can be obtained for the users under other business scenes of Internet sports ads through user portrayal based on consumption psychology. The user consumption dataset mainly covers information about records on user consumption time and price and sports product sellers. Firstly, it is necessary to carry out feature extraction. The qualitative features of the above data cannot be used directly but are converted into quantitative features. Given the vast differences and unknown maximum/minimum of eigenvalues, the original data were normalized by the z-score model. Let v and ρ be the mean and standard deviation of all data samples, respectively. Then, the corresponding conversion function can be established as:

$$u* = \frac{u-v}{\rho} \tag{1}$$

A total of 25 dimensions were selected to analyze the features of user sports consumption. Table 1 explains the meaning of each dimension.

Table 1

Meaning of the characteristic dimensions of user sports consumption

Dimension	Meaning
U	Consumption willingness: 1= low; 2=medium; 3=high
$V1$, $V2$	Consumption frequency; consumption time
$V3$-$V12$	Consumption venues: 10 study areas
	(The number of consumptions of a user in each study area)
$V13$-$V17$	Difference between actual consumption and per-capita consumption
	(Sports food, sports clothing, sports equipment, sports competition services, and sports training services)
$V18$-$V22$	Mean consumption
	(Sports food, sports clothing, sports equipment, sports competition services, and sports training services)
$V23$	Total consumption
$V24$	Consumption evaluation and score

The above user features of sports consumption behaviors lay the foundation for training the prediction model of user sports consumption psychology.

3. Prediction of Internet Sports Consumption Psychology

Ensemble learning, which integrates multiple learners, boasts better generalization performance than a single learner. For the binary classification problem $v \in\{-1, +1\}$ and real function Γ, the error rate of a single classifier l_i is denoted by κ. Then, we have:

$$W\left(l_i(u) \neq \Gamma(u)\right) = \kappa \quad (2)$$

Suppose multiple learners are combined through simple voting. Let H be the total number of base classifiers. If more than H/2 base classifiers predict user consumption psychology correctly, then the ensemble learning model has outputted correct classification results:

$$L(u) = sign(\sum_{i=1}^{H} l_i(u)) \quad (3)$$

Suppose the error rates κ of the different base classifiers are independent of each other. In that case, the classification error rate of the ensemble learning model can be described by Hoeffding's inequality:

$$W\left(L(u) \neq \Gamma(u)\right) = \sum_{S=0}^{[H/2]} \binom{H}{S} (1-\kappa)^S \kappa^{H-S} \leq exp\left(-\frac{1}{2}H(1-2\kappa)^2\right) \quad (4)$$

Formula (4) shows that, with the growing number H of base classifiers, the final error rate of the ensemble learning model declines exponentially and gradually approaches zero.

The user consumption behavior is affected by various factors, which cause differences in consumption psychology. In this paper, the stacking model predicts user consumption psychology, which is highly precise and virtually immune to over-fitting. The details of the model are presented below.

The given training set $F = \{(u_i, v_i), i= 1,2, ..., q\}$ is inputted to the *Level*-0 of the stacking model, where u_i is the eigenvector of the i-th m-dimensional user consumption behavior, and v_i is the class label of user consumption psychology corresponding to the i-th eigenvector. Then, the training set F is decomposed into S nonintersecting subsets of the same scale: $F=F_1, F_2, ..., F_s$ ($s= 1,2, ..., S$). The corresponding test set is $F_{-s}=F-F_s$.

Next, Q different base classifiers $G = G_1, G_2, ..., G_Q$ are established on the *Level*-0 of the stacking model. Each base classifier is trained and tested by F_s and F_{-s}, respectively. Through the training, the test result of F_{-s} is expressed as $W_s=\{(u^s_i, u^s_i), j= 1,2, ..., |F_{-s}|\}$, where $s = 1,2, ..., S$. The test result of the i-th base classifier is the training feature *Ui* of the *Level*-1 of the stacking model. Based on the *Level*-1 training features $U = U_1, U_2, ..., U_Q$ and F_{-s} class labels of the Q base classifiers, it is possible to build a new training set F' for the model.

Then, it comes to *Level* 1 of the stacking model. The base classifiers of that level are selected to train on dataset O' before outputting the test result.

