Heliyon 10 (2024) e37252

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Computing advertising intelligent computing and push based on artificial intelligence in the big data era

Lin Shan

College of Humanities and Communication, Sanya University, Sanya, 572000, Hainan, China

ARTICLE INFO

Keywords:
Advertising push
Artificial intelligence
Computational advertising
Big data
Return on investment

ABSTRACT

In the era of big data, intelligent computation and precise targeting of advertisements rely on artificial intelligence (AI) technology to enhance advertising effectiveness and user experience. As advertising media shift from traditional outlets to the internet and mobile devices, manually filtering and targeting advertisements has become increasingly ineffective due to the vast amount of information and user groups. Consequently, AI-based advertisement computation and targeting technologies have emerged. This paper explores how modern technologies and AI applications can be used to achieve intelligent computation, filtering, and targeting of advertisements, thereby improving ad effectiveness and profitability. The study optimizes ad targeting effectiveness and increases return on investment (ROI) by comparing ad campaigns between Group A, which did not use ad targeting algorithms, and Group B, which employed AI-based targeting algorithms. The experimental results show that the average ROI for tourism, shopping, and rental ads in Group A were 153.03 %, 232.32 %, and 192.57 %, respectively, while Group B's average ROI were 173.96 %, 288.74 %, and 216.12 %. These results indicate that AI-based ad targeting algorithms can significantly improve the ROI of ad campaigns compared to non-algorithmic approaches, suggesting that ad targeting algorithms can help advertisers achieve higher profits.

1. Introduction

With the popularization of the Internet, advertising has become an indispensable part of corporate marketing. The arrival of the "Big data era" has made enterprises face the situation of information explosion when advertising [1]. To find and target target customers in such a huge data set and maximize the advertising effect and ROI, advertising intelligence push technology based on AI and Big data technology should be used [2,3]. Through this technology, it is possible to analyze and model user data, accurately push advertisements that satisfies user interests and needs and help enterprises increase exposure rate and click rate to improve revenue.

The significance of advertising push lies in its ability to help businesses better promote products or services to potential customers, accelerate sales speed and enhance brand awareness. Under the pressure of soaring advertising costs and fierce market competition, more and more online merchants are trying to implement new advertising strategies to improve advertising efficiency and increase sales. Hu, L. and colleagues investigated how e-commerce platforms can leverage membership marketing systems to enhance targeted advertising for SMEs. They found that a narrow fee gap between marketing windows under different cost-bearing mechanisms increases advertising investment and effectiveness. Additionally, they proposed strategies for optimizing the platform's membership marketing system, including discount adjustments and product screening [4]. Gal-Or Esther believes that when a single media content distributor delivers advertising messages on behalf of multiple competing brands, sometimes customized advertising can be used to

E-mail address: linshan@sanyau.edu.cn.

https://doi.org/10.1016/j.heliyon.2024.e37252
Received 3 June 2024; Received in revised form 22 August 2024; Accepted 29 August 2024
Available online 30 August 2024

implement monopoly pricing. Although this monopoly pricing can be achieved through different degrees of commercial advertising customization, when distributors choose the highest feasible level of customization, product revenue and consumer surplus are the highest [5]. When influential individuals post genuine product advertising push, it can reduce advertising awareness and suspicion. Besides, with the continuous growth of digital and social media, the advertising industry is forced to adopt innovative strategies, one of which is influence marketing. Advertising push need to transform the agency process to effectively implement marketing strategies and ensure the success of influence marketing [6,7]. Liu Duen-Ren pointed out that online advertising has brought huge revenue to many websites. He proposed a new advertising push method that considered both the fairness of advertising push and personal interests. His method considered the exposure time of each advertisement and investigated the different user experiences on the website to determine the factors that affected user interest. The experimental results showed that his advertising push method had better performance than traditional methods, could improve the average click rate of all advertisements on the website and ensure the reasonable and fair exposure frequency of each advertisement [8]. Hu Li constructed a stylized model to check whether disclosing or hiding member-only discounts will generate higher revenue, and also incorporated the two-way reference price effect into the customer's utility formation. By comparing the income gap between exogenous and endogenous member-specific discount information disclosure and hiding strategies, it was found that platforms targeting the high-end market charge higher product prices and membership fees, and platforms with higher negative prices but lower regular prices for target customers should implement information hiding strategies [9]. In the above research literature, experts have conducted detailed research on "advertising push" and provided many valuable information.

AI technology plays an important role in advertising push. It can accurately push advertisements and improve the click rate and conversion rate of advertisements by analyzing users' interests, preferences, behavior and other data. WANG Yuan-mei thought that with the proliferation of text information on the Internet and the development of AI technology, the application of sentiment analysis to personalized push systems had become an important part of e-commerce, short videos, online advertising push and other fields. He proposed a ternary sentiment analysis method, which could generate interpretable recommendation items. The application effect of this method was good, greatly improving the interpretability and user experience of the push system [10]. Zhu Sandy believed that intelligent technologies such as AI and big data were booming and various application scenarios had also been launched. Enterprises could use precision marketing to better reach target consumers and push advertisements at appropriate times to promote products and services that consumers were interested in, thereby improving marketing efficiency. Precision marketing had become one of the key areas in the digital marketing industry. He started with an overview of intelligent data, reviewed its definition and development status and analyzed its application in digital advertising marketing [11]. After reading the above articles, it can be seen that these articles have conducted fusion research on AI technology and advertising push. However, in these studies, only theoretical research has been conducted.

In this article, advertising push technology is optimized through an advertising intelligent push algorithm based on AI technology [12–14]. The experimental results showed that the average advertising matching accuracy of the unused advertising push algorithm (Group A) is 64.124 %. The average advertising matching accuracy of the advertising intelligent push algorithm (Group B) based on AI technology is 74.988 %. This indicates that advertising intelligent push algorithms based on AI technology can significantly improve the accuracy of advertising matching, improve the effectiveness of advertising placement and improve user experience. The innovation of this article lies in applying AI technology to the field of advertising intelligent push. Through an advertising intelligent push algorithm based on AI technology, the effectiveness and superiority of the algorithm have been proven through experiments. This can help to recommend advertisements more intelligently and efficiently, improve advertising effectiveness and user experience. Big data era based on the calculation of artificial intelligence advertising intelligent computing and push innovation mainly embodied in the following aspects: 1. Accurate orientation and personalized recommendation, use of large data technology of massive user data analysis, identify the user's interest, demand and behavior patterns, so as to realize personalized advertising, advertising to truly interested user groups.2. Intelligent creative generation: With the help of AI, rich creative advertising copy and design can be automatically generated, reducing the time and cost of manual creation.

