

# **Research** Article

# Accurate Digital Marketing Communication Based on Intelligent Data Analysis

### ZhuoJun Li 🗈

School of Business Administration, Tourism College of Zhejiang, Hangzhou, Zhejiang 311231, China

Correspondence should be addressed to ZhuoJun Li; lzj@tourzj.edu.cn

Received 5 November 2021; Revised 8 December 2021; Accepted 11 December 2021; Published 4 January 2022

Academic Editor: Sikandar Ali

Copyright © 2022 ZhuoJun Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In digital marketing, the core advantages of scientific and technological means such as artificial intelligence and big data analysis gradually appear and pay attention to them. This paper studies the accuracy of digital marketing and proposes an intelligent algorithm based on data analysis, which improves the effect of marketing communication. Through the combination of intelligent algorithms and big data analysis, the data are convincing. Through the comparison and improvement of intelligent algorithm logistic regression and XGBoost, this paper puts forward an improved algorithm of XGBoost based on Bayesian optimization parameters, which can improve the efficiency of digital marketing communication and enhance the social influence of digital marketing.

#### 1. Introduction

In recent years, the purchasing power of the former Millennium baby has gradually appeared in the market. Major businesses are also actively formulating marketing strategies. At present, major Internet platforms are also competing for this market cake and actively applying digital marketing communication strategies to influence the purchasing behavior of this generation and actively communicate with consumers through digital media interactions such as short videos and social media [1]. By adding chat robots and introducing platform intelligent algorithms, each platform actively communicates with target consumers, designs the company's marketing strategy, and collects and analyzes data with consumers before and after purchase to assist the platform to formulate reasonable and effective digital marketing plans and assist decision-making and support [2].

The research on digital marketing has been gradually deepened. Some researchers have incorporated consumer identity into purchase decision support, created consumer demand unconsciously through social platforms, and formed the intervention of social psychological mechanism, which can tap the potential consumption of consumers and provide targeted sales suggestions for manufacturers [3].

In the process of digital marketing, integrating computer and communication technology, innovatively adopting digital tools, carrying out interactive digital marketing research, and providing market decision support for manufacturers through data-driven marketing are a new direction of future development [4].

In the industrial market, some studies put forward the method of evaluating the marketing performance of industrial enterprises through correlation analysis, optimizing the communication efficiency by using digital marketing, and increasing the sales opportunities in the industrial market through statistical analysis of data [5].

In the process of brand marketing communication, through artificial intelligence algorithm, a large number of data are collected based on consumers' personal emotions, and the empathy mechanism in the network environment is to build a communication bridge between consumers and producers and improve brand communication awareness through consumer emotion subdivision. [6] In the development process of digital marketing, artificial intelligence, and big data technology, it is the development trend of digital marketing to effectively evaluate the marketing effect and customize the marketing for you. Compared with the traditional mode, digital marketing embodies high efficiency, measurability, and flexibility [7].

With the prevalence of digitalization, combined with AI algorithm analysis and big data environment, the strategic decision-making of enterprises can be enhanced, and the risks and opportunities of business operations can be discovered early. Through structural equation technology modeling of the questionnaire survey, it is found that big data and AI algorithm have a positive impact on the marketing activities of digital marketing [8].

Network marketing and digital marketing complement each other in modern enterprise marketing, and the current fashion market is gradually migrating from offline to online, establishing corporate brand image online and offline [9]. In the digital marketing environment, this paper combines the common artificial intelligence algorithm logistic regression algorithm and XGBoost algorithm to improve and proposes Bayesian optimization algorithm to improve the parameter optimization process of XGBoost algorithm.

#### 2. Digital Marketing Communication Mode

Marketing model [10] has experienced three historical times, namely, traditional marketing model, network marketing model, and digital marketing model. Fundamentally speaking, the three marketing methods are distinguished based on route, category, and methodology. The changes in each stage are brought about by the changes in consumers' shopping technology and the results of scientific and technological progress.

Digital marketing is a combination of computer, communication, digital multimedia, and other technologies to achieve the purpose of traditional marketing, and its essence is a marketing method derived from traditional marketing methods or replaced by tools [11]. The pursuit of digital marketing is to effectively extract the target customer information from the massive data information and to maximize the use of computer technology to efficiently open up the market and tap the needs of consumers to meet the needs. Digital marketing through marketing matrix [12]: using popular interactive digital multimedia means, such as Douyin short video, Aauto Quicker, Weibo, forum, WeChat circle of friends, select effective targets in the database for targeted marketing to achieve a high conversion rate. Digital marketing acts as a bridge, reduces marketing costs, and embodies customized, targeted, and efficient communication. Through the multichannel cooperation of digital marketing communication, it can effectively interact with consumers, improve the communication efficiency between consumers and manufacturers, and enhance the digital transformation of enterprises. The specific marketing system model is shown in Figure 1.