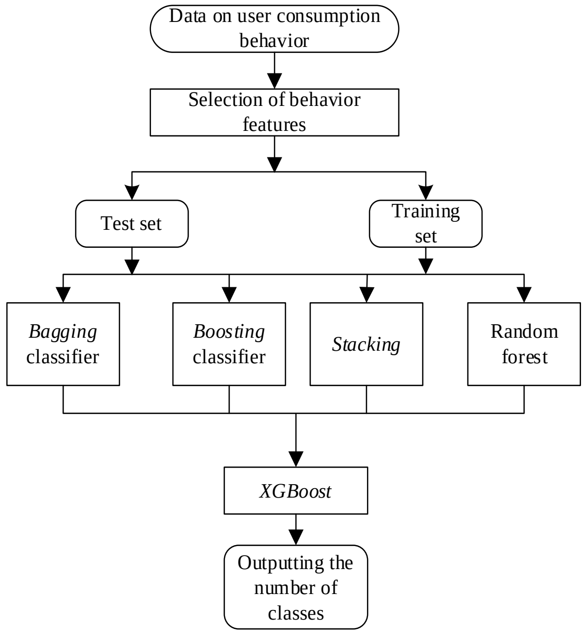

Figure 1. Structure of prediction model for user consumption psychology

Figure 1 illustrates the structure of the prediction model for user consumption psychology. It can be learned that four mainstream classifiers were selected as base classifiers, namely, the bagging classifier, boosting classifier, stacking, and random forest. To optimize the combination of these base classifiers, it is necessary to ensure the diversity and accuracy of the base classifiers in the stacking model. Let R be the generalization error of the prediction by the combination of multiple base classifiers; R* be the weighted mean of the generalization errors of the prophecy by every base classifier; C^* be the weighted divergence, i.e., the weighted mean of the diversity measures of every base classifier. Then, the error-divergence decomposition can be expressed as:

$$R = R^* - C^* \quad (5)$$

This paper measures the difference between the base classifiers by inconsistency. Suppose there are q types of base classifiers. Let G_i and G_j ($i,j=1,2,...,q$, $i\neq j$) be two different types of base classifiers; Q^{11} and Q^{00} be the number of user consumption behavior samples correctly or incorrectly classified by G_i and G_j, respectively; Q^{10} be the number of samples correctly classified by G_i, and incorrectly classified by G_j; Q^{01} be the number of samples correctly classified by G_j, and incorrectly classified by G_i; $Q=Q^{11}+Q^{10}+Q^{01}+Q^{00}$ be the total number of samples. Then, the inconsistency measure $INC_{ij}\in[0,1]$ can be defined as:

$$INC_{ij} = (Q^{10} + Q^{01})/Q \tag{6}$$

Formula (6) shows that the difference between the two types of base classifiers increases with the number of samples classified differently by them. Let o_{ij} be the difference between G_i and G_j. Then, the difference matrix O of q classifiers can be established by:

$$O = \begin{bmatrix} o_{11} & o_{12} & \cdots & o_{1q} \\ o_{21} & o_{22} & & o_{2q} \\ \vdots & & \ddots & \vdots \\ o_{q1} & o_{q2} & \cdots & o_{qq} \end{bmatrix} \tag{7}$$

Then, the difference θ_j of classifier G_j in all classifiers can be expressed as:

$$\theta_i = \sum_{j=1}^{q} o_{ij}/q \tag{8}$$

Then, O can be simplified as:

$$O = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_q \end{bmatrix} \tag{9}$$

Let c_j be the accuracy of classifier G_i. Then, the accuracy matrix C of the q classifiers can be established as:

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_q \end{bmatrix} \tag{10}$$

Let $ES_i=\theta_i+c_i$ be the evaluation score of G_i. After full consideration of the differences and accuracies of the base classifiers, the score matrix of G_i can be established as:

$$K = \begin{bmatrix} ES_1 \\ ES_2 \\ \vdots \\ ES_q \end{bmatrix} \tag{11}$$

Let $ES_{AV}=\Sigma^{q}_{i=1}ES_i/q$ be the mean score of all base classifiers. Thus, the selective ensemble strategy for the base classifiers on *Level*-1 of the stacking model can be determined as follows: First, compute the score of each base classifier. If $ES_i \geq ES_{AV}$, the i-th base classifier will be selected; otherwise, it will be eliminated. In this way, $p(p \leq q)$ base classifiers are chosen for ensemble modeling.

