communication
computing control

CCC Publications

AGORA
UNIVERSITY PRESS

Analysis of Advertising Promotion Strategy Based on Improved Collaborative Filtering Algorithm under Digital Media Technology

Q. Wu

Q. Wu
College of Humanities and Law
Shanghai Business School,Shanghai 200235, China
Corresponding author: wuqin@sbs.edu.cn

Abstract

With the advent of the Internet, individuals have access to an ever-increasing array of external information sources, and as such, advertising information has undergone a significant transformation. In this study, we present a highly accurate, high-throughput, and high-yield ad recommendation system that is capable of precisely targeting users for personalized ad recommendations. Our approach involves enhancing the user interest preference model, introducing ad keywords as labels into the similarity calculation of Query pages, and employing a weighted integrated similarity measure of Query pages to mitigate sparsity within the similarity matrix. Furthermore, we optimize the system process of ad promotion and devise an ad recommendation strategy based on multiple parallel recall and uniform sorting distribution. Our model training results indicate that the proposed algorithm elevates Precision, Recall, and F1 by 27%, 25%, and 28%, respectively, compared to the traditional model. Additionally, our system test results demonstrate enhanced scalability, with approximately four times higher concurrent performance, as well as improvements in high expansion, low latency, and strong stability. Such results hold particular significance in guiding the design of actual ad recommendation system projects.

Keywords: collaborative filtering algorithm; advertising promotion; user preference; multiplex parallelism.

1 Introduction

During the era preceding the widespread adoption of the Internet, advertisements were primarily disseminated through traditional news media channels, such as television, roadside billboards, newspapers, and radio. These conventional advertising display methods necessitated significant human resources for promotion, came with high investment costs and slow returns, and were challenging to manage effectively [1]. However, with the development of mobile technology, Internet advertisements have become increasingly prevalent across major applications. In comparison to traditional communication media, online advertising boasts a vast coverage area, strong targeting capabilities, and an extensive audience base. The Internet serves as a novel and efficient means of connecting consumers with enterprises, rendering online advertising both flexible and interactive [2]. By leveraging the Internet, enterprises can instantaneously collect consumer feedback data, such as search query records,

ad click records, and problem feedback, among other types of information. The amassed data can be quantitatively analyzed on a collected data platform, enabling the identification of issues and the proposal of improvement plans. Such plans may include refinement of placement targets, alteration of placement methods, and adjustment of budget allocation strategies.

For advertisers, failure to identify the appropriate audience group for ad placement can result in information waste and increased costs. Improper recommendations can even result in users developing negative opinions about the website, thereby seriously impacting advertising and website operation effects [3, 4]. Users, in the age of information overload, can easily waste time when they receive excessive amounts of irrelevant information. Thus, there exists practical demand for accurate ad recommendation both for ad deliverers and users. The recommendation system based on user profile information is a feasible practice that can be used to achieve accurate ad push and address market demand [5]. Early research on implicit factor models of ad promotion algorithms focused on users' explicit feedback. This refers to users expressing their personalized preferences by rating ads, such as Douban movie ratings, Taobao product ratings, and so on. Given explicit user feedback, the recommendation problem is generally formalized as a rating prediction problem. That is, the user and ad embedding vectors in the model are iteratively optimized by minimizing the difference between predicted and true ratings. However, in practical Internet application scenarios, obtaining explicit user rating feedback is often difficult. Therefore, more and more research works are focusing on how to mine and use implicit feedback information from users' browsing of products, clicking on links, and other user history behaviors to model their personalized preferences [6, 7].