A product is alive, and this problem should be fully considered in the process of digital marketing. From

entering the market, a product will enter the market from 0 blank (opening period, introduction period) to 1 market recognition (growth period, maturity period), and to 0 exit from the market (recession period). This life cycle mainly includes four stages. Marketing strategies and algorithms are combined with the life cycle of different products, so as to enhance the activity of products in the life cycle and, at the same time, make consumers resonate and have a high sense of identity with products. The following is a comparison of product life cycle and product sales curve, as shown in Figure 2.

#### 3. Intelligent Algorithm

By understanding the needs of consumers, Peter Drucker, a famous management guru, pointed out that the ultimate goal of marketing is to gain business insight and understand users' needs. With the development of artificial intelligence and big data technology [13], enterprises can understand the actual needs of users, improve the service quality of enterprises, and establish the brand image of enterprises through technical means. Through big data technology, we analyze the consumption and purchase behavior of each user, compile the "portrait" of users, distinguish customers, and provide differentiated and valuable targeted services [14]. For example, Xiao Li bought a pair of running shoes online, and his purchase records were recorded in the database. Then, he would push all kinds of socks, sportswear, running-related items, and fitness knowledge for Xiao Li. Through the data recorded by users' daily online consumption, purchase, and browsing, combined with artificial intelligence algorithms and other technologies, the consumption logic of consumers' purchase behavior is excavated, which provides decision-making basis for digital marketing and promotes market transactions. [15] The daily behavior data of users through technical means are analyzed, predicted the future behavior consumption path, gradually enriched the "portraits" of characters, spread the company's cultural values, aroused consumers' recognition, improved marketing efficiency, and achieved marketing effects. Figure 3 shows a digital marketing feedback training model for consumer demand, forming a closed-loop effect of interactive feedback.

Mobile Internet makes digital marketing grow rapidly. The massive data generated by major Internet platforms every day are the real content of public actual participation. Enterprises are closely concerned about making good use of the data of major Internet platforms and carrying out targeted marketing communication. Nowadays, short videos and online dramas are very popular with the younger generation [16], and most of these videos are entertaining. By recording users' usage habits and combining with corresponding artificial intelligence algorithms, the platform makes targeted recommendations to target customers, creating fan product economy, achieving accurate marketing, and realizing traffic realization while harvesting large traffic [17]. In order to achieve accurate marketing, decisionmaking will be more and more based on data and analysis, and enterprises will optimize their operations based on data

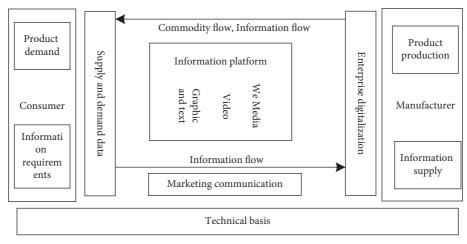


FIGURE 1: Marketing system model.

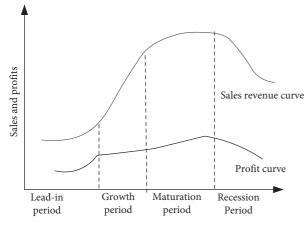


FIGURE 2: Product life cycle and product sales curve.

analysis [18]. The essence of machine learning is to count and find similar data with certain statistical laws, and its goal is to find an optimized model from the regular data space [19]. The framework of digital marketing design by machine learning algorithm is shown in Figure 4. Evaluation criteria and marketing communication performance are important basis for measuring the quality of a model. Combined with specific needs, the evaluation indicators and communication performance reference parameters suitable for this problem are selected. Intelligent algorithm can improve the accuracy of digital marketing, improve personalization and recommendation system, optimize paid advertising, optimize the time and channel of marketing promotion, improve the automation of marketing process, and understand and predict user behavior, etc.

