

Retraction

Retracted: Research on the Effect of Online Marketing Based on Multimodel Fusion and Artificial Intelligence in the Context of Big Data

Security and Communication Networks

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. 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.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] H. Zhao, F. Lyu, and Y. Luo, "Research on the Effect of Online Marketing Based on Multimodel Fusion and Artificial Intelligence in the Context of Big Data," *Security and Communication Networks*, vol. 2022, Article ID 1516543, 9 pages, 2022.

Research Article

Research on the Effect of Online Marketing Based on Multimodel Fusion and Artificial Intelligence in the Context of Big Data

Hongyu Zhao,¹ Fang Lyu ,² and Yalan Luo¹

¹Business School, Jeonju University, Jeonju 55069, Republic of Korea

²School of Business Administration, Shanxi University of Finance and Economics, Taiyuan 030006, Shanxi, China

Correspondence should be addressed to Fang Lyu; 631505040101@mails.cqjtu.edu.cn

Received 9 November 2021; Revised 8 December 2021; Accepted 16 December 2021; Published 5 January 2022

Academic Editor: Jian Su

Copyright © 2022 Hongyu Zhao et al. 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.

Traditional online marketing methods use a single model to predict the advertising conversion rate, but the prediction results are not accurate, and users are not satisfied with the recommendation results. Therefore, this paper proposes an online marketing method based on multimodel fusion and artificial intelligence algorithms under the background of big data. First, it introduces big data technology and analyzes the characteristics of network advertising marketing model (RTB). Second, combined with multitask learning and fusion technology to improve the single model in advertising conversion rate prediction effect, prediction results to further improve the accuracy of results. Then, tF-IDF technology in artificial intelligence algorithm is used to measure the importance of advertising words in online marketing and calculate the contribution degree. Finally, according to XGBoost technology, the multitask fusion model of online marketing effect is classified. Experiments are used to analyze the effect of online marketing. Experimental results show that the proposed method can improve the accuracy of advertising conversion rate prediction and online sales of goods.

1. Introduction

With the rapid development of high and new technology, science and technology have been more and more integrated into and affect our life. At present, information technology has been widely used in all walks of life and has been fully applied in the marketing industry, especially in the digital marketing industry. With the advent of the 5G era, everything becomes connected, and the distance between time and space is rapidly compressed. Human beings are stepping towards the vision of a global village [1]. The development of artificial intelligence has greatly improved the operational efficiency of each industry and its segments. Especially, in recent years, the research and application of artificial intelligence in customer service, marketing, and other fields have gradually been deepened, bringing opportunities and challenges to this industry and market. In this era of rapid technological development, the pace of society's functioning, our ability to process information, and the impact of technology on social progress are all accelerating at an

unprecedented rate. With the change and development of artificial intelligence, big data, cloud computing, VR/AR (virtual reality), and other technologies, artificial intelligence technology group, as the focus of China's scientific and technological construction, has become a national strategy and has been applied in all walks of life, such as medical care, education, e-commerce marketing, and so on [2]. The "cooperation" between AI and the advertising industry will not only be the application and practice of AI technology but also the "optimization and reconstruction of advertising industry". With the deep power of artificial intelligence to advertising production, audience, product, and market, advertising marketing has ushered in the intelligent marketing 4.0 era [3]. Intelligent marketing is a new theory born with the application momentum of artificial intelligence; the academic circle has not formed a consistent definition of the concept, showing the trend of "benevolence and wisdom". At present, the industry gradually formed such a consensus. Smart marketing uses human creativity to create advanced computers, networks, mobile Internet, Internet of Things,

integrated technology, etc., and apply them to new thinking, new ideas, new methods, and new tools in the field of contemporary brand marketing. It includes intelligent matching, smart tags, intelligent acquisition and intelligent implementation aspects of [4].

In order to improve user recommendation satisfaction, improve advertising conversion rate prediction accuracy, and accurately analyze online marketing effects, this paper proposes an online marketing effect analysis method based on multimodel fusion and artificial intelligence algorithm in the context of big data and verifies the effectiveness of the method in this paper through experiments.

2. Analysis of Advertising Marketing Model in the Context of Big Data

2.1. Online Marketing Status of Big Data. There is no doubt that the arrival of big data has caused great changes in the ecology of various systems in the social field. Big data is like a “catfish,” stirring a pool of spring water. With the dual promotion of technology and demand, more and more government agencies, companies, and individuals will realize that data is a huge economic asset. Like money or gold, it will bring a new entrepreneurial direction, business model, and investment opportunities [5]. “Big data,” as an information mine in the network era, undoubtedly contains great value. According to statistics, different groups have their own characteristics in receiving information:

- (1) The audience of traditional media (TV and newspaper) is the post-1970s, who have their own long-term browsing habits and are difficult to change
- (2) For the post-1980s and post-1990s, the main channel to receive news information is the network

Especially the post-1990s people, who have strong purchasing power, have been exposed to new media such as the Internet since their birth. Their information acquisition methods are reflected in the fact that they like to browse, dislike reading, like online games, rely on virtual online emotions, and are willing to express their opinions in cyberspace [6]. In their life, entertainment occupies the main leisure time and has strong commercial development value.