To control the complexity of the stacking model, extreme gradient boosting (XGBoost) was applied on Level 1. XGBoost is improved from the loss function of the gradient boosting decision tree (GBDT) by adding a regular term. The loss function of XGBoost can be expressed as:

$$\Phi(\phi) = \sum_i \Delta(v_i, \hat{v}_1) + \sum_s \Psi(f_s) \tag{12}$$

where,

$$\Psi(f) = \mu H + \frac{1}{2}\upsilon\|\zeta\|^2 \tag{13}$$

To improve model accuracy, the loss function (12) undergoes the Taylor expansion of the second order:

$$\Phi^{(h)} \approx \sum_{i=1}^{q}\left[\Delta\left(v_j\hat{v}^{(h-1)}\right) + t_i g_h(u_i) + \frac{1}{2}t_i {g^2}_h(u_i)\right] + \Psi(g_h) \tag{14}$$

where,

$$t_i = \theta_{\hat{v}^{(h-1)}}\Delta\left(v_i\hat{v}^{(h-1)}\right) \quad l_i = \theta^2{}_{\hat{v}^{(h-1)}}\Delta\left(v_i\hat{v}^{(h-1)}\right) \tag{15}$$

To extract the features of user consumption behavior, the behavior features of the continuous numerical dimension were selected for correlation analysis. Let RU and RV be the mathematical expectations of the variables corresponding to the consumption behaviors of two users, respectively; Var and $Var^{1/2}$ be the variance, and standard deviation, respectively; $COV(U, V)=D((U-RU)\ (V-RV))$ be the covariance of random variables U, and V. The Pearson's correlation coefficient can be defined as:

$$\sigma_{UV} = \frac{COV(U,V)}{\sqrt{Var_U}\sqrt{Var_V}} = \frac{D((U-RU)(V-RV))}{\sqrt{Var_U}\sqrt{Var_V}} \tag{16}$$

Suppose the known user consumption behavior dataset is the superposition of multiple multivariate Gaussian distributions. Then, each data sample is generated by the model superimposed by S l-variate Gaussian distributions S, l, β_1, β_2, ..., β_S, Σ_1, Σ_2, ..., Σ_S, π_1, π_2, ..., π_S. Let S be the number of multivariate Gaussian distributions; l be the number of variates; β_i and Σ_i be the mean and covariance matrix of the i-th l-variate Gaussian distribution, respectively; φ_i be the proportion of the samples generated by the i-th dl-variate Gaussian distribution in all samples. Then, we have:

$$w(u;\delta) = \sum_{s=1}^{S}\frac{\phi_s}{2\phi^{\frac{l}{2}}|\sum s|^{\frac{1}{2}}}e^{\frac{1}{2}(u-\beta_s)^K\sum_s^{-1}(u-\beta_s)} \tag{17}$$

Since the model is superimposed from S l-variate Gaussian distributions, sample u may be generated by any of the S l-variate Gaussian distributions. Then, we have:

$$w(u;\delta) = \sum_{s=1}^{S}\phi_s Q(u;\beta_s,\sum s) \tag{18}$$

Let $Q(u;\beta_s,\Sigma_s)$ be the probability density function (PDF) of the multivariate Gaussian distributions with a mean of β_s and a covariance of Σ_s. Then, the PDF of the Gaussian mixture model can be given by:

$$Q(u;\beta_s,\sum s) = \frac{1}{2\phi^{\frac{l}{2}}|\sum s|^{\frac{1}{2}}}e^{-\frac{1}{2}(u-\beta_s)^H\sum_s^{-1}(u-\beta_s)} \tag{19}$$

The parameters of the above model can be solved by maximum likelihood estimation. The optimal local solution of each parameter can be found with the ultimate expectation algorithm.