In this paper, several classification algorithms which are relatively frequently used in real life are selected, and the optimal combination method suitable for digital marketing is found by modeling, comparative analysis, and comparison results, and quantitative and qualitative indicators are distinguished [20]. The selection algorithm and improved algorithm are introduced below. 3.1. Multivariate Logistic Regression. This paper is based on the classification of digital marketing user categories. If there are N categories, they can be expressed as  $\{1, 2, ..., k\}$ . Suppose the dataset  $D = \{(x_i, y_i)\}_{i=1}^n = 1, x_i = (1, x_{i1}, ..., x_i p)T \in \mathbb{R}^{p+1}; y_i \in \{1, ..., k\}, i = 1, ..., n.$ 

The binary logistic regression [21] model assumes that the logarithm of the ratio of posterior probabilities belonging to two categories 0 and 1 is linear, as shown by the following formula:

$$\log \frac{\Pr\left(y=1|x\right)}{\Pr\left(y=0|x\right)} = \beta^{T} x. \tag{1}$$

By the logic of the above formula, the logistic regression model can also be called the logarithmic probability model [22]. It is a model with Class 0 as the base class. By analogy, extending to multiple classes, and setting class K as the base class, the logarithmic probability of K-1 is also as follows:

$$\begin{cases} \log \frac{\Pr(y = 1|x)}{\Pr(y = 1|x)(y = K|x)} = \beta_1^T x, \\ \log \frac{\Pr(y = 2|x)}{\Pr(y = K|x)} = \beta_2^T x, \\ \dots \\ \log \frac{\Pr(y = K - 1|x)}{\Pr(y = K|x)} = \beta_{K-1}^T x, \end{cases}$$
(2)

where  $\beta_1, \beta_2, ..., \beta_{K-1} \in \mathbb{R}^{p+1}$ .

$$\begin{cases} \Pr(y = 1|x) = e^{\beta_1^T x} \Pr(y = K|x), \\ \Pr(y = 2|x) = e^{\beta_2^T x} \Pr(y = K|x), \\ \dots \\ \Pr(y = K - 1|x) = e^{\beta_{K-1}^T x} \Pr(y = K|x). \end{cases}$$
(3)

Formulas (4) and (5) are derived from formula (3) as follows:

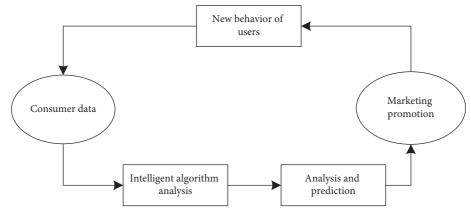


FIGURE 3: Consumer demand feedback training.

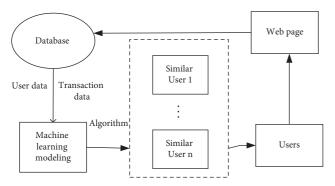


FIGURE 4: Machine learning metrics framework.

$$\sum_{K=1}^{K} \Pr(y = K | x) = 1,$$
(4)

$$pl(x,\beta) = \Pr\left(y = l|x\right) = \frac{\exp\left(\beta_l^T x\right)}{\sum_{j=1}^K \exp\left(\beta_j^T x\right)}.$$
 (5)

From the above derivation, we can know that logarithmic probability can be reflected by the prediction results of the linear regression model in multivariate logistic regression. This model is called "logarithmic probability regression," also known as logistic regression [23], and the logistic regression function will not be affected by wrong assumptions, so data modeling can be carried out directly.

*3.2. XGBoost.* XGBoost algorithm [24], whose full name is the limit gradient lifting algorithm, is an improved learning algorithm of the decision tree. It was proposed by Dr. Chen Tianqi of Washington University in 2014. This algorithm has excellent ability and is widely used in industry, medicine, aviation, economy, and other fields. The algorithm is characterized by parallel computing, giving full play to the advantages of CPU multithreading, and improving the efficiency of data propagation through the improved algorithm.

Given *n* data samples  $D = \{(x_I, y_i)\} (|D| = n, X_i \in \mathbb{R}^m, y_i \in \mathbb{R})$ , with *m* characteristic attributes, the additive model is used to predict the output.

$$\widehat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F},$$
(6)

where the regression tree space is represented by  $F = \{f(x) = w_{q(x)}\} (q: \mathbb{R}^m \longrightarrow T, w \in \mathbb{R}^T)$  which represents all trees, and the structure of each tree is denoted by Q, and the total number of leaves on the tree is denoted by T. Each learner  $f_k$  corresponds to an independent tree structure q and leaf weight w, and it is determined whether the selection algorithm parameters are the best effect [25] by using an objective function defined as follows:

$$\operatorname{Obj}(\theta) = \sum_{i=1}^{n} l(y_i, \, \widehat{y}_i) + \sum_{i=1}^{n} \Omega(f_k).$$
(7)

As can be seen from the above formula, it contains two parts. The left part of the plus sign is the loss function  $\Omega(f) = \gamma T + (1/2)\lambda w^2$ , also known as the differentiable loss function, and it is used to measure the difference between the predicted value  $\hat{y}_i$  and actual value  $y_i$ . This result is also called training error. The right part of the plus sign is represented as a regular term, which punishes the weight of leaf nodes and achieves the purpose of optimizing the objective function through addition training [26] to avoid the influence of overfitting.