At present, the relatively clear commercial value development of big data takes place in the field of Internet advertising and precision marketing. After experiencing the hustle and bustle of mass communication and the prosperity of mass communication, traditional advertising marketing began to usher in a new window of change, precision marketing for specific internet users (some people call it “mass communication”) [7]. At present, the changes brought by big data to advertising marketing are taking place. Then, what impact will the era of big data have on the advertising media industry? In the context of big data, what kind of marketing philosophy and wisdom should the advertising marketing model uphold? Furthermore, under the background of big data, what characteristics and development trends will the future online advertising marketing model present? What are the operational business models in

the field of big data? These problems, which we must attach great importance to, are also the focus of this study.

2.2. Big Data under the Characteristics of Network Advertising Marketing Model (RTB). In the era of big data, the network advertising marketing model has been changed, and a new model represented by RTB has emerged. So what is RTB? What is the difference between the RTB model and the traditional advertising marketing model? What are its characteristics? RTB advertising is a new model of network advertising marketing but also a new thing in the field of advertising. RTB is short for real-time bidding, which stands for real-time bidding [8]. It is a new type of advertising that conducts real-time bidding based on the exposure of each AD display. Compared with traditional PPC ads (charging per user), CPM ads (charging per thousand views), CPC ads (charging per click), Monthly Flat, Daily Flat, and so on, RTB bids on the basis of each AD display exposure [9]. The basic process can be described as follows: advertisers put forward a set of advertising positioning requirements to the Internet platform providers and agencies. The Internet platform providers find out which attributes match potential consumers based on consumers’ browsing history data on their platforms, and then ask advertisers, advertising agencies, provide feedback in accordance with the principles of auction and bidding transactions. The highest bidder gets to display ads on the platform. It is essentially the dynamic process of advertising and the diversification of means. This process of RTB, due to the tracking of user attributes and timely dynamic delivery, improves the conversion rate of advertising, increases the opportunity to display advertising, and becomes popular, winning the hearts of many advertisers. From the occurrence process of RTB mentioned above, the original advertising ecological chain has undergone subversive changes in the RTB advertising marketing model [10]. Compared with the traditional advertising marketing model, the biggest difference has four aspects.

First, it changed the relationship between advertisers and network platform operators and advertisers. In the RTB model, the relationship between network platform operators and advertisers is changed. Before 2002, most advertising operators mostly used the mode of breaking and fixed advertising space. In RTB model, advertisers have absolute initiative and discourse power. Advertisers no longer have to go to network platform operators and advertisers to seek cooperation; instead, they will directly provide one-to-one services in a timely manner. In essence, as long as advertisers have a clear advertising orientation, there is no need to consider the previous advertising location and channel [11]. However, there is little room for advertisers to choose from traditional PPC ads (charging per user), CPM ads (charging per thousand views), CPC ads (charging per click), Monthly Flat, and Daily Flat. And must discuss the rules of the game in advance, can be advertising. This is very unfair to advertisers.

Second, introduce real free market rules-real-time bidding. In the past, the traditional advertising marketing model is based on traffic, page views, the number of people,

or time and on the like. These data are actually not clear, and there is a possibility of falsehood. In addition, regardless of consumers; purchase intention and effect, it is a “one-size-fits-all” charging [12]. RTB model is the implementation of real-time bidding mode. The biggest advantage of this model is openness, justice, fairness, transparency, and efficiency. If the traditional advertising marketing model may be in the number of people, visits, and other aspects of the existence of “water”, so RTB model in this respect is an accurate guide, all “dry goods.” For example, if an advertiser wants to advertise to a specific group of users and another advertiser wants to advertise to that group, then you actually need to “bid.” In the bidding process, the party with the highest bid will get the advertisement [13].

Third, it directly tracks user needs, rather than the traditional sense of advertising space and billboards. Therefore, it improves the real effect and efficiency of advertising so as to realize the bidirectional benign interactive logic of advertisers “what I buy is what I want” and consumers “what you recommend is what I buy.” The consequence of this logic is that it will continuously bring a better delivery effect and continuously stimulate and guide consumption [14]. This is because every advertisement of advertisers is targeted and is based on the analysis of big data and the tracking of the group attributes of consumers; therefore, now more and more advertisers began to be interested in this kind of effective precise advertising. Although compared with the traditional advertising marketing model, the coverage rate and coverage of RTB model advertising may not be as large as before, but obviously, in the efficiency of advertising, the real effect, and the degree of accuracy, the RTB model is far superior to the previous advertising marketing model.

Fourth, it improves the relationship between advertising and consumers and shortens the distance between them. We have all had the experience that under the traditional advertising model, when we watch TV or a wonderful drama, some advertisements are often interrupted [15]. We often feel disgusted and revulsion and even turn off the TV and leaving. Studies have pointed out that: “The broadcast time of advertisements is positively correlated with the flow of sewage pipes. When the advertisements are broadcast, the flow of sewage surges, and after the advertisements end, the flow tends to normal.” As a result, this kind of advertising produces little effect. However, the RTB model is based on the accurate grasp of customers and the association of benefits. Therefore, according to consumers’ interest preference and shopping preference, some products that are exactly what consumers need are recommended. In this way, consumers will not be disgusted but will have positive emotions [16]. For example, we now have apps on our phones, such as maps. When users use mapping software, they want to obtain geographic information, including routes and so on. If at this time, based on the user’s search for geographic information, the user can also be presented with very convenient information such as food and lodging, shopping, and public transportation. In this case, users do not hate advertisements but will actively seek the information of the advertisement, which is beyond the reach

of the traditional advertising trading era, let alone unimaginable.