This article first introduces the research background, and then studies the advertising intelligent computing push technology in the second chapter. This part mainly includes data collection and processing in advertising push calculation, user model construction in advertising push calculation, prediction and recommendation algorithms in advertising push calculation; the third chapter mainly talks about calculating advertising push strategies, this part mainly includes targeted advertising technology and advertising push algorithms based on artificial intelligence technology; the fourth chapter mainly introduces advertising push experiments based on artificial intelligence technology, this part mainly includes the selection of advertising data sets, experimental evaluation indicators, and experimental results; and finally summarizes.

2. Advertising intelligent computing push technology

In the study of advertising intelligent computing and push technology, the data types include user behavior data, advertising data, social media data, and search data. User behavior data encompasses click records, browsing history, search keywords, and shopping cart activity, which are obtained from the company's own websites and mobile applications, with some data sourced from third-party data providers and advertising platforms. Advertising data involves metrics such as ad display frequency, click-through rates, and conversion rates, which are extracted from advertising platforms and marketing tools. Social media data includes user interaction information on social platforms, such as likes, comments, and shares, which helps in understanding user interests and social behavior. Search data refers to user query records from search engines, obtained from search engine data providers. During data collection, this study adheres to the Personal Information Protection Law and the Cybersecurity Law to ensure user privacy is protected and also complies with the policy requirements of data providers.

Data processing and cleansing are critical steps, including removing duplicate data, correcting format errors, filling in missing values, and data structuring. Removing duplicate data ensures the uniqueness and accuracy of the data, while correcting format errors, such as unifying date formats, prevents impacts on subsequent analysis results. Missing values are handled using mean imputation methods to reduce their impact on analysis. Data structuring involves converting raw data into tabular formats and filtering out irrelevant information to enable more effective management and analysis of the data.

Regarding the choice and application of artificial intelligence algorithms, Support Vector Machines (SVM) and Deep Neural Networks (DNN) are selected as the primary modeling algorithms. SVM is used for classification tasks such as predicting ad clicks, while DNN is employed for more complex tasks such as user profiling and ad recommendation. The model training process includes data preparation, feature extraction, and training. Data preparation involves splitting the data into training, validation, and test sets to ensure comprehensive and effective model training. Feature extraction includes deriving key features from user behavior and advertising data, such as user interest labels and ad visual features. During training, the model is trained using the training set, and parameters are adjusted to optimize performance, with hyperparameters determined through cross-validation.

2.1. Data collection and processing in advertising push calculation

In advertising push calculation, data collection and processing are crucial steps that determine the quality and effectiveness of advertising push [15,16]. Data collection can collect user behavior data, interests and other useful information to better understand user needs. Data processing can help extract valuable information from massive amounts of data, analyze and model it and thus bring more accurate results to advertising push. The following is a detailed explanation of the data collection and processing steps.

2.1.1. Data collection

The process of data collection is to collect various data required for advertising push for subsequent processing and analysis. In this process, data needs to be obtained from multiple channels, including user behavior data, advertising data, social media data, search data and other sources of data. To effectively collect data on advertising push, the following points need to be considered:

Determining the required data type and source: Before starting data collection, it is necessary to clarify which data is needed and from which sources it can be obtained. For example, if user behavior data is required, it can be obtained from one's own website or application or purchased from third-party data providers or advertising platforms.

Ensuring data collection compliance: During the data collection process, relevant laws, regulations and privacy policies must be followed. This includes but is not limited to complying with relevant laws and regulations such as Personal Information Protection Law and Cybersecurity Law, as well as the policy requirements in the location of the enterprise and the user.

Maintaining the accuracy and real-time of data collection: Data collection should be an ongoing process that requires regular checks of data accuracy and real-time, necessary corrections and updates.

2.1.2. Data cleansing

Data cleansing refers to the process of making data more standardized and useful by detecting, repairing or deleting damaged, inaccurate, incomplete, duplicate or irrelevant records. After collecting the data required for advertising push, data cleansing is required to ensure the quality of the data. This step is a committed step to ensure data quality and a necessary prerequisite for data analysis. Only after data cleansing, can the data be more accurate and complete, then providing reliable data basis for enterprises to correctly formulate advertising push strategies. Data cleansing includes the following aspects:

Removing duplicate data: When collecting data, there may be duplicate data that can affect the accuracy of the analysis and model. Therefore, when performing Data cleansing, duplicate data should be removed.

Correcting formatting errors: Incorrect data format can affect the accuracy of data processing and analysis. For example, inconsistent data types and inconsistent date formats all need to be corrected.

Filling in missing values: In some cases, there may be missing values in the data, such as null or unknown values, which can affect data accuracy. To maintain data integrity, these missing values should be filled in during data cleansing.

2.1.3. Data structuring

Data structuring is the process of reorganizing and processing the original data according to certain rules and methods, with the purpose of more convenient and efficient management, storage and use of data. This process usually includes data cleansing, conversion, integration and other steps to enable data to be classified, sorted, retrieved and analyzed according to certain rules. By structuring data, it is possible to efficiently utilize data resources, obtain more information and value, standardize data management and usage, reduce error rates in data processing and improve work efficiency. Here is a detailed introduction to these methods: Converting data into table format: This is the most common method and one of the basic forms of data structuring. By decomposing the original data into rows and columns according to certain rules, it can be converted into table form for easy processing and analysis.

Eliminating unnecessary information: There may be a lot of irrelevant information in the original data, such as noise, abnormal data, etc. This information may interfere with the accuracy of the calculation results, so it needs to be filtered and eliminated.

2.1.4. Data dissection

Data analysis is a crucial step in the advertising push process. Through data analysis, useful information about users and advertisements can be obtained, thereby improving the effectiveness of advertising push and making them more in line with user needs.

Specifically, data analysis involves the following aspects:

User behavior analysis: This analysis mainly focuses on the user behavior data on mobile devices, such as user search history, text input, scrolling traffic, etc. By collecting and analyzing these data, it is possible to understand the users 'interests, and the types of ads they may be interested in, and then recommend the ads that meet the users' needs and interests [17]. Advertising effect analysis: Data such as click through volume, conversion rate and browsing volume of advertisements is analyzed. By collecting and analyzing this data, it is possible to understand how advertisements perform and whether they can attract users' attention, promoting their purchasing behavior [18,19].This can help optimize and set various types of advertising content to achieve higher click rate and conversion rate.

The detailed steps for collecting and processing advertising data are shown in Fig. 1.

2.2. User profile construction in advertising push calculation

The construction of user profiles in advertising push calculation refers to the collection and integration of multi-dimensional data [20,21]. By describing the basic information, behavioral habits, interests and preferences of users, a user database is established and users are classified, to more accurately match user needs during advertising push, improving the effectiveness of advertising placement and user conversion rate. In detail, the steps for constructing user profiles are as follows.

2.2.1. User data integration

User data integration is one of the important steps in constructing user profiles, which requires collecting various aspects of user data through various channels, including basic information, behavioral habits, interest preferences, etc [22,23]. Among them, basic information can include the user's name, gender, age, occupation, etc. Behavioral habits can contain users' browsing history, search records and purchasing behavior on various platforms. Interest preferences can include users' favorite movies, music, books, travel destinations, etc. After collecting various aspects of user data, deduplication, error correction and standardization processing should be performed to ensure the accuracy and reliability of the data.