Incorporating the predicted value  $\hat{y}_i^{(t)}$  into the following formula minimizes the loss function:

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \ \widehat{y_i}^{(t-1)} +) + f_i(x_i) + \Omega(f_i).$$
(8)

Put the loss term at  $\hat{y_i}^{(t-1)}$  [27], as shown in the following formula:

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} \left[ l(y_i, \hat{y}_i^{(t-1)} +) + g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_i).$$
(9)

Here  $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ ,  $h_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$  are the first and second partial derivatives of the loss function *L*, respectively [28]. The constant term is removed and changed it to the following formula:

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} \left[ g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_i).$$
(10)

Through the function defined above, the regularization term is defined as

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_{j^2}.$$
(11)

In formula (9), the smaller the values of  $\gamma$  and  $\lambda$ , the more complex the model is. Let  $I_j = \{i | q(x_i = j)\}$  be the sample of leaf node *j*, and the formula can be changed as follows:

$$\mathscr{L}^{(t)} \approx \sum_{i=1}^{n} \left[ g_{i} f_{i} \left( x_{i} \right) + \frac{1}{2} h_{i} f_{i}^{2} \left( x_{i} \right) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} \omega_{j^{2}}$$

$$= \sum_{j=1}^{T} \left[ \left( \sum_{i \in I_{j}} g_{i} \right) w_{j} + \frac{1}{2} \left( \sum_{i \in I_{j}} h_{I} + \lambda \right) w_{j}^{2} \right] + \gamma T.$$
(12)

The weight  $w_j^*$  of j in the value corresponding to the objective function is as follows:

$$w_j^* = -\frac{G_j}{H_j + \lambda}.$$
 (13)

The optimal solution can be obtained by bringing the above formula into the objective function as follows:

$$\mathscr{L}^{t}(q) = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_{j}^{2}}{H_{j} + \lambda} + \gamma T.$$
(14)

*3.3. Algorithm Improvement.* Through the idea of optimizing model parameters, we can choose Bayesian optimization algorithm to improve parameter optimization. By setting good values, we train the model for Bayesian optimization and finally form the results for classification evaluation. The optimal parameters are substituted into XGBoost algorithm [29]. Specific steps are as follows:

- (1) Let t = 0, the parameter initial population p (0).
- (2) Select the candidate solution S(T) from P(T) is selected.
- (3) The Bayesian mesh is constructed according to the following formula:

$$x_t = \operatorname{argmax}_{x \in D} \mu_{t-1}(x) + \beta_t^{(1/2)} \sigma_{t-1}(x).$$
(15)

(1.10)

- (4) A new solution O(T) is generated according to the above formula.
- (5) A new population P(T+1) is deduced from O(T).
- (6) Cycle calculation terminates when the calculation is over and outputs the final optimization result. If it is

not finished, jump to (2) and continue to execute in sequence until the optimal result is found at the end of the calculation. By optimizing the XGBoost algorithm model with the above Bayesian optimization parameter method, literature [30] can increase the prediction accuracy of the final result.

#### 4. Data Modeling

The data acquisition model is shown in Figure 5.

Through network data crawling and data modeling, according to descriptive statistics, we screened out key indicators as follows: category browsing times, purchase times, brand browsing times, brand last stay viewing time, and last viewing time of this category, and user's gender, level, age, marriage, education level, occupation, unit price, and purchasing power of each customer in the latest month, including 13 indicators and predicted value  $\gamma$  [31].

4.1. Logistic Regression Modeling. Through data processing and cleaning, the program is executed by *R* language software, and 33 variables of parameter estimated value are obtained, as shown in Table 1. From the table, it can be found that 23 variables with significant changes are retained in the table, for example, category browsing times  $x_1$ , additional purchase times  $x_4$ , purchase times  $x_5$ , etc.; consider removing insignificant variables, such as brand purchase times  $x_2$ , age  $x_{13}$ , unit price per customer  $x_{17}$  in the latest month, etc., which are not significant, so as to test whether the new model fits better. In the new fitting results obtained by experiments, it is found that every regression coefficient in the new model is very significant (P < 0.5) [32].