3. Research on Online Marketing Effect Based on Multimodel Fusion and Artificial Intelligence Algorithm in the Context of Big Data

3.1. Research Status of Multitasking Learning. Usually, most machines only learn one task at a time. When multiple interrelated tasks are studied together, it is called multitask learning (MTL). Multitask learning can improve the overall performance of the model by learning the effective information of multiple associated tasks. Multitasking learning is used in the early stage to improve the generalization ability of the model as a whole. Generally speaking, when we need to find a representation suitable for all tasks, the more difficult it is to fit with the original corpus representation, the higher the generalization ability. Caruana et al. summarized multitask learning as improving generalization ability by using domain-specific information contained in related tasks [17]. At the same time, Baxter and others also show that the risk is less than that of fitting specific task parameters by fitting multitask shared parameters. The most commonly used method in multitask learning is to share shallow representation, and its paper can be traced back to the literature. This paper greatly reduces the risk of overfitting by sharing shallow representation, but when the correlation between tasks is relatively loose, its sharing mechanism will fail. The literature also improves the generalization ability of the model through the idea of parameter sharing in a neural network.

Nowadays, multitask learning has become one of the most promising directions in machine learning. However, if machine learning wants to train a high-performance learner, it needs a lot of labeled data [18]. As a typical representative of machine learning, each training network in deep learning implies a large number of parameters to be learned. The high cost of manual labeling cannot meet the training needs. Using multitask learning method to obtain useful information between tasks has become a better choice.

In the field of natural language processing, multitask learning is mainly manifested in finding a better task level. Hashimoto et al. proposed an association model for multitask learning based on the hierarchical structure of some predetermined natural language processing tasks. Different from the learning sharing structure, Kendall et al. considered the uncertainty of the task by using the orthogonal method and further improved the accuracy of multitask learning by updating the weight value according to the derivative loss function. Wang and others collected and sorted out the text data of large retail stores. Aiming at the demand of sellers for buyers’ characteristic attributes in the offline transaction scenario, they proposed a multitask expression learning model. Compared with the single classification task, the model captures the association relationship between different tasks by sharing the underlying expression, effectively avoiding the problem of insufficient training data of some

single classification tasks. The model makes use of the deep network structure and obtains better user expression through the method of supervised learning. However, multitask learning for the underlying expression depends more on the accuracy of the original data expression. Using the text features of microblogs and comments published by users and the social features of users' personal information and paying attention to the relationship of concern, Chen Jing and others proposed an age regression method based on dual-channel LSTM, added a new level (merge layer) to the network, and learned the text feature representation and social feature representation sets generated by the two LSTM channels again. This fusion method overcomes the shortcomings of single model training, but its fusion method is the same model fusion of different data sources, which ignores the impact of the differences of different model training methods on the final fusion results to a certain extent. It can be seen that multitask learning and fusion technology are conducive to sharing information, strengthening learning ability, obtaining better generalization effect than single task and single model, so as to further improve the accuracy of results.

3.2. A Method for Measuring the Importance of Advertising Words. The fundamental of user attribute inference is the effective acquisition of user information, and advertising is the main carrier of user demand information. Therefore, how to extract the key information in advertising words is particularly important. The importance measurement method of advertising words is one of the ways to extract key information of the text. Term frequency-inverse document frequency (TF-IDF) is a simple and widely used measure of word importance, which is used to measure the contribution of an advertising word to its advertising.

Suppose the advertising set is D and N represents the total number of advertisements in D , and the calculation formula is as follows:

$$\begin{aligned}
 W_{TF-IDF} &= TF \times IDF, \\
 TF &= \frac{\text{count}(t, d)}{\text{size}(d)}, \\
 IDF &= \log\left(\frac{N}{\text{docs}(t, D)}\right),
 \end{aligned} \tag{1}$$

where TF (term frequency) represents word frequency, which is used to calculate the proportion of $\text{count}(t, d)$ of advertisement word t appearing in advertisement D and $\text{size}(d)$ of total word number in document D , that is, the frequency of occurrence in document D . IDF represents the inverse document frequency of the word t in the whole document set D and refers to the sparsity of the word's distribution in other advertisements. $\text{docs}(t, D)$ represents the number of documents containing the word t in document set D .

$TF-IDF$ word importance measurement algorithm can effectively extract the words that can highly summarize the content and theme of advertisements, namely keywords.

Different from the standardized keywords, keywords obtained through the importance measurement algorithm of advertising words can be obtained in advertising without being limited by the thesaurus, and the cost is small. However, the selection process can only be based on the statistics of the existing advertising words, without considering the semantic meaning of the text above and below the advertising words. At the same time, the statistics are completely dependent on the result of word segmentation and are not sensitive to the emergence of new words with periodic changes.