Finally, the cleaned data should be integrated into a user profile to establish a user database. User profile is a visual description constructed based on user data, which includes various aspects of user information, behavioral habits, interests and preferences and can comprehensively and accurately reflect the characteristics and needs of users. The user database is a collection of user profiles that can be easily managed and analyzed, providing support for applications such as advertising push.

2.2.2. User data dissection

User data analysis requires the use of data mining and other technologies to conduct in-depth analysis of user profiles, to extract information such as user feature labels and behavior patterns. When conducting user data analysis, the type and quantity of user data must be first understood. Common types of user data include quantitative data (such as user age, purchase amount, etc.) and qualitative data (such as user preferences, hobbies, etc.) and understanding the different characteristics of these data types is crucial for subsequent analysis [24,25].

After completing the processing of basic data, various algorithms and models can be used to analyze user profiles. For example, clustering analysis and other techniques can be used to divide users into different groups, with each group representing a type of user. Association analysis and other techniques can also be used to explore the behavioral connections between users and understand their preferences and behavioral patterns. In addition, algorithms such as decision trees and support vector machines can also be used to construct user classification models to achieve more accurate user profile construction. Ultimately, through user data analysis, information such as feature labels and behavior patterns can be extracted, further optimizing the construction process of user profiles and improving the accuracy and effectiveness of advertising push.

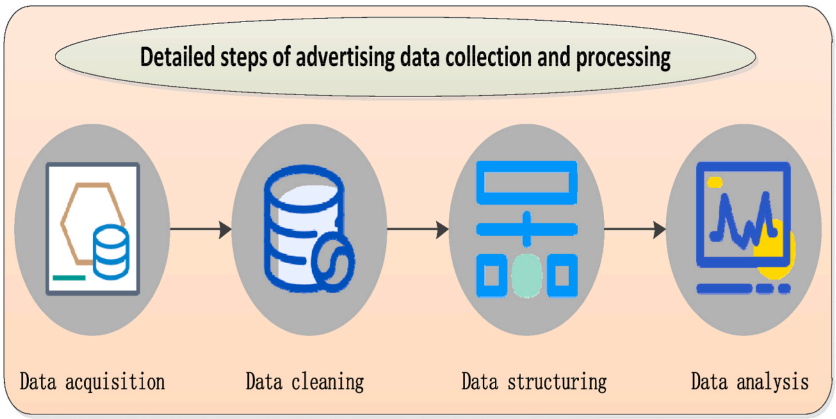

Fig. 1. Detailed steps for advertising data collection and processing.

2.2.3. User profile classification

User profile classification is the process of classifying users based on their feature labels and behavior patterns. By analyzing and mining user data, a large amount of valuable user information can be obtained, such as user age, gender, consumption preferences, interests, social circles, etc. This information can be used to construct user profiles and for classification.

In the process of user model classification, various methods can be used, such as cluster analysis, association rule mining, decision tree analysis, etc. These are the commonly used classification methods. In this paper, cluster analysis is selected for classification, which can divide users with similar characteristics into the same group to better understand their needs and recommend a more personalized product or service [26–28].

2.2.4. User profile optimization

User profile optimization refers to the process of adjusting and improving user profiles based on indicators such as advertising push effectiveness and conversion rate. By continuously iterating and refining user profiles, the accuracy and effectiveness of advertising push can be improved, thereby achieving more accurate and effective advertising marketing. The core of user profile optimization lies in data abstraction and generalization. Firstly, it is necessary to abstract and summarize user data and extract useful features and information, to facilitate more accurate user classification and behavior prediction. Secondly, technologies such as data mining and machine learning are utilized to analyze and optimize user profiles, while combining with actual marketing scenarios to quickly and accurately identify target users and provide personalized services and advertisements for them.

The construction steps of user profiles are shown in Fig. 2.

2.3. Prediction and recommendation algorithms in advertising push calculation

In advertising push calculation, prediction and push algorithms are very important [29,30]. These algorithms utilize machine learning and deep learning techniques to analyze user profiles and advertising materials, predict user reactions to advertisements and make personalized recommendations. Specifically, these typically take the following steps.

2.3.1. Modeling and training

In advertising push calculation, modeling and training are the core parts of prediction and push algorithms. This process is based on user profiles and advertising material data, establishing a model and training it to predict user reactions to advertisements [31]. The modeling process typically involves the following steps. Firstly, a large amount of user data is collected, including information such as user behavior, preferences, interests and purchase history and user profiles are constructed based on this data [32,33]. Secondly, advertising materials are analyzed, including advertising content, language style, images, audio and video and features are extracted from them. Next, the features of user profiles and advertising materials are taken as inputs and machine learning and deep learning techniques are used to establish a model and conduct training. The purpose of training the model is to learn the correlation between user profiles and advertising materials, thereby predicting the user's response to a certain advertisement.

The trained model can be applied to practical scenarios to predict users' reactions to advertisements and recommend the advertisements that best meet their needs and preferences based on the predicted results [34,35]. As the collected data becomes more abundant, the model can also be continuously improved and optimized to improve accuracy and effectiveness.

2.3.2. Model evaluation and adjustment

After modeling and training, the predictive performance of the trained model should be evaluated and tested. Simultaneously, existing problems should be identified, adjusted and optimized to improve the accuracy and robustness of the model.

Firstly, appropriate evaluation indicators should be selected to measure the performance of the model. Common model evaluation indicators include accuracy, recall and so on. Then, a test dataset is used to validate the predictive performance of the model, which can

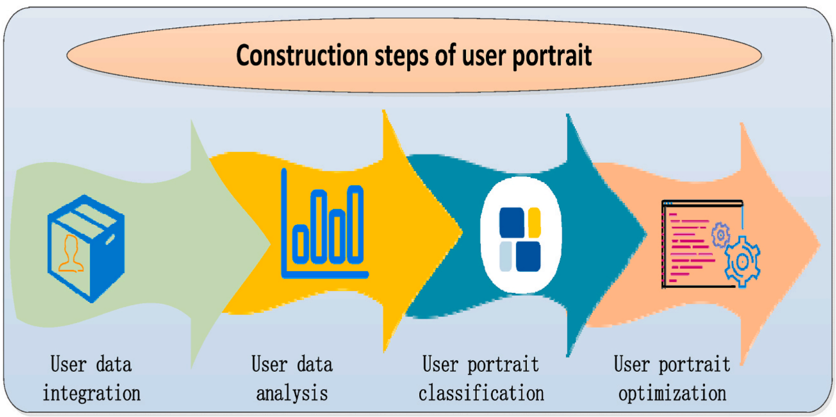

Fig. 2. Construction steps of user profile.

help determine its effectiveness in practical scenarios. If the performance of the model is poor, the accuracy and robustness of the model are improved by adjusting and optimizing algorithm parameters. For example, the vocabulary vector representation in the model can be adjusted; regularization coefficients can be adjusted; the size of the training dataset can be changed. Besides, some special skills, such as data enhancement and transfer learning, can also be used to improve the performance of the model.