To apply the Gaussian mixture model, it is necessary to determine the number of Gaussian distributions. This paper uses the Bayesian information criterion (BIC) to determine the S value. Let s be the number of parameters in the model, q be the number of data samples, and G be the likelihood function of the model. Then, we have:

$$BAY = s\,ln(q) - 2\,ln(G) \tag{20}$$

4. Strategy for Internet Sports Ads Creation Based on User Consumption Portrait

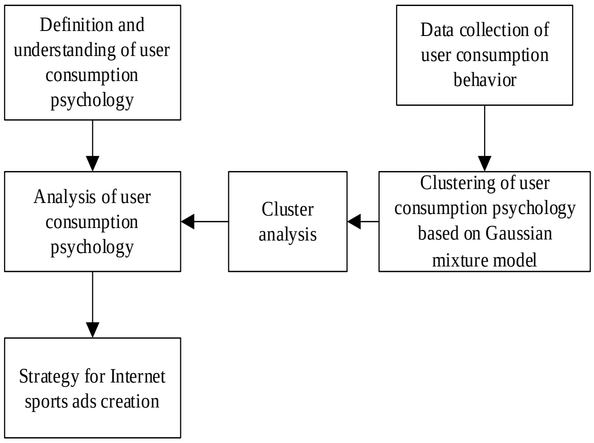

Figure 2. Generation flow of the strategy for Internet sports ads creation

To generate Internet sports ads creation that resonates with user consumption psychology, this paper proposes and establishes a mining and analysis model for user consumption psychology. The model first clusters the users with the same consumption psychology by the Gaussian mixture model and classifies the users based on the clustering results. Next, Internet sports ads are created according to user consumption behavior and psychology laws.

During the cluster analysis, the number of Gaussian distributions in the Gaussian mixture model is determined by the formula (18), i.e., the number of classes for user consumption behavior. Next, k-means clustering (KMC) is adopted to cluster user consumption behavior. For the given dataset on user consumption behavior M = $\{u_1,u_2,...,u_m\}$, every sample has p attributes $b_i = \{b_{i1},b_{i2},...b_{ip}\}$. The Euclidean distance between any two samples b_i and b_j can be calculated by:

$$DIS(b_i, b_j) = \sqrt{\sum_{g=1}^{p}(b_{ig}, u_{jg})^2} \quad (21)$$

The sum DIS_w of the distances between sample pairs can be calculated by:

$$DIS_w = \sum_{i=1}^{q-1}\sum_{j=i+1}^{q} DIS(b_i, b_j) \quad (22)$$

Threshold α can be calculated by:

$$\alpha = \frac{DIS_w}{[q\times(q-1)]/2} \quad (23)$$

There is a neighborhood for any sample. Whether the other samples fall in the neighborhood of sample C can be judged by:

$$DIS(b_i, C) < \alpha \quad (24)$$

If sample b_i satisfies formula (24), then the sample will be reserved in the neighborhood of sample C; otherwise, the sample will be eliminated. Hence, the number of samples q_C within the neighborhood of sample C can be determined. The probability w_C for a sample to fall in the neighborhood of sample C can be calculated by:

$$w_C = \frac{q_C}{q} \quad (25)$$

The samples with a relatively high probability are often selected as cluster centers. If sample b_i is chosen as a cluster center, and if the probability of b_i is similar to that of b_j ($w_{bi}>w_{bj}$), then b_i cannot serve as the center of another cluster. Firstly, it is necessary to compare the distance $DIS(u_i,u_j)$ between b_i and b_j with threshold α. If $DIS(b_i,b_j)>\alpha$, then b_j will be reserved, and the other samples will be judged; otherwise, b_j will be reserved, and the other samples will be considered. The above process is repeated until all cluster centers are selected. Figure 3 explains the flow of cluster analysis. Figure 4 presents the architecture of the strategy generation system.

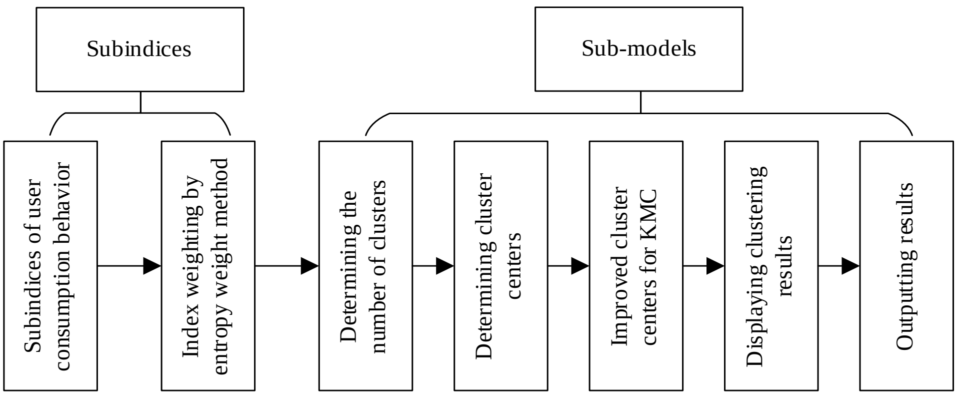

Figure 3. The flow of cluster analysis

The above figures show that the user willingness for consumption is jointly affected by the distance of consumer psychology and the performance of Internet sports ads between users. The elements and forms of ads and user cognition of Internet sports ads are the leading influencing factors.