3.3. Vector Representation Model of Advertising Words word2vec. Different from the measurement of document importance, the vector representation model of advertising words is a description of document semantics. The difficulty of natural language processing lies in the complexity of semantics and representation. It is usually necessary to mathematicize the language. Vectorization is a good method. There are two common word vector representation methods: one hot representation and distributed representation. In one hot presentation, each advertising word can be represented as a vector whose dimension is equal to the size of the dictionary. Only one component of the vector is 1, the index position of the component in the dictionary, and all the other components are 0. However, this method cannot well describe the similarity between advertising words while being affected by the high dimension.

Distributed representation was first proposed by Hinton in 1986. It can obtain word vector representation with prespecified dimensions through training so as to transform words into machine computable form and solve the gap between advertising word semantics and machine language. Because the length of the advertisement vector is relatively shorter than that of one hot representation, the amount of calculation is relatively low. At the same time, the similarity between advertising words can be easily obtained by calculating the distance between advertising word vectors. word2vec is a popular way of representation. In 2013, Google proposed an open-source toolkit word2vec that can realize low-dimensional real vector representation of advertising words. Its simple and efficient processing method has attracted extensive attention in the industry after its release. There are two training methods: CBOW (continuous bag of words) and Skip gram. The two models are very similar. They remove the hidden layer in the neural probabilistic language model and only retain the three-layer network structure of the input layer, projection layer, and output layer. However, the CBOW model predicts the target word through the context information, while the Skip gram model predicts the context information through the target word. Usually, the choice of the model needs to be determined according to the size of the corpus. When the corpus size is relatively small, the CBOW model is generally selected because the CBOW model averages all advertising words in the context in the training process, which is similar to the smoothing of word vector; Skip-gram regards all contextual advertising words as a new result. When the amount of data is large, the effect is

often better. When it is impossible to determine which model is better, the combined training of the two models can achieve better results.

3.4. Skip-Gram Model. The framework diagram of the Skip-gram model is shown in Figure 1. The training method is to

$$\begin{aligned}
\vartheta &= \sum_{w \in C} \log p(\text{Context}(w) | w) = \sum_{w \in C} \log \prod_{j=2}^{\mu} p(f_j^u | v(w), \theta_{j-1}^u), \\
&= \sum_{w \in C} \log \prod_{u \in \text{Context}(w)} \prod_{j=2}^{\mu} \left\{ [\sigma(v(w)^T \theta_{j-1}^u)]^{1-f_j^u} \cdot [1 - \sigma(v(w)^T \theta_{j-1}^u)]^{f_j^u} \right\}, \\
&= \sum_{w \in C} \prod_{u \in \text{Context}(w)} \prod_{j=2}^{\mu} \left\{ (1-f_j^u) \cdot \log[\sigma(v(w)^T \theta_{j-1}^u)] + f_j^u \cdot \log[1 - \sigma(v(w)^T \theta_{j-1}^u)] \right\}.
\end{aligned} \tag{2}$$

Among them, the parameters to be optimized are the same as CBOW. So θ_{j-1}^u the updated formula is as follows:

$$v(w) := v(w) + \eta \prod_{u \in \text{Context}(w)} \prod_{j=2}^{\mu} \left\{ [1 - f_j^u - \sigma(v(w)^T \theta_{j-1}^u)] \theta_{j-1}^u \right\}. \tag{3}$$

The pseudocode of random gradient rise of the Skip-gram model is shown in Table 1.

3.5. Multitask and Multimodel Fusion Representation. In the model fusion stage, the classification results obtained by multiple models are fused, which can obtain a better classification effect than the single model training results and reduce the possibility of model overfitting. Single-task multimodel fusion can be divided into the following two stages:

The first stage: it is usually necessary to train every single classifier, and the training process adopts cross-validation to avoid the deviation of training results caused by uneven data distribution. Let the user document set be D and $D = \{(y_t, x_t), t = 1, \dots, T\}$, where x_t represents the historical search term of the t -th user and y_t represents its category (age 2, gender 6, and education 6).

If N -fold cross-validation is adopted, the user corpus D is randomly divided into N equal training subsets, which are D_1, D_2, \dots, D_N . At the same time, D_n is defined as the test set for n -fold cross-validation, and the training set $D^{\text{train}} = D - D_n$. Assuming that there are K classifiers in the first stage, which are C_1, C_2, \dots, C_K , the model obtained by training D^{train} through the K -th classifier C_K is $M_K, k = 1, \dots, K$. Then, for the sample users in each test set, the model will generate its corresponding inference results z_{tK} . When K classifiers complete training, a new data set for the second stage input will be formed as follows:

$$D' = \{(y_t, z_{t1}, z_{t2}, \dots, z_{tK}), t = 1, \dots, T\}. \tag{4}$$

predict contextual information through current words. The objective function of Skip-gram is as follows:

According to the hierarchical softmax framework, similar to the CBOW model, its log-likelihood function can be expanded into

The second stage: the new data set D' formed by the training combination in the first stage is added to the learning in the second stage as a feature vector, and a classification algorithm Ψ is used to classify the new data to obtain the model M' , which is used to represent the relationship between the inferred value and the real result in the first stage. The pseudocode of the model is shown in Table 2.