3. Computing advertising push strategies

In this study, focusing on advertising push technology in the era of big data, we first selected artificial intelligence-based advertising push algorithms, including collaborative filtering models and content recommendation models, to accommodate different types of advertisements and data characteristics. To improve the accuracy of the models, data preprocessing techniques were employed, including data cleaning and feature extraction. Specifically, the TF-IDF algorithm was used to analyze the keyword weights in textual data and to construct user profiles based on user interests and behavioral data. During the model training phase, cross-validation methods were used to evaluate the models to ensure their stability and accuracy. In particular, k-fold cross-validation was conducted, with multiple experiments performed on training and testing data to optimize model parameters and avoid overfitting issues.

3.1. Targeted advertising technology

Targeted advertising is a technology that improves the accuracy and effectiveness of advertising by selecting suitable target audiences based on user profiles and advertiser needs [36,37]. In targeted advertising placement, by analyzing user interests, behaviors, preferences, geographical location and other information, the user group that best meets the advertiser's goals is selected and advertised. This method has higher efficiency and accuracy compared to traditional casual advertising placement. Targeted advertising usually requires the following steps.

3.1.1. Determining placement area

In the advertising placement plan, determining the placement area is a very important step. Different advertisers may have different regional advertising needs, such as local service advertisements targeting specific regions or internet service advertisements targeting global users. Therefore, selecting an appropriate regional range plays an important role in ensuring advertising effectiveness and saving placement costs.

3.1.2. Determining launch time

Determining the advertising placement time refers to the strategy of selecting the most suitable time period for advertising placement based on factors such as the purpose of advertising promotion, audience characteristics and market competition [38–40]. Through this approach, it is possible to maximize exposure to target users and improve advertising effectiveness.

3.1.3. Choosing a platform and media for advertising

Advertisers should have a detailed understanding of the characteristics and advantages of different advertising platforms and media, such as search engines, social media, video websites, etc [41]. Then the most suitable advertising platform and media are chosen based on factors such as product or service type, target audience, promotional purpose, budget, etc., to achieve the best advertising effect.

3.1.4. Conducting advertising and evaluation

The final step in advertising strategy is to conduct advertising placement and effectiveness evaluation. After the start of advertising placement, advertisers should closely monitor the effectiveness of the advertisement and adjust and optimize the advertising placement strategy in a timely manner.

3.2. Advertising push algorithm based on AI technology

In advertising platforms, user text data should be analyzed, discovering their interests and hobbies. The most important aspect of text processing is to identify keywords that can be used in the text and identify the topic of the text, which mainly includes two parts: text segmentation and text feature extraction.

First, user features are extracted. The topic model is mainly used to mine the hidden topics in the text. The probability of word occurrence in each document is as shown in Equation (1):

$$P(w|d) = \sum_{t} P(w|t) * P(t|d) \tag{1}$$

Among them, $P(w|t)$ is the probability of words in the topic and $P(t|d)$ is the probability of the topic in the document.

Then, text feature extraction is performed using the TF-IDF (term frequency inverse document frequency) method [42,43]. The frequency of a word appearing in a document is higher than that in the overall database. The larger the ratio, the more important it is and the greater the weight.The calculation formula is as shown in (2):

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{2}$$

Among them, $n_{i,j}$ is the number of times that the word appears in the file.The expression for IDF_i is shown in formula _ (3):

$$IDF_i = log\frac{D}{1 + M_{t_i}} \tag{3}$$

Among them, M_{t_i} is the total number of files containing words and D is the number of text files.

The calculated expression for this $TF - IDF_{i,j}$ is shown in Equation (4):

$$TF - IDF_{i,j} = TF_{i,j} * IDF_i \tag{4}$$

Finally, personalized advertising placement is conducted. User similarity is analyzed and calculated to identify similar users. Formula (5) is as follows:

$$sim(u,v) = \frac{\sum_i \frac{1}{ln(1+W_i)}}{\sqrt{N(u) * N(v)}} \tag{5}$$

Among them, W_i is the user group interested in product i. N(u) represents the product set that user u is interested in and N(v) represents the product set that user v is interested in.

Then, products that the target user is interested in but not exposed to within a similar user group is looked for. The probability expression that product I recommends to the user u is shown in Equation (6):

$$P(u,i) = \sum_{u,v,i} W_{uv} * r_{vi} \tag{6}$$

Among them, W_{uv} is the similarity between users u and v. r_{vi} is the level of interest of user v in product i.

Through this method, users' historical behavior data, interests, social networks and other information are analyzed to predict the advertising content that users may be interested in and relevant advertisements are presented to users. It can make advertisements more accurately match users' needs and improve advertising effectiveness and ROI, while also improving user experience and reducing the degree of interference of advertisements on users.

4. Experiment on advertising push based on AI technology

In this experiment, to better reflect the results of the experiment, A/B Scientific control was used for testing. Among them, group A did not use advertising push algorithms; group B used an advertising intelligence push algorithm based on AI technology; these two groups were used for advertising placement testing. By testing the results of two experimental groups, a clear understanding of the actual advertising push performance of the algorithm could be got.

In this study, to thoroughly assess the effectiveness of the advertising push algorithms, a rigorous data analysis process was conducted, including the application of statistical methods and ensuring data reliability. To verify the significance of the experimental results, statistical significance tests such as t-tests and Analysis of Variance (ANOVA) were employed. These methods were used to compare the differences between Group A and Group B in terms of advertising match accuracy, recall rates, and ROI, to ensure that the observed effects were not due to random fluctuations. T-tests on the advertising match accuracy and recall rate data from Groups A and B showed that Group B's advertising match accuracy and recall rate were significantly higher than those of Group A, thereby validating the effectiveness of the AI-based advertising push algorithms.

To ensure data reliability, several measures were taken: During the data collection phase, care was taken to use advertising datasets with sufficient scale and representativeness to avoid data bias affecting the experimental results. The datasets chosen were Avazu, Ali_Display_Ad_Click, and Display Advertising Challenge, each with different characteristics, to provide a comprehensive set of experimental results. During data processing, strict adherence to data cleaning and preprocessing standards was maintained to ensure data accuracy and consistency. To prevent the results from being due to chance, multiple repeated experiments were conducted, and statistical indicators such as mean and standard deviation were used to assess the stability of the experimental results. These steps ensured the reliability of the experimental data and enhanced the credibility of the advertising push algorithm evaluations.