According to the results of variance analysis in Table 2, the chi-square value is not significant (p = 0.28), which shows that the fitting effect of the new model with 23 independent variables is basically consistent with that of the model with 33 complete variables. Therefore, we can choose a simple model to calculate.

The comparison diagram of the two models is shown in Figure 6. Under different standards of all independent variables, the important values of independent variables corresponding to different independent variables are selected.

The corresponding relationship between algorithm and error evolved from decision tree is shown in Figure 7.

The probability prediction value is obtained by logistic regression model, the probability prediction value of counterexample is less than or equal to 0.5, and the probability prediction value of positive example is greater than 0.5, and the training set error matrix is obtained, as shown in Table 3.

Using this model in the test set, the error matrix of the test set is obtained, as shown in Table 4.

The logistic regression model was analyzed and calculated, and the results are shown in Figure 8.

4.2. XGBoost Modeling. The training set is also modeled and used in the XGBoost algorithm, and the error matrix is shown in Table 5.

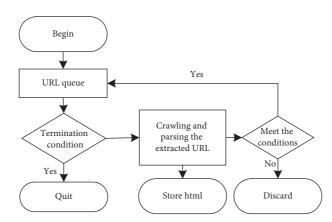


FIGURE 5: Online data acquisition model.

TABLE 1: Estimated	values	of parameters.
--------------------	--------	----------------

	Estimate	Std. error	Z value	$Pr>( \mathbf{z} )$	
(Intercept)	-6.12 E + 00	3.13 <i>E</i> – 01	-19.716	<2 <i>e</i> – 16	* * *
x1	1.11 E - 01	1.83 E - 02	5.804	6.11 E - 09	* * *
x2	1.67 E - 05.	4.41 E - 04	0.038	0.968183	
x3	9.66 E - 03	2.52 E - 03	3.833	0.000128	* * *
x5	8.43 E - 01	2.48 E - 01	3.439	0.000586	* * *
x6	1.43 E - 01	5.78 E - 03	24.506	<2 <i>e</i> – 16	* * *
x7	3.03 E - 02	1.86 E - 03	16.302	<2 <i>e</i> – 16	* * *
x10	4.32 E - 01	2.13 E - 02	20.433	<2 <i>e</i> – 16	* * *
x11	-5.83 E - 01	4.45 E - 02	-13.15	<2 <i>e</i> – 16	* * *
x12	-9.39 E - 02	4.38 E - 02	-2.16	0.031549	*
x13	-9.15 E - 03	1.08 E - 01	-0.841	0.399845	
x14	-7.48 E - 01	3.59 E - 01	-2.088	0.036693	*
x15	-8.15 E - 01	3.59 E - 01	-2.286	0.022199	*
x16	4.83 E - 01	2.56 E - 02	18.645	<2 <i>e</i> – 16	* * *
x17	1.94 E - 01	6.43 E - 02	3.023	0.002508	* *
x18	-3.53 E - 01	5.58 E - 02	-6.30	3.19 E - 10	* * *
x19	5.83 E - 02	5.23 E - 02	1.11	0.267071	
x20	-1.23 E - 01	7.41 E - 02	-1.615	0.106332	
x21	-4.19 E - 02	4.71 E - 02	-0.886	0.374191	
x22	1.63 E - 01	7.19 E - 02	2.281	0.022343	*
x23	2.02 E - 02	1.02 E - 01	0.195	0.843986	
x24	8.34 E - 02	1.01 E - 01	0.812	0.415882	
x25	-6.61 E - 01	1.25 E - 01	-5.215	1.84 E - 07	
x26	-7.31 E - 01	1.31 E - 01	-5.501	3.76 E - 08	
x27	-9.20 E - 01	1.42 E - 01	-6.481	9.02 E - 11	
x28	-1.01 E + 00	2.71 E - 01	-3.776	0.000156	
x29	-1.81 E - 05	3.52 E - 05	-0.52	0.603296	
x30	5.32 E - 01	5.19 E - 01	1.021	0.306568	
x31	5.40 E - 01	3.40 E - 01	1.586	0.112511	
x32	1.08 E + 00	2.64 E - 01	4.148	3.37 E - 05	* * *
x33	1.12 E + 00	2.64 E - 01	4.255	2.08 E - 05	* * *
x34	7.53 E - 01	2.68 E - 01	2.806	0.004987	* *
x35	3.75 E - 03	5.51 E - 04	6.821	9.05 E - 12	* * *
x36	3.21 E - 01	1.55 E - 02	20.535	<2e-16	* * *

TABLE 2: Analysis of variance values.