The above fusion framework can improve the classification effect only for a single task. In order to achieve the mutual restriction between user multitask factors, on this basis, this paper adds the above fusion framework to multitask factors, which is conducive to sharing the shallow representation of users and forming a multitask fusion framework for user attribute inference. The specific user attribute inference method is shown in Figure 2. The two stages of this framework complete the following tasks, respectively: in the first stage, single model inference, user-level vector representation is realized by using the learning methods based on text semantics (multi_DBOW and multi_DM) and keyword based on text word frequency (NW_TF-IDF) proposed in Section 3 according to user data, and then M feature distribution probabilities of users are trained through the model. In the second stage, based on multitask and multimodel fusion inference, that is, the results of each representation in the first stage are combined as a new overall representation of the user, the classification training in the second stage is carried out through the classification model, and the attribute values of multiple tasks of the user are finally obtained through fusion learning.

When H starts to work, $Q_1, Q_2, Q_3, \dots, Q_h$ represents the data parameter. Each task has S_1, S_2, \dots, S_h classification tags, a new data set dimension can be obtained after the first stage of training, which is the sum of the number of classification tags of each task, that is, the user vector representation of $(S_1 + S_2 + \dots + S_h) * H$ dimension. In the second stage, the multitask vector generated in the first stage training is used as the input data in the second stage and then trained through XGBoost. Because the user representation in the second stage involves all tasks, shallow semantic sharing

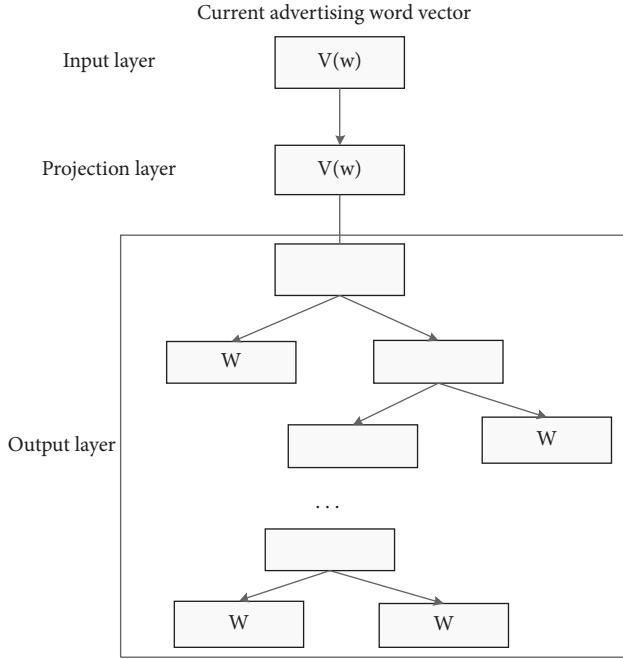


FIGURE 1: Schematic diagram of Skip-gram model structure.

TABLE 1: Description of Skip-gram model algorithm.

Algorithm	Skip-gram model stochastic gradient ascent algorithm
	For $u \in \text{Context}(w)$ do
	$k = 0$
	For $j = 2:u$ l do
	$q = \sigma(v(w)^T \theta_{j-1}^u)$
	$g = \eta(1 - f_j^u - q)$
	$k := k + g \theta_{j-1}^u$
	$\theta_{j-1}^u := \theta_{j-1}^u + gv(w)$
	End for
	$v(w) := v(w) + k$
	End for

can be realized in the parameter training process in this stage. It is conducive to the joint learning of related tasks. The multitask ensemble inference model framework is shown in Figure 2.

3.6. Multitask Fusion Model Classification. In the second stage, the classification model used in this paper is XGBoost (extreme gradient boosting). XGBoost is one of the boosting algorithms. Based on the gradient lifting decision tree, the goal is to establish K regression trees to make the predicted value of the tree group as close as possible to the real value. Mathematically, the objective function of XGBoost can be defined as follows:

$$L(\phi) = \sum_i l(y'_i - y_i) + \sum_j \Omega(f_i), \quad (5)$$

where i represents the i -th sample and $l(y'_i - y_i)$ represents the prediction error of the i -th sample. The error value should be kept as small as possible in training. $\sum_j \Omega(f_i)$ is a complexity function to measure the

TABLE 2: Ensemble algorithm description.