4.1. Selection of advertising datasets

In this article, three advertising data sets are used to evaluate and test the performance of the push algorithm.These three data sets include the Avazu data set, the Ali_Display_Ad_Click data set, and the Display Advertising Challenge data set. Their acquisition strictly complies with relevant laws and regulations and data use agreements to ensure the legality and ethics of data use.The Avazu dataset is a publicly available data source that contains more than 1 billion records and is often used in the study of advertising click-through rate prediction.In order to protect user privacy, the sensitive personal information in the data set has been anonymous, and some characteristic information may also be deleted or replaced.This data set is mainly obtained through the public data set platform.The Ali_Display_Ad_Click data set is published by Taobao and covers 700 million shopping records. It contains a wealth of user and advertising information, which is of great value for studying user behavior and advertising effectiveness.This data set can be obtained through applications within Alibaba Group or cooperative research institutions.As for the Display Advertising Challenge data set, it is part of the advertising click-through rate prediction competition.This data set provides continuous and non-fixed sampling rate data, ensuring a balanced proportion of positive and negative samples, and is very suitable for predictive research on advertising click-through rates.This data set can be obtained through the official website of the competition.

When conducting advertising push testing experiments, the selection of advertising datasets is crucial. The quality and scale of advertising datasets directly affect the performance and effectiveness of push systems. To better evaluate the performance and effectiveness of push algorithms, some advertising datasets should be selected. The following is the advertising dataset selected in this article and its basic information is shown in Table 1.

According to the information in Table 1, this article has selected three advertising datasets for evaluation and testing of push algorithms, namely the Avazu dataset, Ali_ Display_ Ad_ Click dataset and Display Advertising Challenge dataset. These three datasets each have their own focus and can provide some reference for the performance and effectiveness of advertising push algorithms.

4.2. Experimental evaluation indicators

The accuracy of advertisement matching: It is the ratio of the number of advertisements that accurately match a user's interest profile to the number of all advertisements that match that profile.

Recall rate of advertisement matching: It is the ratio of the number of advertisements that accurately match a user's interest profile to the number of advertisements of that type in the advertisement library.

Ad Matching Specificity (Sp): The proportion of actual negative ads that are correctly identified as negative by the model. It is calculated as the number of true negatives divided by the sum of true negatives and false positives.

Ad Matching Sensitivity (Sn): The proportion of actual positive ads that are correctly identified as positive by the model. It is calculated as the number of true positives divided by the sum of true positives and false negatives.

F1 Score: The harmonic mean of precision and recall, used to provide a balanced evaluation of the model's performance.

Matthews Correlation Coefficient (MCC): A measure for assessing the quality of binary classification models, ranging from −1 to 1. A value of 1 indicates perfect prediction, −1 indicates total disagreement, and 0 indicates performance equivalent to random guessing.

Return on Investment (ROI) and Statistical Significance: ROI data indicates advertising effectiveness, while the p-value is used for significance testing, with $p < 0.01$ indicating high statistical significance.

4.3. Experimental results

Firstly, the accuracy of advertising matching between groups A and B was tested and calculated. The detailed data is shown in Fig. 3.

In Fig. 3, the horizontal axis represents the experimental number. A total of 10 experiments were conducted. The vertical axis is the accuracy, which represents the accuracy of advertising matching. After calculation, it can be seen that the average accuracy of advertising matching in Group A was 64.124 %. The average accuracy of advertising matching in Group B was 74.988 %. This indicates that Group B had better advertising matching accuracy, could better meet the needs of advertisers and users and improve the effectiveness and revenue of advertising placement.

Then, the recall rate of advertising matching between groups A and B was tested and calculated and the detailed data is shown in Fig. 4.

In Fig. 4, the horizontal axis represents the experimental number. A total of 10 experiments were conducted. The vertical axis represents the recall rate of advertising matching. After calculation, it can be seen that the average recall rate for advertising matching in Group A was 67.45 %. The average recall rate for advertising matching in Group B was 78.943 %. This indicates that the average recall rate of advertising matching in Group B was significantly higher than that in Group A, indicating that the advertising push algorithm adopted by Group B was more effective and could better match the needs and interests of users.

Table 1
Basic information of the advertising dataset.

Name	Advantage	Disadvantage	Statistical information
Avazu	The data set is large and contains over 1 billion records, representative and covering multiple time periods and regions.	The data is anonymized and some feature information is deleted or replaced, which may affect the model accuracy.	Released in 2014, contains 10 days of data totaling over 1 billion records. Ability to optimize the performance of the advertisement click-through rate prediction model.
Ali_Display_Ad_Click	Users and advertising information is rich, covering a more comprehensive user behavior information, a wide range of applications.	The classification standard of user information is not clear and the user behavior sequence lacks the desensitized advertising identification.	Posted by Taobao, including the basic information of all advertisements, covering a total of 700 million shopping records of all users within 22 days.
Display Advertising Challenge	Data is continuous performance. The sampling rate is not fixed and the proportion of positive and negative samples is not different.	Limited application scope	containing 40 million training samples, 5 million test samples, features including 13 numerical features and 26 category features.

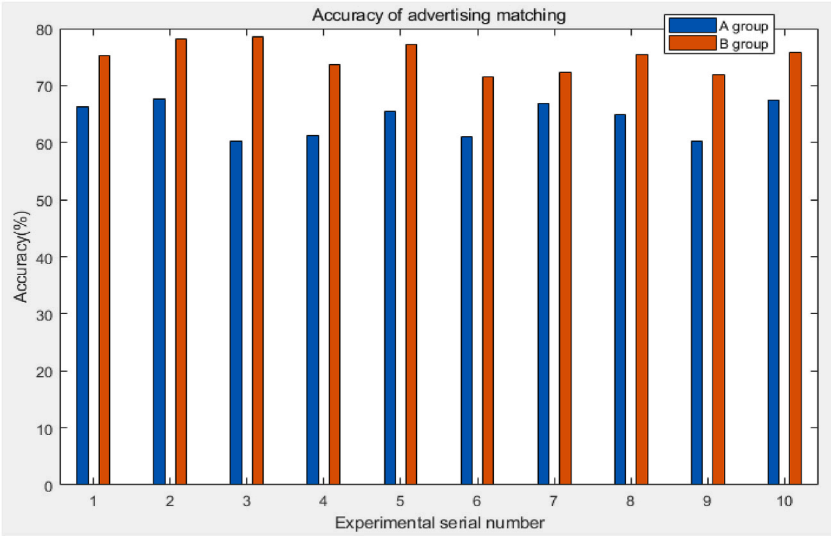

Fig. 3. Accuracy of advertising matching.