There are many types of Internet sports. Only creative ones can deeply impress potential users and stimulate their purchase desires. Under its superiority in information dissemination, the Internet can provide a precise information propagation path for users with an obvious consumption demand and enable bidirectional exchanges between users and ads designers and between users and ads service providers.

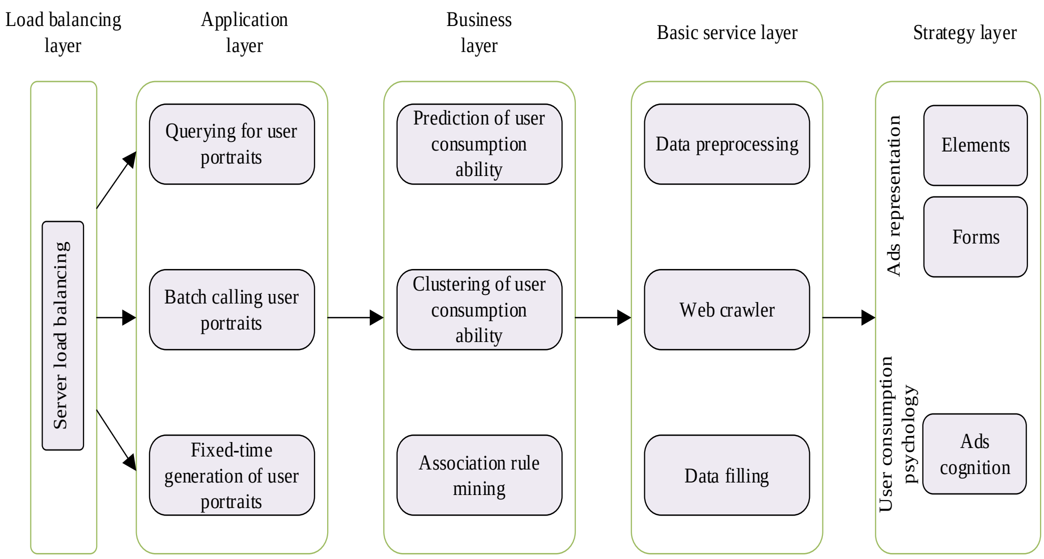

Figure 4. The architecture of the strategy generation system

When it comes to the complicated psychological transformations of sports consumers, there is a clear law governing their behavior. Therefore, it would be practical and effective to build an advertising design strategy from enterprise-controllable production and sales aspects based on the outcomes of an investigation of user consumption behavior and psychology and to utilize this strategy to impact user willingness to consume. Existing Internet sports advertisements lack credibility and technological content. These flaws can be efficiently avoided by establishing an interactive and humanized advertising design and management approach that takes user consumption psychology into account in its entirety.

5. Experiments and Results Analysis

Table 2

Model clustering effects of different combinations

Model	Without XGBoost			With XGBoost		
User willingness for consumption	High	Medium	Low	High	Medium	Low
Silhouette coefficient (SC)	0.42	0.53	0.47	0.54	0.61	0.55
Calinski-Harabasz (CH) index	1447.58	5002.62	1352.85	158.92	441.37	763.84

The research data were obtained from a publicly available dataset on December 10, 2020. The collection includes 664,728 consumption records for 3,566 Internet users and eight attribute indices: invoice number, commodity number, commodity name, consumption date, unit price, consumption quantity, consumer ID, and consumption area. Data collection was followed by feature extraction and preprocessing. Then, the effective and relevant characteristics were chosen and put into our network training model.