The cold-start issue within ad recommendation primarily arises due to new users or ads, where the former has not visited any ads, and the latter have not been viewed by any users. This predicament poses a significant challenge in the recommendation system. For ads' cold start issue, content attributes such as the ad topic, topic category, and keywords are predominantly explored. For instance, ads can be converted into keyword vectors, and by evaluating the vectors' similarity, ads can be clustered and suggested to users interested in similar ads. Additionally, deep learning techniques are employed to circumvent the neural network's reliance on comparable information by incorporating interaction or rating information of some advertisements to make it robust towards new items. In addressing the system's cold start issue, expert labeling is introduced via expert knowledge [8]. Given the numerous user interest possibilities, it is vital to provide users with representative, diversified, and distinct ads to ensure high coverage, which can expedite the collection of user feedback and construct user portraits. Randomly recommending ads with different topics or contents to users and subsequently evaluating their feedback can aid in achieving this goal. The foundation of information flow recommendation hinges on traffic, and the system's rapid traffic growth poses the problem of system throughput. Reduced system throughput can diminish the number of ads placed, which, in turn, can affect advertising revenue. With an increasing number of users, the disparity in experience between users can be substantial. Presently, many internet companies' information flow recommendation systems are deficient in accuracy and do not cater to personalized recommendations for users [9]. In this study, through the design of a personalized advertising recommendation system, advertisements changed from the original large-scale advertising to personalized recommendation, which mainly weakened users' objectionable attitude towards advertisements and increased advertisers' economic benefits so that advertising was more targeted. Therefore, this paper's main contributions are as follows: (1) Enhanced the user interest preference model by incorporating advertising keywords as labels in the similarity computation of query pages, which reduced the sparsity of the similarity matrix by employing a Query page weighted integrated similarity measure. (2) Leveraging a multi-parallel recall algorithm, this approach primarily employs multi-threading technology to fully utilize the server's multi-core performance in parallel selection and recommendation of more material pairs under limited time consumption.

2 Related works

Clustering algorithms bear a resemblance to user-based or item-based collaborative filtering since clustering of users or items is executed based on the distance metric values [10]. In user-based cluster-

ing, users can be segregated into various target groups based on metric values, and items with higher scores in the same target group can be recommended to the target user, while item-based clustering is akin to item-based collaborative filtering [11]. Currently, collaborative filtering using logistic regression is another advertising recommendation algorithm with a broad audience. The core concept of this algorithm is to transform the recommendation problem into a dichotomous problem, and classification recommendation algorithms such as logistic regression and plain Bayes deliver superior performance in terms of interpretation. Consequently, their usage is more feasible, and their final outcome sorts the data based on probability values by employing feature engineering, thereby achieving tuning effects [12]. The regression algorithm aligns broadly with the classification algorithm and can efficiently handle continuous rating data to predict ratings for users and items alike.

With the advent of machine learning and deep learning, the use of model-based advertising recommendation algorithms has become more prevalent. In today's industrial environment, there are a plethora of recommendation models available, including traditional machine learning models [13] such as LR, FM, FFM, GBDT+LR, and more advanced deep learning models [14] such as DNN, DSSM and its variants like Wide and Deep, Deep FM, DIN, etc. Among these models, logistic regression, an important recommendation base model, has progressed significantly with deep learning. It can be transformed into a CTR (Click Through Rate) problem in recommendation systems, where it is used to predict the probability of a user clicking on an item. This transforms the recommendation problem into a classification problem, which is ultimately predicted by the model to get a continuous value that ranges between 0 and 1. The continuous value is sorted according to the likelihood represented by it, or a threshold value is set to determine whether to click or not [15]. Liu et al. [16] proposed a combined model of GBDT+LR based on logistic regression, which is also widely used in industrial settings. The model's core idea is to automatically filter features and combinations, eventually generating a discrete feature vector. This vector is then fed into the LR model for CTR prediction. In the early advertising recommendation system, the recommendation technology is too simple, which mainly realizes the collaborative filtering technology of recommendation by comparing the user similarity. At the same time, there are some problems in the application of these technologies, such as cold startup, sparse user feedback matrix and high dimension. Therefore, it is not enough to meet the differences between users' needs and preferences.

Multi-Armed Bandit (MAB) is a practical and widely used algorithm in reinforcement learning that seeks to maximize the cumulative reward for the whole process by finding the arm with the largest reward expectation [17, 18]. The exploration-exploitation (E&E) problem is central to MAB algorithms, where exploration involves discovering new and unknown information to improve the recommendation model, and exploitation entails using known commodity information to select an optimal choice under current conditions [19]. Contextual-Bandit is a variant of MAB that considers the user's product characteristics and consists mainly of the LinUCB algorithm and its improved versions. Among these, LinUCB is the most popular algorithm in current dynamic news recommendation. Compared to the UCB algorithm, LinUCB accounts for the user's characteristics and incorporates them into the mean estimation and variance calculation [20]. The LinUCB algorithm assumes a generalized linear relationship between the user's feedback score and their characteristics [22 **?**]. However, this generalized linear model may not fit all models, and thus, the relationship between user features and predicted feedback has become a direction of improvement for the LinUCB algorithm. Ciucanu et al. [23] used the concept of implicit features to represent additional information about users using unobserved implicit information. They combined this information with LinUCB's original linear feature processing to improve user feedback rate. Erginbas et al. [24] used an association matrix to represent user similarity information between users and constructed a linear relationship between users and feedback based on user similarity to further improve the prediction accuracy.