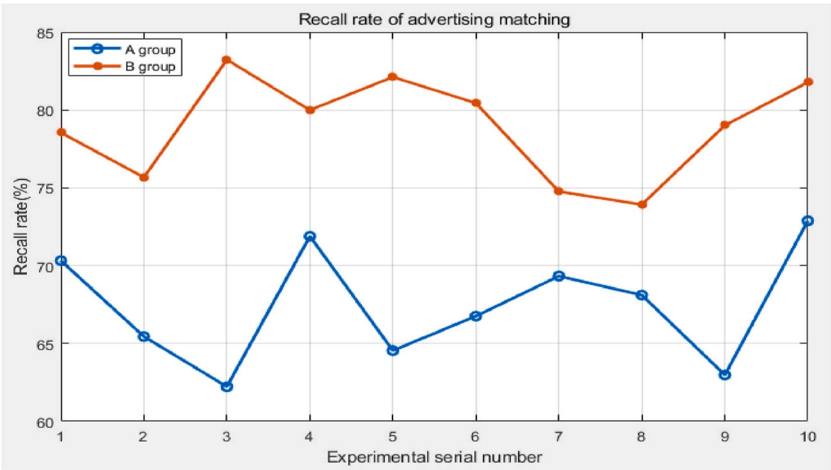

Fig. 4. Recall rate of advertising matching.

Finally, three advertising types, namely tourism, shopping and renting, were selected for placement testing and the ROI of advertising placement in groups A and B was calculated. The detailed data is shown in Fig. 5.

In Fig. 5, the horizontal axis represents the type of advertisement. The vertical axis is the ROI value, which represents the ROI of advertising placement. Black dots in the figure represent the number of experiments. After calculation, it can be concluded that in Group A, the average ROI of advertising for tourism, shopping and renting was 153.03 %, 232.32 % and 192.57 %, respectively. In Group B, the average ROI of advertising for tourism, shopping and renting was 173.96 %, 288.74 % and 216.12 %, respectively. This indicates that using AI based advertising intelligent push algorithms (Group B) has improved the average ROI of advertising placement in three types of advertising: tourism, shopping and renting compared to not using advertising push algorithms (Group A), helping advertisers gain more revenue.

Advertising push tests were conducted in two experimental groups. Group A did not use any advertising push algorithms, while Group B employed an AI-based advertising intelligent push algorithm. The experiment involved three types of advertisements: tourism, shopping, and rental. For each type of advertisement, multiple repeated experiments were conducted to ensure data stability. In each experiment, metrics such as advertising match specificity, sensitivity, F1 score, and Matthews Correlation Coefficient (MCC) were recorded. The experimental data underwent rigorous data cleaning and preprocessing to ensure accuracy. Statistical significance tests, including t-tests and Analysis of Variance (ANOVA), were applied to compare the differences between Group A and Group B on these metrics. The results are shown in Table 2.

In the experiment, multiple repeated tests were conducted on three types of advertisements—tourism, shopping, and rental. Key metrics such as advertising match specificity (Sp), sensitivity (Sn), F1 score, and Matthews Correlation Coefficient (MCC) were

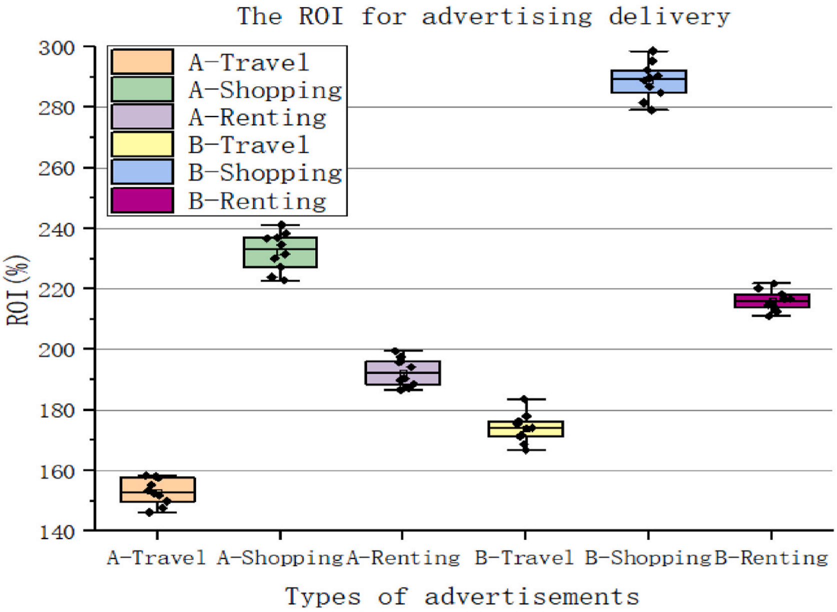

Fig. 5. ROI of advertising placement.

Table 2
Experimental results of the advertising push algorithm.

Advertisement Type	Group	Ad Matching Specificity (Sp)	Ad Matching Sensitivity (Sn)	F1 Score	MCC	Return on Investment (ROI)	Statistical Significance
Travel Ads	Group A	62.35 %	65.20 %	63.75 %	0.56	153.03 %	–
Travel Ads	Group B	73.85 %	77.45 %	75.58 %	0.72	173.96 %	$p < 0.01$
Shopping Ads	Group A	61.80 %	64.10 %	62.92 %	0.55	232.32 %	–
Shopping Ads	Group B	72.90 %	76.60 %	74.68 %	0.70	288.74 %	$p < 0.01$
Rental Ads	Group A	60.45 %	63.90 %	62.14 %	0.53	192.57 %	–
Rental Ads	Group B	71.95 %	75.20 %	73.54 %	0.68	216.12 %	$p < 0.01$
Average Metrics	Group A	Group B	p-value	Group A	Group B	p-value	–
Average Precision (Sp)	61.87 %	72.90 %	$p < 0.01$	–	–	–	–
Average Recall (Sn)	64.07 %	76.42 %	$p < 0.01$	–	–	–	–
Average F1 Score	62.27 %	74.27 %	$p < 0.01$	–	–	–	–
Average MCC	0.55	0.70	$p < 0.01$	–	–	–	–
Average ROI	192.57 %	216.12 %	$p < 0.05$	–	–	–	–

recorded, resulting in detailed data as shown in Table 2. For tourism advertisements, Group B, which utilized the AI-based intelligent push algorithm, showed an approximate 11.5 percentage point increase in specificity compared to Group A, reaching 73.85 %. Additionally, sensitivity improved from 65.20 % to 77.45 %. These improvements were also reflected in the F1 score and MCC, which reached 75.58 % and 0.72, respectively. In terms of Return on Investment (ROI), the average ROI for tourism ads in Group B was 173.96 %, a significant increase from Group A's 153.03 %, with statistical significance ($p < 0.01$). Similarly, results for shopping and rental ads showed a consistent trend, with the use of AI-based intelligent push algorithms leading to improvements in all metrics. The average ROI for shopping ads increased from 232.32 % to 288.74 %, and for rental ads, from 192.57 % to 216.12 %, both showing statistically significant differences ($p < 0.01$). Overall, both in terms of advertising match accuracy and economic return, the AI-based intelligent push algorithm demonstrated clear advantages.