This paper measures the effects of model clustering using the silhouette coefficient and CH index. Table 2 depicts our model's mean SCs and CHs following 10 iterations of clustering. The model with XGBoost had greater CHs than the model without XGBoost, indicating that the covariance of the model without XGBoost is insufficiently significant to highlight the inter-class difference of user consumption behavior, or that the covariance is insufficiently small, i.e., the intra-class density of user consumption behavior samples is insufficiently high. In user groups with a strong consumption propensity, the model with XGBoost had higher SCs than the model without XGBoost. This indicates that user groups with a strong inclination to consume are dispersed in location, whereas samples within the same class are primarily centralized. Overall, the addition of XGBoost boosts the model's CH and SC. This is because the XGBoost model features a covariance matrix, which corresponds nicely with the fact that data from different classes are somewhat correlated. Figure 5 depicts the grouping effect more directly and visually.

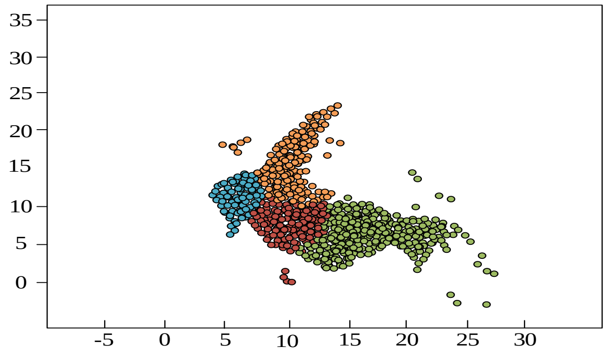

Figure 5. Visualization of the clustering effect

Table 4

Association rules of user sports consumption under different consumption psychologies

Willingness for consumption of user group	High		Medium		Low
Antecedent	Afternoon, sports clothing	Evening, sports clothing	Evening, sports equipment	Afternoon, sports clothing	Noon, sports clothing
Consequent	Sports clothing	Sports equipment	Sports equipment	Sports training	Sports equipment
Support	0.15	0.11	0.16	0.12	0.17
Confidence	0.35	0.22	0.45	0.36	0.47
Lift	1.35	0.48	0.75	1.37	0.58

Table 5

Association rules of user sports consumption under different consumption items

User group number	1	2		3	4		5
Number of users	524	263		89	1152		22
Antecedent	Sports competitio n	Sports equipment	Sports clothing	Sports clothing, sports equipment	Sports competition	Sports clothing	Sports competition
Consequent	Sports clothing	Sports competition	Sports equipment	Sports clothing	Sports training	Sports equipment	Sports training
Support	0.26	0.22	0.14	0.25	0.18	0.16	0.13
Confidence	0.38	0.87	0.55	0.73	0.36	0.45	0.27
Lift	1.14	1.18	1.95	0.84	1.52	1.37	1.85

Tables 4 and 5 illustrate the association rules for sports consumer consumption under various consumption psychologies and consumption items, respectively. Due to the uninformed distribution of experimental data on regional sports consumption, it can be noticed that the majority of the consumption behavior of user groups with varying levels of willingness to consume occurred in the areas of sports apparel, sports competition, and sports equipment.

The experimental findings indicate that the hypothesized association rules for user consuming behavior and psychology correspond to the user consumption times. The majority of Internet sports product purchases occurred in the afternoon and evening. Regardless of clustering results, consumers with a high consumption propensity like a sports competition, sport apparel, and sports equipment. Based on the characteristics of user sports consumption behavior and psychology, it is possible to determine the dimensions favored by sports consumption users and then choose more appropriate and successful ideals and methods for sports design.

Figure 6. Examples of Internet sports ads designed by famous brands

To increase product visibility, most sports firms would pay for endorsements from international athletes (Figure 6). Considering user sports consumption behavior and psychology, Internet sports advertisements must inspire emotional resonance, depict the beauty of strength, and have vibrant colors and unique content. Facing varied consuming domains, Internet advertisements for sports products must combine a high-end aesthetic with product distinction through well-coordinated visuals, words, and colors. The expressiveness of ads should be strengthened in light of the characteristics of user sports-consuming behavior and psychology to satisfy both commercial and spiritual pursuits of sports.

The preceding analysis demonstrates that Internet users outnumber Internet consumers. To build an Internet-appropriate advertising plan, sports product/service businesses, and marketers must be aware of this minor distinction. According to prior research, the specific responsibilities of modern Internet users are crucial to developing Internet sports advertisements by advertisers and entrepreneurs. This article reveals that the effectiveness of Internet sports advertisements is contingent upon eliciting emotional resonance, envisioning the beauty of power, utilizing vibrant colors, and incorporating individualized elements.