Model 1	,	2 + x3 + x5 + x6 + x7 + x10 + x 22 + x23 + x24 + x25 + x26 + x			
Model 2	y ~ x1 +	$x^{3} + x^{5} + x^{6} + x^{7} + x^{10} + x^{10}$ $x^{21} + x^{22}$	1 + x12 + x13 + x14 + x14 + x23 + x24 + x25 + x24 + x25 + x24 + x25 + x24 + x25 +		+ x20 +
	Resid. Df	Resid. Dev	Df	Deviance	Pr (>Chi)
1	32358	25684			
2	32373	25702	-15	-17.66	0.2819

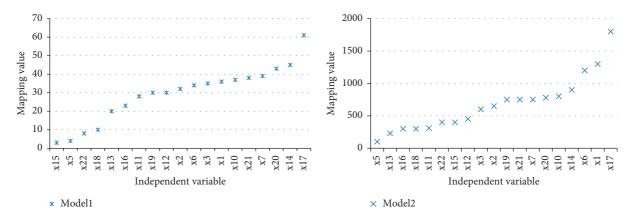


FIGURE 6: Comparison diagram of models 1 and 2.

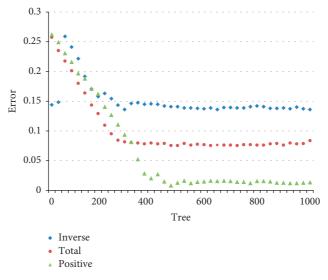


FIGURE 7: Relationship between decision tree and error mapping.

Real situation	Positive example	Negative example
Positive example	13128	3015
Negative example	3135	13138

TABLE 3: Error matrix of logistic regression training set.

INDEE I. EITOT INALIA OF TOGISTIC TEGTESSION LEST SEL		
Real situation	Positive example	Negative example
Positive example	5678	1335
Negative example	1298	5596

TABLE 4. Error matrix of logistic regression test set

Similarly, when the XGBoost model is used for the test set, the resulting error matrix is shown in Table 6.

The model accuracy and recall on the training set and test set are calculated by analyzing the data in Table 6 as shown in Figure 9.

The method proposed in this paper solves the classification accuracy of digital marketing. The XGBoost takes into account the case that the training data are sparse, which can improve the efficiency of the algorithm. The XGBoost algorithm model is simpler and prevents overfitting. The XGBoost uses not only the first derivative but also the second derivative.

4.3. Improved Bayesian Optimization Algorithm XGBoost. Through Bayesian optimization parameters, XGBoost is improved, and some important attribute parameters are selected as shown in Figure 10.

From Figure 10, the importance of algorithm features is improved and compared according to the selection of optimized parameters.

The training set is also modeled and used in the improved XGBoost algorithm, and the error matrix is shown in Table 7.

Similarly, using the XGBoost model for the test set, the resulting error matrix is shown in Table 8.

The data are compared as shown in Figure 11.

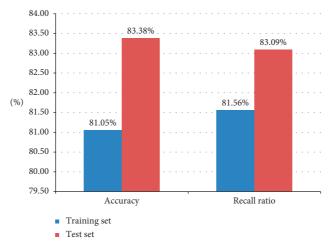
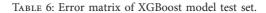


FIGURE 8: Logistic regression model training set and test set.

TABLE 5: Error matrix of XGBoost model training se	set.
--	------

	Positive example	Negative example
Positive example	15239	906
Negative example	1448	14816



Positive example	Positive example positive example	
	5856	1155
Negative example	1392	5494

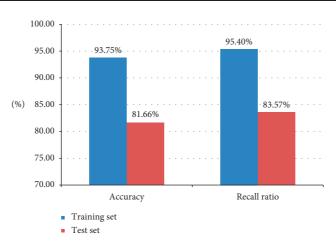


FIGURE 9: XGBoost model training set and test set.

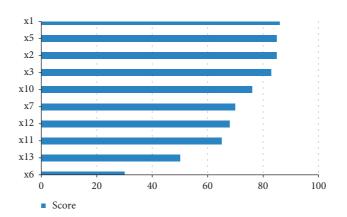


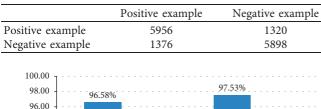
FIGURE 10: Comparison of feature importance of improved algorithm.