Input: train data $D = \{(y_t, x_t), t = 1, \dots, T\}$
One-level classification algorithm C_1, C_2, \dots, C_K
Two-level classification algorithm Ψ
For $k = 1, 2, \dots, K$ do
$M_k = C_k(D)$
end for
$D' = \emptyset$
For $t = 1, 2, 3, \dots, T$ do
For $k = 1, 2, 3, \dots, K$ do
$z_{tK} = M_k(X_t)$;
end for
$D' = D' \cup ((y_t, z_{t1}, z_{t2}, \dots, z_{tK}))$;
end for
$M' = \Psi(D')$;
Output: $M(x) = M'(M_1(x), M_2(x), \dots, M_K(x))$

generalization ability of the model. The lower the complexity, the stronger the generalization ability. The expression is as follows:

$$\Omega(f_i) = \gamma N + \frac{1}{2} \lambda \|v\|^2, \quad (6)$$

where N represents the number of leaf nodes in the tree and v represents the value of nodes. When the tree is a regression tree, it corresponds to the corresponding value. If it is a classification tree, it corresponds to a classification label. In this experiment, the tree is a regression tree, corresponding to the corresponding values for fusion. XGBoost forms a tree group by splitting, and the tree formed each time is obtained by splitting or building a tree based on the best prediction of the last time. Each process is similar to the greedy algorithm. The advantage of using XGBoost is that regularization is added to the objective function to prevent overfitting of the training model. At the same time, parallel processing can be realized when selecting the best splitting point, which is conducive to improving the operational efficiency. Therefore, the research on online marketing effect based on multimodel fusion and artificial intelligence algorithm under the background of big data is realized.

4. Experiment

4.1. Introduction to Experimental Environment and Tools. *Python* is a powerful and perfect programming language. With its advantages of easy to read and rich tool library, it has been favored by data scientists and is widely used in the field of data science. This paper uses *Python* language to code each link of the prediction method and then verifies the prediction effect through experiments. The development environment and various important tool libraries used in the experiment are introduced as follows.

4.1.1. Spyder. There are many integrated development environments (ides) supporting *Python* language, such as Pycharm, Spyder, Eclipse, and so on. Spyder is specially developed for the field of data science. Its interface is simple,

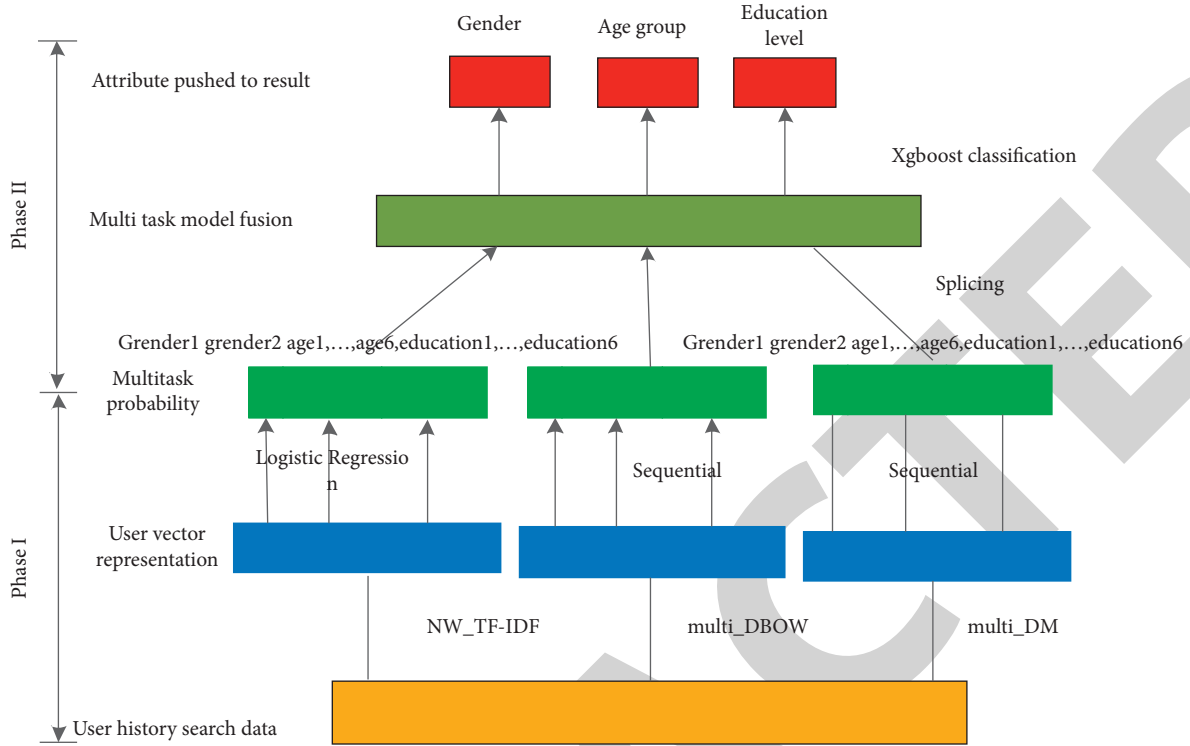


FIGURE 2: Multitask ensemble inference model framework.

and its function partition is obvious. It can view the data in the current memory in real time. It also has built-in necessary tools for data science such as Numpy and Pandas.

4.1.2. Matplotlib. Matplotlib is a 2D drawing library in Python. It can draw various statistical charts such as column chart, broken line chart, and box chart, which provides a solid foundation for exploratory data analysis in data science. When Matplotlib and Numpy are used together, they can effectively replace the functions of MATLAB.