4.4. Discussion

The data from the experiments clearly illustrate the effectiveness of AI-based advertising push algorithms in enhancing ad performance metrics and economic returns. Specifically, the results indicate substantial improvements in advertising match accuracy, recall rates, and return on investment (ROI) when using AI-driven algorithms compared to traditional methods.

Firstly, the average accuracy of ad matching in Group B, which utilized AI algorithms, was 74.988 %, significantly higher than the 64.124 % observed in Group A. This 10.864 percentage point difference highlights the AI algorithms' superior ability to match ads with users' interests, resulting in more relevant advertisements. This improvement is further corroborated by the higher recall rate in Group B, which was 78.943 % compared to 67.45 % in Group A. The higher recall rate indicates that AI algorithms were more effective in identifying and delivering relevant ads, thus enhancing overall advertising efficiency.

The ROI data further underscores the advantages of AI-based algorithms. For tourism ads, Group B achieved an average ROI of 173.96 %, compared to 153.03 % in Group A. Similarly, for shopping ads, the ROI increased from 232.32 % in Group A to 288.74 % in Group B, and for rental ads, it improved from 192.57 % to 216.12 %. These increases in ROI are statistically significant, with p-values less than 0.01, indicating that the observed improvements are not due to random chance. The consistent rise in ROI across all three types of ads—tourism, shopping, and rental—demonstrates that AI-based algorithms can effectively enhance the profitability of various advertising campaigns.

The findings suggest that AI algorithms not only improve the precision and relevance of ad targeting but also provide a significant boost to economic returns. This is indicative of the algorithms' capability to optimize ad placements, leading to higher engagement rates and better alignment with user preferences. Moreover, the enhanced ROI reflects the efficiency of AI algorithms in maximizing advertising investments, ultimately helping advertisers achieve greater profitability.

In practical terms, these results emphasize the importance of integrating AI technologies into advertising strategies. The superior performance metrics and financial returns associated with AI-driven ad targeting provide a compelling case for advertisers to adopt these advanced technologies. By leveraging AI, businesses can achieve more precise targeting, improve ad relevance, and increase overall return on investment, thus gaining a competitive edge in the digital advertising landscape. The study highlights the significant potential of AI-based advertising solutions to transform marketing practices and drive better outcomes for advertisers.

5. Conclusions

In today's rapidly developing information technology, computational advertising has become an indispensable part of the marketing field. With the continuous emergence of emerging technologies such as internet applications, mobile devices and social media, the scenarios and forms of commercial activities are also constantly changing, making the accuracy and effectiveness of computing advertising increasingly crucial. Therefore, AI based computational advertising intelligent computing and push has emerged. In this article, experiments were conducted on advertising push using AI technology's advertising intelligent push algorithm. Through experimental data, it could be concluded that advertising intelligent push algorithms based on AI technology have shown better results and benefits in terms of advertising matching accuracy, recall and ROI. This showed that in the era of Big data, AI based computing advertising intelligence computing and push has become the trend and key technology of industry development. It could better serve advertisers and users through its own intelligent analysis and learning, improving the effectiveness and revenue of advertising placement. In the future, with the continuous development and improvement of technology, computational advertising based on AI is more widely applied and promoted.

The computer and advertising industries are changing at a rate never seen before due to the quick development of big data and artificial intelligence technologies. The accuracy and user experience of advertising have been greatly enhanced by the application of AI, particularly in the areas of intelligent computing and push. Nonetheless, there are still a lot of obstacles and restrictions facing this field's study and use.1. Risk of privacy leakage: Since the AI intelligent recommendation system uses a lot of user data, there is a chance that some of it will leak during processing and transmission, endangering the privacy of the user.2. Algorithm deviation: Due to algorithmic variance or issues with data quality, the intelligent recommendation system may produce biased or erroneous recommendation results. In order to make informed advertising decisions, advertisers will increasingly rely on the gathering, analyzing, and application of large-scale data due to the ongoing growth of data and technological advancements. Advertisers can enhance their advertising performance by optimizing their campaigns and gaining a comprehensive understanding of consumer behavior, preferences, and interests through the use of AI and machine learning algorithms. In the future, personalized and tailored advertising will gain significant traction in the advertising sector. Advertisers will be able to deliver more relevant and eye-catching advertising content based on the user's interests, personality traits, and behavior patterns. Advertisers are able to deliver a highly tailored advertising experience by capturing real-time changes in customer demands and interests through the intelligent recommendation system.

Funding

This work was supported by University of Sanya talent introduction project "Analysis on the Development Strategy of Big Data Advertising Talents in the 5G Era", project No. USYRC23-04.