In the prediction process of neural network method, the parameter scale and characteristic dimension of the model will be very large, and the training time will increase sharply. Therefore, we can try to deal with this phenomenon by solving the problem of parameter miscellaneous.

Using attention mechanism to predict CTR can achieve good results compared with the traditional divine Scripture network method. However, at present, there are few researches on the application of attention mechanism to CTR prediction, which can be carried out in this field and try to apply

attention mechanism to recommendation systems in various fields to improve recommendation quality. It should also be noted that in the deep learning method, data noise will affect the model's acquisition of feature information, and the parameter learning process is subject to various restrictions. When the training data is interfered by uncertainty, the CTR prediction effect will be greatly affected, that is, the robustness is poor. Therefore, CTR prediction in noisy environment still needs further study.

3   Advertising promotion system based on user personalized recommendation

3.1   Overall design

The business request process for the ad recommendation system based on personalized user data can be roughly outlined as follows: initially, the user's request activates the online ad traffic gateway, which then accepts and parses the request, while simultaneously acquiring the user's corresponding profile and the experimental information associated with the traffic. Utilizing this data, the recommendation system performs multiple parallel operations including recall, sorting, rule filtering, among others, to ultimately return the most relevant ad content to the user. The recall process is essential to ensure that high-quality content participates in the sorting and bidding process. As users interact with the recommended ads, their feedback is collected in the log system, which serves as a fundamental source of data to assess the system's accuracy.

The recommendation system implemented in this design is founded on parallel multiplex recall and uniform sorting distribution. Specifically, multiplexed parallel recall incorporates content recall, collaborative recall, model recall, and tag recall, concurrently executed to widen the range of eligible content for the campaign. Coarse ranking employs models to identify and eliminate poor quality candidates based on ad candidate recall. Fine ranking, on the other hand, seeks to select the most captivating and engaging ad content with the highest probability of interaction with users. The rules applied in the system encompass those related to the advertising business, anti-fraud measures, anti-cheating protocols for traffic, among others. Advertisers, also referred to as ad producers, submit their ad content to the platform for review and approval, after which it enters the candidate database. This database is responsible for storing, calculating, and analyzing the ad material entered by ad producers, aiding in recall by the ad engine. The recommendation effect is ultimately improved through the application of the effect evaluation system, log collection system, and model service system. The present paper introduces an ad recommendation system based on personalized user data, with Figure 1 below demonstrating the specific request link and the interaction flow between the various systems.

3.2   User Interest Preference Model

User-based collaborative filtering recommendation is mainly to study the relationship between users [25], using the nearest neighbor technique, by analyzing the interest preference information of target users to recommend items that match their interest preferences, the specific process is shown in Figure 2.

The CTR is a reflection of a user's input request preference (i.e., the output Query page) for the search engine's selection of displayed ads. The ad keyword factor, which not only impacts ad display, but also plays a vital role in relevance calculations affecting the number of clicks, must be taken into account [26, 27, 28, 29]. To achieve this, we propose augmenting the two-dimensional interest preference model of Query pages and advertisements with the ad keyword factor, creating a comprehensive three-dimensional interest preference model that accurately describes Query page preference levels. The Q-K-A interest model based on ad keywords must consider three factors: the information of ad keywords matching the Query page, the information of ads shown by the Query page, and the relationship between the ad keywords matching the Query page and the chosen ads displayed. To facilitate this, we introduce the following definition.