6. Conclusions

This research analyzes the production of Internet sports advertisements based on user consumption behavior and psychology. First, the characteristics of Internet sports consumption behavior were retrieved, then each characteristic dimension of sports consumption was explained. The modeling procedure was described in detail after establishing a stacking model to forecast user consumption psychology. In addition, the authors presented the generating process for Internet sports advertisement development strategy and clustered consumer consuming psychology using the Gaussian mixture model. Experiments were conducted to determine the clustering impacts of our model using various combinations, and the clustering results were displayed visually. Under distinct consumption psychologies and consumption goods, the association rules of user sports consumption were investigated. It was discovered that most of the consumption behavior of user groups with varying levels of propensity to consume happened in sports apparel, sporting events, and sports equipment. In light of user sports consumption behavior and consumption psychology, the authors described the design concepts of Internet sports advertisements based on advertisements created by well-known firms.

Our study findings are easily understood and applicable by sports product and service companies. Based on our recommendations, these businesses can watch Internet purchasing activity, design and produce more effective sports advertisements, and push these advertisements to specific user groups. In addition, the abundance of data makes it easier for businesses to divide customers into smaller groups through data mining.

Future research will increase the sample size. This study's research data describe the one-year consumption patterns of an Internet platform's users. Although the data volume satisfies the study objective, it may restrict the scientific rigor of the investigation. Consequently, the following study will collect samples from a broader range to dynamically identify the most valuable Internet users for sports goods and service businesses.

References

Beusker, E., Stoy, C., & Pollalis, S. N. (2012). Estimation model and benchmarks for heating energy consumption of schools and sport facilities in Germany. *Building and Environment, 49*, 324-335. https://doi.org/10.1016/j.buildenv.2011.08.006

Bi, D. (2012). On the Three-dimensional Relationship of Design, Consumption Psychology and Social Development in Green Packaging. In *Advanced Materials Research* (Vol. 591, pp. 115-118). Trans Tech Publications Ltd. https://doi.org/10.4028/www.scientific.net/AMR.591-593.115

Fu, C., & Miller, C. (2022). Using Google Trends as a proxy for occupant behavior to predict building energy consumption. *Applied Energy, 310*, 118343. https://doi.org/10.1016/j.apenergy.2021.118343

Hora, C., Dan, F. C., Bendea, G., & Secui, C. (2022). Residential Short-Term Load Forecasting during Atypical Consumption Behavior. *Energies, 15*(1), 291. https://doi.org/10.3390/en15010291

Huang, P. (2014). Analysis on sport consumption behavior under the environment of E-commerce. In *Applied Mechanics and Materials* (Vol. 687, pp. 2003-2006). Trans Tech Publications Ltd. https://doi.org/10.4028/www.scientific.net/AMM.687-691.2003

Ibrahim, H. (2014). Technology acceptance model: Extension to sport consumption. *Procedia Engineering, 69*, 1534-1540. https://doi.org/10.1016/j.proeng.2014.03.152

Khan, K. N., Scepanovic, S., Niemi, T., Nurminen, J. K., Von Alfthan, S., & Lehto, O.-P. (2019). Analyzing the power consumption behavior of a large scale data center. *SICS Software-Intensive Cyber-Physical Systems, 34*, 61-70. https://doi.org/10.1007/s00450-018-0394-7

Khoironi, A., Anggoro, S., & Sudarno, S. (2019). Community behavior and single-use plastic bottle consumption. *IOP Conference Series: Earth and Environmental Science, 293*(1), 012002. https://doi.org/10.1088/1755-1315/293/1/012002

Lemas, D. J., Loop, M. S., Duong, M., Schleffer, A., Collins, C., Bowden, J. A., Du, X., Patel, K., Ciesielski, A. L., & Ridge, Z. (2021). Estimating drug consumption during a college sporting event from wastewater using liquid chromatography mass spectrometry. *Science of the total environment, 764*, 143963. https://doi.org/10.1016/j.scitotenv.2020.143963