4.1.3. Scikit-Learn. Scikit-learn, also known as Sklearn, is the most famous machine learning toolkit at present, with built-in implementation functions of most machine learning algorithms. In addition to supervised learning and unsupervised learning tasks through algorithms, the basic functions of Sklearn also include data preprocessing, data dimensionality reduction, model selection, and so on. Sklearn is an indispensable part of data science research.

4.1.4. X Learn. X learn is also an integrated machine learning algorithm library, which mainly includes LR, FM, FFM, and other algorithms commonly used for online advertising conversion rate prediction. Compared with the traditional toolkits that also support these algorithms such as liblinear, libfm, libffm, X learn not only ensures performance but also greatly improves time efficiency.

4.1.5. XGBoost. XGBoost algorithm library is an open-source toolkit developed to realize the functions of XGBoost algorithm. It is suitable for a variety of language environments including Python and applicable to various operating systems. In addition, XGBoost library also supports various distributed processing frameworks, and its excellent performance is widely praised.

4.2. Experimental Index. In this paper, the logarithmic loss function (logistic loss, logloss) is used as the evaluation index to measure the effect of prediction. Its standard definition form is as follows:

$$\log loss = -\log P\left(\frac{Y}{X}\right). \quad (7)$$

The prediction of online advertising conversion rate belongs to one of the two classification problems, and the corresponding calculation formula is as follows:

$$\log loss = \frac{1}{G} \sum_i^G (y_i \cdot \log p_i + (1 - y_i) \log(1 - p_i)), \quad (8)$$

where G represents the total number of samples in the measurement set, y_i represents the real label of the i -th sample, and p_i represents the probability that the prediction label of the i -th sample is 1. The logloss value can be used to judge the difference between the real probability distribution and the predicted probability distribution. The smaller the value, the more accurate the prediction result. For the test set, the prediction results

TABLE 3: Comparison of prediction results of single model and ensemble model.

Algorithm model	Validation set logloss	Test set logloss
GBDT + LR	0.1728712	0.14264
GBDT + FM	0.1731887	0.14275
XGBoost	0.1718231	0.14128
Multimodel fusion	0.1705332	0.13990

TABLE 4: Comparison of product online marketing effect (10,000 pieces).

Sales time/day	Online sales of goods (10,000 pieces)			
	GBDT + LR algorithm	GBDT + FM algorithm	XGBoost algorithm	Multimodel fusion algorithm
1	0.15	0.08	0.12	1.32
5	0.65	0.11	0.37	2.97
15	0.87	0.28	0.65	5.57
30	1.15	0.56	0.89	6.12
60	1.32	0.65	1.27	7.54
90	1.87	0.86	1.65	8.68

can be submitted to the Alibaba Tianchi platform for online verification.

4.3. Experimental Result

4.3.1. Prediction Results of Single Model and Fusion Model.

The single model is constructed for the gGBDT + LR algorithm, GBDT + FM algorithm, XGBoost algorithm, and multimodel fusion algorithm. The same training set is used to measure the results together with the loss values of the verification and test sets. The experimental results are shown in Table 3.

According to Table 3, XGBoost has the best effect among the three algorithms, followed by GBDT + LR, and GBDT + FM is slightly inferior to GBDT + LR, but the gap is very small. It can be considered that the effects of the latter two are basically the same. Because the data set used in this paper has time factors, if Stacking is used for fusion, it may lead to data crossing, so the fusion method adopts weighted average fusion. XGBoost becomes the most important model with the lowest predicted loss value and gives the largest weight value. In the first mock exam, we found that the prediction results of the fusion model were better than the single model. When the XGBoost weights were 0.6, the weights of GBDT + LR and GBDT + FM were 0.2; the results obtained from multimodel fusion had lower loss values on the validation set and test set, and the value was the final result of the multimodel fusion prediction method.

4.3.2. Comparison of Product Online Marketing Effects.

In order to verify the effect of this method in commodity online marketing, GBDT + LR algorithm, GBDT + FM algorithm, XGBoost algorithm, and multimodel fusion algorithm are used to obtain the online sales of various commodities. The results are shown in Table 4.

By analyzing Table 4, we can see that there are differences in the effect of commodity online marketing under different algorithms. When the sales time is 1 day, the online

marketing sales volume of goods of GBDT + LR algorithm is 1,500 pieces; that of GBDT + FM algorithm is 8,000 pieces; that of XGBoost algorithm is 12,000 pieces; and that of multimodel fusion algorithm is 13,200 pieces. When the sales time is 60 days, the online marketing sales volume of the GBDT + LR algorithm is 13,200 pieces; that of GBDT + FM algorithm is 6,500 pieces; that of XGBoost algorithm is 12,700 pieces; and that of multimodel fusion algorithm is 75,400 pieces. The first mock exam is more efficient than other single model algorithms, which indicates that this method can improve online marketing and online marketing.