Assuming IS $\equiv$ (Q,K,A,C) is an information system, where. $Q \equiv \{q_1, q_2, ..., q_m\}$ is the set of Query pages. m is the total number of Query pages. $K \equiv \{k_1, k_2, ..., k_n\}$ is the set of ad keywords. n is the total number of ad keywords. $A \equiv \{a_1, a_2, ..., a_r\}$ is the set of ads. r is the total number of ads.
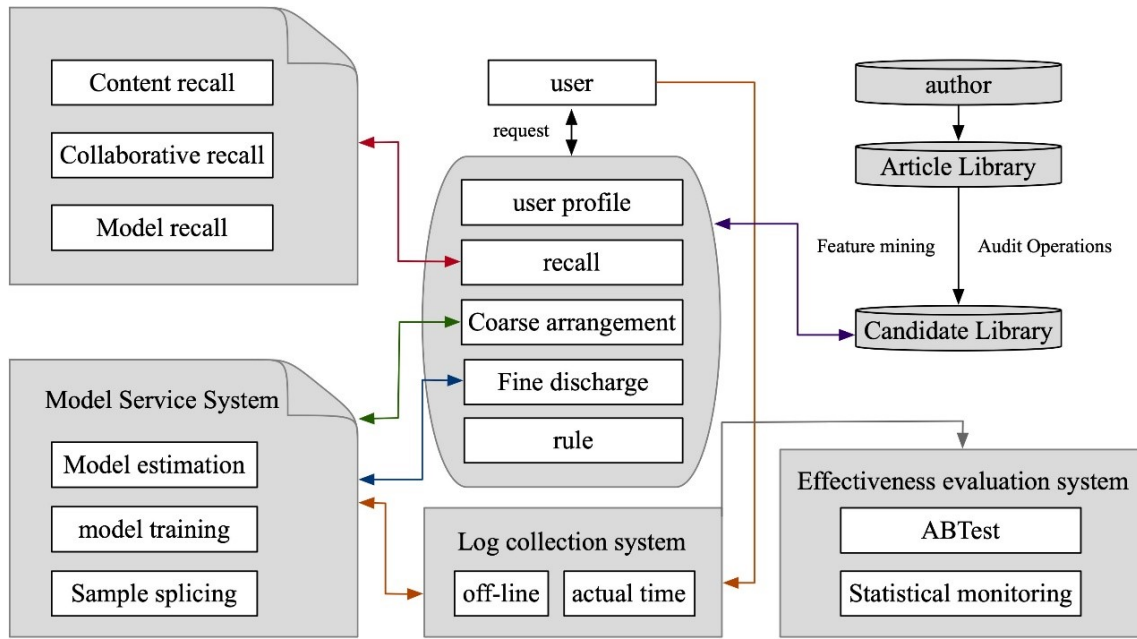
Figure 1: Advertising promotion system architecture

$C \equiv \{c_{(q,a)} \| q \in Q, a \in A\}, c_{(q,a)}$ denotes the number of elements in Q element in q on which the A elements in a generated by the click-through rate.

Let $T = t_1, t_2, ..., t_s$ be the set of ad tags, where s is the total number of ad tags. For any $t\_i \in T.k\_j \in K(1 is, 1 jr)$, there is $t_i = k_j$, when and only when i=k .

Given an information system IS=(Q,K,A,C), the non-empty finite set represents the set of all advertising clicks, for $x \in D$, with $x = <click, impression.a, q, k>$. for any $a_i \in A$. $q_j \in Q(1 \le i \le r, 1 \le j \le m), x \in D, if x.a = a_i$, Then $q = q_j$. displays the click rate $c_{(q_j, a_i)}$ generated by a_i on $q_j.c_{(q_j, a_i)}$ is defined as follows:

$$c_{q_j, a_i} = \frac{\sum_{x \in D} x.click}{\sum_{x \in D} x.impression} \qquad (1)$$

Where $\sum_{x \in D} x.click$ represents that the q_j display a_i and the total number of clicks generated. $\sum_{x \in D} x.impression$ indicates the total number of clicks generated on q_j show on a_i shows the total number of impressions generated. Equation 1 shows that. $c_{(q_j, a_i)}$ is determined by the number of clicks and impressions it corresponds to.

The key aspect of the user-based collaborative filtering algorithm is to find users with similar preferences to the target users. In this paper, we use a modified cosine similarity measure, which not only keeps the correlation degree similar in value, but also makes all the rating curves smooth. Let $I_x y$ be the combination of user and co-rated item vectors. I_x and I_y denote the vector of rated items of user and user respectively, then the similarity between user and user can be expressed as follows.

$$sim(R_{x,i}, R_{y,i}) = \frac{\sum_{i \in I_{xy}} (R_{x,i} - \overline{R_x}) \cdot (R_{y,i} - \overline{R_y})}{\sqrt{\sum_{i \in I_x} (R_{x,i} - \overline{R_x})^2} \sqrt{(R_{y,i} - \overline{R_y})^2}} \qquad (2)$$

In the ad tag-based search ad recommendation system, the similarity of Query pages mainly depends on three factors: the co-hit similarity between Query pages, the co-label similarity and the co-contained relationship similarity.