Lilje, J., & Mosler, H.-J. (2017). Socio-psychological determinants for safe drinking water consumption behaviors: a multi-country review. *Journal of Water, Sanitation and Hygiene for Development, 7*(1), 13-24. https://doi.org/10.2166/washdev.2017.080

Lüthje, C. (2004). Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers. *Technovation, 24*(9), 683-695. https://doi.org/10.1016/S0166-4972(02)00150-5

Mäntymäki, M., & Salo, J. (2015). Why do teens spend real money in virtual worlds? A consumption values and developmental psychology perspective on virtual consumption. *International Journal of Information Management, 35*(1), 124-134. https://doi.org/10.1016/j.ijinfomgt.2014.10.004

Riantini, R. E., Andini, S., Florencia, M. M., & Rabiah, A. S. (2019). E-Marketing strategy analysis of consumer purchase decision in Indonesia online sports stores. In *2019 International Conference on Information Management and Technology (ICIMTech)* (Vol. 1, pp. 426-431). IEEE. https://doi.org/10.1109/ICIMTech.2019.8843808

Sánchez-Zambrano, D. F., Zárate, L. F. A., & Torán, M. M. (2020). Influencing Factors in the Consumer's Loyalty Towards Sports Brands. In *2020 6th International Conference on Information Management (ICIM)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICIM49319.2020.244661

Sanchez, F., Alduan, M., Alvarez, F., Menéndez, J. M., & Baez, O. (2012). Recommender system for sport videos based on user audiovisual consumption. *IEEE Transactions on Multimedia, 14*(6), 1546-1557. https://doi.org/10.1109/TMM.2012.2217121

Santos, J., & Amado, P. (2016). The influence of audiovisual sports advertising in sedentary individuals. In *2016 1st International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW)* (pp. 1-8). IEEE. https://doi.org/10.1109/TISHW.2016.7847778

Singh, S., & Yassine, A. (2017). Mining energy consumption behavior patterns for households in smart grid. *IEEE Transactions on Emerging Topics in Computing, 7*(3), 404-419. https://doi.org/10.1109/TETC.2017.2692098

Singhvi, V., & Srivastava, P. (2020). Comparative Assessment of Consumer Perception on Sports Brands Using Data Mined from Online Communities. In A. Saha, N. Kar, & S. Deb (Eds.), *Advances in Computational Intelligence, Security and Internet of Things* (pp. 417-429). Springer Singapore. https://doi.org/10.1007/978-981-15-3666-3_34

Song, Y. (2021). Sports Consumer Behavior Based on Integrated Data in the Context of Big Data. In *The Sixth International Conference on Information Management and Technology* (pp. 1-4). Association for Computing Machinery, New York, NY, United States. https://doi.org/10.1145/3465631.3465642

Toma, A. R., Olteanu, M. V., Gheorghe, C. M., & Dumitrescu, A.-M. (2019). Electric energy consumption behavior of university students. In *2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE)* (pp. 1-6). IEEE. https://doi.org/10.1109/ATEE.2019.8724885

Xiao, Y., Liu, Y., Zhang, L., & Li, N. (2009). Research on the virtual forecast method of background advertisement exposal parameter of sport events. In *2009 5th International Conference on Wireless Communications, Networking and Mobile Computing* (pp. 1-4). IEEE. https://doi.org/10.1109/WICOM.2009.5300930

Yang, C., & Ming, H. (2021). Detection of sports energy consumption based on IoTs and cloud computing. *Sustainable Energy Technologies and Assessments, 46*, 101224. https://doi.org/10.1016/j.seta.2021.101224

Yi-fan, J., & Chang, X. (2012). Brief Study on the Economic Value of Advertising in Great Sports Event. In J. Luo (Ed.), *Soft Computing in Information Communication Technology* (pp. 455-458). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29148-7_64

Yi, X., Lin, Z., & Fengyun, Z. (2010). Research on the Simulation System of Background Advertisement Setting of Sport Events. In *2010 Second International Conference on Information Technology and Computer Science* (pp. 313-316). IEEE. https://doi.org/10.1109/ITCS.2010.83

Zhiqiang, Z. (2010). Research on network advertising of China's sports enterprises. In *2010 International Conference on Computer and Communication Technologies in Agriculture Engineering* (Vol. 1, pp. 162-165). IEEE. https://doi.org/10.1109/CCTAE.2010.5545302