#### Scientific Programming

TABLE 7: Error matrix of training set of improved algorithm model.

_	Positive example	Negative example
Positive example	16235	980
Negative example	1513	15818

TABLE 8: Error matrix of improved algorithm model test set.



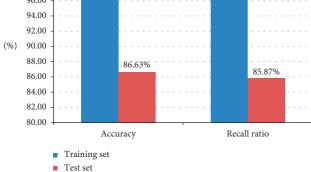


FIGURE 11: Improved algorithm model training set and test set.

#### 5. Conclusion

In this paper, the intelligent data analysis algorithm logistic regression and XGBoost algorithm, and the improved Bayesian optimization parameter XGBoost algorithm are integrated into digital marketing. Combined with the product life cycle, the data that have been crawled through the network are processed, and the data set is modeled, in which 60% of the data are randomly selected as training sets, 40% as test sets, and accuracy and recall are used as performance measurement methods of the model. Through verification, the improved XGBoost algorithm based on Bayesian parameter selection is obviously superior to logistic regression in accuracy and recall, and its computational efficiency is also obviously superior to logistic regression, which shows certain advantages in data processing and analysis.

#### **Data Availability**

The experimental data used to support the findings of this study are available upon request to the author.

## **Conflicts of Interest**

The author declares that there are no conflicts of interest regarding this work.

#### References

 A. Munsch, "Millennial and generation Z digital marketing communication and advertising effectiveness: a qualitative exploration," *Journal of Global Scholars of Marketing Science*, vol. 31, no. 1, pp. 10–29, 2021.

- [2] N. Vilkaite-Vaitone and S. Solo Gu Bas, "Strategic planning of marketing communications in the digital age: an empirical study of small and medium freight transport companies," *Journal of Business Administration Research*, vol. 002, no. 4, pp. 1–7, 2019.
- [3] G. V. Dovzhik, V. N. Dovzhik, and O. V. Kurasova, Empathy and Identification as an Online Technology of Blogger's Communication in Digital Marketing, Springer International Publishing, Heidelberg, Germany, 2021.
- [4] A. S. Krishen, Y. K. Dwivedi, N. Bindu, and K. S. Kumar, "A broad overview of interactive digital marketing: a bibliometric network analysis," *Journal of Business Research*, vol. 131, no. 2, pp. 183–195, 2021.
- [5] N. Izakova, L. Kapustina, and A. Drevalev, "Performance of social media marketing communications of industrial companies," SHS Web of Conferences, vol. 93, no. 2, Article ID 03027, 2021.
- [6] A. N. Timokhovich and O. S. Bulycheva, "Technologies for personalization of brand marketing communications using artificial intelligence," *Digital Sociology*, vol. 3, no. 4, pp. 19–24, 2021.
- [7] L. Amorim and B. Sousa, "Digital marketing and tourism trends: an exploratory study in the context of barcelos (Portugal)," *African Journal of Hospitality Tourism and Leisure*, vol. 9, no. 6, pp. 1–10, 2021.
- [8] T. Saheb and B. Amini, The Impact of Artificial Intelligence Analytics in Enhancing Digital Marketing: The Role of Open Big Data and Ai Analytics Competencies, Research Square, Durham, NC, USA, 2021.
- [9] V. Bilyk, O. Serhiienko, and I. Krupenna, Digital Marketing Tools in The Conditions of Transformation of Communications of The Modern Organization, vol. 825, Ecovis, Berlin, Germany, 2020.
- [10] M. A. Nindyta, S. Hesti, R. Briandana, and P. M. Kurniasari, "Digital marketing communication strategies: the case of Indonesian new's portals," *International Journal of Economics and Business Administration*, vol. VIII, no. 3, pp. 307–316, 2020.
- [11] D. Munandar, "The model of digital marketing communication strategy, customer relationship management and product quality to encourage purchase decision," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 4, pp. 486–492, 2020.
- [12] M. Rozaq, S. Hastjarjo, and Y. Slamet, "How novice entrepreneurs learn about digital marketing communication strategies?" *Jurnal ASPIKOM*, vol. 6, no. 2, p. 345, 2021.
- [13] H. Wang, N. A. Dembsey, B. J. Meacham, S. Liu, and A. Simeoni, "Comparison of sensitivity matrix method, power functionbased response surface method, and artificial neural network in the analysis of building fire egress performance," *Journal of Building Engineering*, vol. 43, no. 2, Article ID 102860, 2021.
- [14] Z. Sun and J. Shao, "Preliminary study on innovative design of bamboo furniture based on users' big data," *Journal of Physics: Conference Series*, vol. 1774, no. 1, Article ID 012017, 2021.
- [15] Y. Li, X. Chen, T. Mao, and G. Huang, "User portrait for archival talents based on recruitment," in *Proceedings of the* 2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), April 2021.
- [16] X. Chi, Z.-P. Fan, and X. Wang, "Pricing mode selection for the online short video platform," *Soft Computing*, vol. 25, no. 7, pp. 5105–5120, 2021.
- [17] X. Wang, "Accurate marketing strategies based on data analytics," *Journal of Physics: Conference Series*, vol. 1744, no. 4, Article ID 042201, 2021.