5. Conclusion

This paper proposes an online marketing method based on multimodel fusion and artificial intelligence algorithms under the background of big data. Big data technology is introduced; the characteristics of online advertising marketing mode are analyzed; the importance of advertising words in online marketing by using TF-IDF technology in artificial intelligence algorithm is measured, and the online marketing effect by multitask fusion model according to XGBoost technology is classified; and experiments are used to analyze the effect of online marketing. The following conclusions are drawn through experiments:

- (1) In the first mock exam, we found that the prediction results of the fusion model were better than the single model. When the XGBoost weights were 0.6 and the weights of GBDT + LR and GBDT + FM were 0.2, the results obtained from multimodel fusion had lower loss values on the validation and test sets, and the value was the final result of the multimodel fusion prediction method.
- (2) When the sales time is 60 days, the product online marketing sales of multimodel fusion algorithm is 75,400 pieces. The first mock exam is more efficient than other single model algorithms, which indicates

that this method can improve online marketing and online marketing.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] B. Mjla, Z. A. Jian, C. Rmw, and L. Qi, "Simulation research on online marketing strategies of branded agricultural products based on the difference in opinion leader attitudes," *Information Processing in Agriculture*, vol. 12, no. 21, pp. 132–154, 2020.
- [2] R. Swathika and T. S. Sharmila, "Multi-model fusion based satellite image classification using versatile unsupervised vector zone (VUVZ) fusion and intensive pragmatic blossoms (IPB) technique," *Multimedia Tools and Applications*, vol. 79, no. 6, pp. 4239–4260, 2020.
- [3] H. Liao and S. B. Tsai, "Engineering M . Research on the B2C online marketing effect based on the," *LS-SVM Algorithm and Multimodel Fusion*, vol. 16, no. 32, pp. 54–68, 2021.
- [4] R. T. Kreuzer, A. Rumler, and B. Wille-Baumkauff, *Instrumente des Online-Marketings*, vol. 45, no. 43, pp. 76–88, 2020.
- [5] T. Cui, Y. Wang, and B. Namih, "Build an intelligent online marketing system: an overview," *IEEE internet computing*, vol. 23, no. 4, pp. 53–60, 2019.
- [6] P. Lagrée, O. Cappé, B. Cautis, and S. Maniu, "Algorithms for online influencer marketing," *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 1, pp. 1–30, 2019.
- [7] S. C. e Silva, P. Duarte, J. C. Machado, and C. Martins, "Cause-related marketing in online environment: the role of brand-cause fit, perceived value, and trust," *International Review on Public and Nonprofit Marketing*, vol. 17, no. 2, pp. 135–157, 2020.
- [8] S. Silva, P. Duarte, J. C. Machado, and C. Martins, "Cause-related marketing in online environment: the role of brand-cause fit, perceived value, and trust," *International Review on Public and Nonprofit Marketing*, vol. 20, no. 2, pp. 17–36, 2020.
- [9] V. Herrera, E. Carrillo, S. Herrera, and F. R. Castillo Villar, "Influence of experiential marketing on online engagement of the consumer in the fashion industry in the city of aguas-calientes," *Advances in Management and Applied Economics*, vol. 16, no. 24, pp. 10–19, 2020.
- [10] S. G. Li, Y. Q. Zhang, Z. X. Yu, and F. Liu, "Economical user-generated content (UGC) marketing for online stores based on a fine-grained joint model of the consumer purchase decision process," *Electronic Commerce Research*, vol. 21, no. 4, pp. 1083–1112, 2020.
- [11] W. Li and R. Huang, "Research on the model of real estate online community marketing based on big data technology," *Journal of Technical Economics & Management*, no. 9, pp. 28–32, 2019.
- [12] A. Abid, P. Harrigan, and S. K. Roy, "Online relationship marketing through content creation and curation," *Marketing Intelligence & Planning*, vol. 38, no. 6, pp. 699–712, 2019.
- [13] E. Naryono, "Aida analysis (attention, interest, desire, action) toward advertising effectiveness as A basic online shop marketing strategy in sukabumi city," *OSF Preprints*, vol. 54, no. 6, pp. 454–466, 2019.
- [14] J. Salminen, V. Yoganathan, J. Corporan, B. J. Jansen, and S.-G. Jung, "Machine learning approach to auto-tagging online content for content marketing efficiency: a comparative analysis between methods and content type," *Journal of Business Research*, vol. 101, no. 8, pp. 203–217, 2019.
- [15] Y. Xie, L. I. Chunqing, P. Gao, and Y. Liu, "The effect and mechanism of social presence in live marketing on online herd consumption from behavioral and neurophysiological perspectives," *Advances in Psychological Science*, vol. 27, no. 6, p. 990, 2019.
- [16] J. A. Castaeda, D. M. Frías-Jamilena, M. A. Rodríguez-Molina, and A. Jones, "Online Marketing Effectiveness - the influence of information load and digital literacy, a cross-country comparison," *Electronic Markets*, vol. 30, no. 2, pp. 78–89, 2019.
- [17] Z. Yu, L. Chang, and B. Qian, "A belief-rule-based model for information fusion with insufficient multi-sensor data and domain knowledge using evolutionary algorithms with operator recommendations," *Soft Computing*, vol. 23, no. 13, pp. 5129–5142, 2019.
- [18] Y. Geng, X. Wang, and P. Jiang, "Prediction of the cement grate cooler pressure in the cooling process based on a multi-model fusion neural network," *IEEE Access*, vol. 8, pp. 115028–115040, 2020.