CRediT authorship contribution statement

Lin Shan: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Waymond Rodgers, Tam Nguyen, Advertising benefits from ethical artificial intelligence algorithmic purchase decision pathways, J. Bus. Ethics 178 (4) (2022) 1043–1061.
[2] Linwan Wu, Understanding Twitter conversations about artificial intelligence in advertising based on natural language processing, Int. J. Advert. 41 (4) (2022) 685–702.
[3] Tran Minh Tung, A systematic analysis of artificial intelligence's usage in online advertising, Mig. Let. 21 (S6) (2024) 892–900.
[4] L. Hu, W. Zhang, S. Du, X. Sun, How to achieve targeted advertising with the e-commerce platform's membership system? Omega 103156 (2024).
[5] Esther Gal-Or, Mordechai Gal-Or, Customized advertising via a common media distributor, Market. Sci. 24 (2) (2005) 241–253.
[6] Marijke De Veirman, Liselot Hudders, Disclosing sponsored Instagram posts: the role of material connection with the brand and message-sidedness when disclosing covert advertising, Int. J. Advert. 39 (1) (2020) 94–130.
[7] Courtney Carpenter Childers, Laura L. Lemon, Mariea G. Hoy, # Sponsored# Ad: agency perspective on influencer marketing campaigns, J. Curr. Issues Res. Advert. 40 (3) (2019) 258–274.
[8] Duen-Ren Liu, Liao Yu-Shan, Chung Ya-Han, Kuan-Yu Chen, Advertisement recommendation based on personal interests and ad push fairness, Kybernetes 48 (8) (2019) 1586–1605.
[9] Li Hu, H.L. Ma, L. Wang, Y. Liu, Hiding or disclosing? Information discrimination in member-only discounts, Transport. Res. E Logist. Transport. Rev. 171 (2023) 103026.
[10] Yuan-mei Wang, W.A.N.G. Ya-jun, Z.H.A.O. Shi-bo, Personalized explainable recommendation based on BERT, J. Artif. Intell. 5 (1) (2023) 24–38.
[11] Sandy Zhu, A discussion on how to apply smart data for precision marketing in the global economy: asian context, Economit Journal: Scientific Journal of Accountancy, Management and Finance 1 (2) (2021) 70–84.
[12] Qi Xin, The implementation of an AI-driven advertising push system based on a NLP algorithm, Int. J. Comput. Sci. Inf. Technol. 1 (1) (2023) 30–37.
[13] Amir Zaib Abbasi, Advertising value of vlogs on destination visit intention: the mediating role of place attachment among Pakistani tourists, Journal of Hospitality and Tourism Technology 13 (5) (2022) 816–834.
[14] Sun Joo Ahn, Jooyoung Kim, Jaemin Kim, The bifold triadic relationships framework: a theoretical primer for advertising research in the metaverse, J. Advert. 51 (5) (2022) 592–607.
[15] Lichun Zhou, Product advertising recommendation in e-commerce based on deep learning and distributed expression, Electron. Commer. Res. 20 (2) (2020) 321–342.
[16] Liang Wang, Yu Zhiwen, Bin Guo, Dingqi Yang, Lianbo Ma, Zhidan Liu, et al., Data-driven targeted advertising recommendation system for outdoor billboard, ACM Transactions on Intelligent Systems and Technology (TIST) 13 (2) (2022) 1–23.
[17] Federico Germani, Nikola Biller-Andorno, The anti-vaccination infodemic on social media: a behavioral analysis, PLoS One 16 (3) (2021) e0247642.
[18] Hui Li, Q. Chen, Z. Zhong, R. Gong, G. Han, E-word of mouth sentiment analysis for user behavior studies, Inf. Process. Manag. 59 (1) (2022) 102784.
[19] Lu Zhang, J. Shen, J. Zhang, J. Xu, Z. Li, Y. Yao, et al., Multimodal marketing intent analysis for effective targeted advertising, IEEE Trans. Multimed. 24 (2021) 1830–1843.
[20] Y.J. Zhang, Z. Dong, X.W. Meng, Research on personalized advertising recommendation systems and their applications, Chin. J. Comput. 44 (3) (2021) 531–561.
[21] Zi-yan Jin, Juan Zhang, L.I. Xiang-jun, A collaborative filtering advertising recommendation algorithm with tag, Comput. Eng. 44 (4) (2018) 236–242.
[22] Zhaoxiang Cai, Machine learning for multi-omics data integration in cancer, iScience 25 (2022) 2.
[23] Malte D. Luecken, Benchmarking atlas-level data integration in single-cell genomics, Nat. Methods 19 (1) (2022) 41–50.
[24] Zhabiz Gharibshah, Xingquan Zhu, User response prediction in online advertising, ACM Comput. Surv. 54 (3) (2021) 1–43.
[25] Jannick Kirk Sorensen, Hilde Van den Bulck, Public service media online, advertising and the third-party user data business: a trade versus trust dilemma? Convergence 26 (2) (2020) 421–447.
[26] E.S. Dalmaijer, C.L. Nord, D.E. Astle, Statistical power for cluster analysis, BMC Bioinf. 23 (1) (2022) 205.
[27] M. Benassi, S. Garofalo, F. Ambrosini, R.P. Sant'Angelo, R. Raggini, G. De Paoli, G. Piraccini, Using two-step cluster analysis and latent class cluster analysis to classify the cognitive heterogeneity of cross-diagnostic psychiatric inpatients, Front. Psychol. 11 (2020) 1085.
[28] S. Ghosal, R. Bhattacharyya, M. Majumder, Impact of complete lockdown on total infection and death rates: a hierarchical cluster analysis, Diabetes Metabol. Syndr.: Clin. Res. Rev. 14 (4) (2020) 707–711.
[29] Yulu Liu, Xuesong Dai, Huan Zhang, Zhanxin Gu, Xi Xihang, Wifi-based advertising push method, World Scientific Research Journal 5 (7) (2019) 66–69.
[30] Liming Wang, Yanni Liu, Jicheng Wu, Research on financial advertisement personalised recommendation method based on customer segmentation, Int. J. Wireless Mobile Comput. 14 (1) (2018) 97–101.
[31] Katherine White, Rishad Habib, David J. Hardisty, How to SHIFT consumer behaviors to be more sustainable: a literature review and guiding framework, J. Market. 83 (3) (2019) 22–49.
[32] Minou Weijs-Perree, J. Van De Koevering, R. Appel-Meulenbroek, T. Arentze, Analysing user preferences for co-working space characteristics, Build. Res. Inf. 47 (5) (2019) 534–548.
[33] Vladislav Ilin, The Role of User Preferences in Engagement with Online Learning, E-Learning and Digital Media 19.2, 2022, pp. 189–208.
[34] Lianbo Ma, H. Kang, G. Yu, Q. Li, Q. He, Single-domain generalized predictor for neural architecture search system, IEEE Trans. Comput. (2024).
[35] Jiawei Chen, H. Dong, X. Wang, F. Feng, M. Wang, X. He, Bias and debias in recommender system: a survey and future directions, ACM Trans. Inf. Syst. 41 (3) (2023) 1–39.
[36] Daniel Yang, Saman Grice, Research on the design of E-commerce recommendation system, Int. J. Smart Bus. Technol 6 (1) (2018) 15–30.
[37] Y.X. Hou, G.C. Yang, X.M. Fan, Research on vehicle network advertising delivery based on friends relationship, Open J. Transport. Technol. 7 (5) (2018) 359–369.
[38] Jūra Liaukonytė, Alminas Žaldokas, Background noise? TV advertising affects real-time investor behavior, Manag. Sci. 68 (4) (2022) 2465–2484.
[39] Mark G. Maffett, Jorge O. Brusa, THE effect of online advertising on consumer buying interest in online selling applications with customer satisfaction as an intervening variable (case study of medical doctor and nurses in Chicago), MEDALION JOURNAL: Medical Research, Nursing, Health and Midwife Participation 3 (4) (2022) 136–143.
[40] Rumen Pozharliev, Matteo De Angelis, Dario Rossi, The effect of augmented reality versus traditional advertising: a comparison between neurophysiological and self-reported measures, Market. Lett. 33 (1) (2022) 113–128.

[41] Lianbo Ma, N. Li, G. Yu, X. Geng, S. Cheng, X. Wang, et al., Pareto-wise ranking classifier for multi-objective evolutionary neural architecture search, IEEE Trans. Evol. Comput. (2023).
[42] Denis Eka Cahyani, Irene Patasik, Performance comparison of tf-idf and word2vec models for emotion text classification, Bulletin of Electrical Engineering and Informatics 10 (5) (2021) 2780–2788.
[43] Shahzad Qaiser, Ramsha Ali, Text mining: use of TF-IDF to examine the relevance of words to documents, Int. J. Comput. Appl. 181 (1) (2018) 25–29.

Lin Shan was born in Harbin, Heilongjiang, P.R. China, in 1989. She received a master's degree from East China Normal University. Now, she works in the School of Media and Cultural Industry of Sanya University. Her research interests include computing advertising, big data advertising and digital marketing. E-mail: linshan@sanyau.edu.cn