- [18] L. Kang, S. Liu, D. Gong, and M. Tang, "A personalized pointof-interest recommendation system for O2O commerce," *Electronic Markets*, vol. 31, pp. 253–267, 2020.
- [19] H. Wang, "Multi-sensor fusion module for perceptual target recognition for intelligent machine learning visual feature extraction," *IEEE Sensors Journal*, vol. 21, no. 99, p. 1, 2021.
- [20] S. Qi and J. Feng, "Performance evaluation of micro-blog marketing based on fuzzy theory," *Journal of Intelligent and Fuzzy Systems*, no. 6, pp. 1–11, 2021.
- [21] P. Bradulov, V. Ordynskyi, and L. Romanenko, "Evaluation of digital marketing communications – quantitative and qualitative indicators," *Herald UNU International Economic Relations and World Economy*, vol. 33, no. 5, 2020.
- [22] Z. A. A. Al\_Bairmani and A. A. Ismael, "Using logistic regression model to study the most important factors which affects diabetes for the elderly in the city of h," *Journal of Physics: Conference Series*, vol. 1818, no. 1, Article ID 012016, 2021.
- [23] C. Song, L. Wang, and Z. Xu, "An optimized logistic regression model based on the maximum entropy estimation under the hesitant fuzzy environment," *International Journal* of *Information Technology and Decision Making*, pp. 1–25, 2021.
- [24] L. Torlay, M. Perrone-Bertolotti, E. Thomas, and M. Baciu, "Machine learning–XGBoost analysis of language networks to classify patients with epilepsy," *Brain Informatics*, vol. 4, no. 4, pp. 159–169, 2017.
- [25] J. L. Rodriguez-Alvarez, R. Lopez-Herrera, I. E. Villalon-Turrubiates, G. Grijalva-avila, and J. L. García Alcaraz, "Modeling and parameter optimization of the papermaking processes by using regression tree model and full factorial design," *Tappi Journal*, vol. 20, no. 2, pp. 123–137, 2021.
- [26] H. Toraya, "Finding the best-fit background function for whole owder attern fitting using LASSO combined with tree search," *Journal of Applied Crystallography*, vol. 54, no. 2, pp. 1–11, 2021.
- [27] W.-D. Xi, L. Huang, C.-D. Wang, Y.-Y. Zheng, and J.-H. Lai, "Deep rating and review neural network for item recommendation," *IEEE Transactions on Neural Networks and Learning Systems*, no. 99, pp. 1–11, 2021.
- [28] Y. U. Haiyan, Z. Zhou, and Z. Zhao, "A method for rapid matching based on second order partial derivative," *Wireless Sensor Network*, vol. 2, pp. 37–42, 2010.
- [29] J. Jeong and H. Shin, "Bayesian optimization for a multiplecomponent system with target values," *Computers & Industrial Engineering*, vol. 157, Article ID 107310, 2021.
- [30] J. H. Elsas, N. A. G. Casaprima, P. H. S. Cardoso, and I. F. M. Menezes, "Bayesian optimization of riser configurations," *Ocean Engineering*, vol. 236, no. 5, Article ID 109402, 2021.
- [31] J. Wang, Q. Li, J. Zhou, Y. Wen, Y. Zhang, and P. Lin, "News crawling based on Python crawler," *Journal of Physics: Conference Series*, vol. 1757, no. 1, Article ID 012117, 2021.
- [32] N. Ekong, I. Moffat, A. Usoro, and I. Matthew, "A comparative study of the impact of dummy variables on regression coefficients and canonical correlation indices," *An Empirical Perspective*, vol. 19, no. 4, pp. 576–586, 2021.