To represent user rating differences in Query pages accurately, this paper utilizes the modified cosine similarity measure to calculate the three aforementioned factors, namely, the proportion of ad keywords matching between two Query pages and the proportion of ad keywords matching between pages and ad-related relationships. Additionally, a comprehensive weighting method is used to calculate the similarity between Query pages in the recommendation algorithm to reduce sparsity in the similarity matrix calculation. Furthermore, the Top-N strategy is employed to decrease the size of the
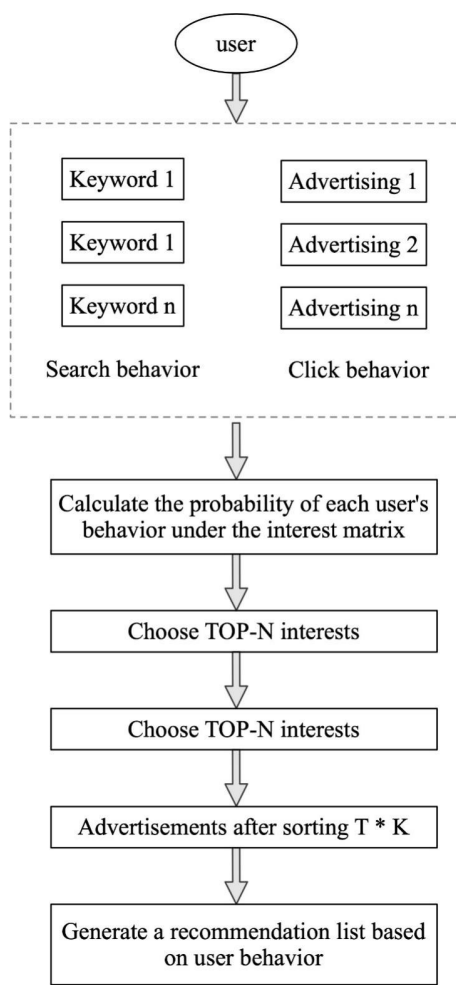
Figure 2: Ad recommendation process based on user behavior preference

Query pages. Specifically, the size of the nearest neighbor candidate set, K, is also reduced using the Top-N strategy.

3.3 Multiplex Recall

he ad recall process is responsible for selecting ad candidates with the objective of recalling as many relevant materials as possible while consuming limited time. Its aim is to choose the top-level ads from a pool of hundreds of thousands of ads to present to the user, and then, after the final business strategy filtering, return 1-3 ads with the strongest relevance. Ad recall content can be user-based or ad-based. User-based recall includes hashtag-based recall, user social network-based recall, and target group pack-based recall.

Traditional recommendation systems typically use multiple serial recalls and prioritize the recall link strategy, which can result in long recall times, large computational effort, and few candidate materials recalled by the system. To address these issues, this system is designed based on a multiple parallel recall algorithm that utilizes multi-threading technology, making full use of server multi-core performance to obtain more pairs of materials for selection and recommendation in parallel with limited time consumption. The unified sorting and distribution incorporates the mandatory intervention of product traffic policy, which enables the system to disable the distribution of ad content placed by a certain advertiser in real time and suppress certain ads during the sorting and distribution. The detailed pseudo-code of the algorithm is provided below. User traffic characteristics contain user profiles, basic data, traffic base devices, time, and other information, while ad primarily consists of unique ad identifiers.

The algorithm then performs ad candidate recall using multiple parallel recall modules, such as content recall, collaborative recall, model recall, and tag recall, in order to obtain a large number of

---

Algorithm 1 Parallel Recall

Input: UQ:User traffic characteristics
Output: AD:The set of ads for this user

1: Parse user traffic characteristics UQ Obtain user unique identifiers UID and ad space characteristics AP
2: Initialize variables $AD \leftarrow \emptyset$
3: Completing the user profile according to the unique user identification UID UP
4: Generate new features for users $UF \leftarrow AP \cup UP$
5: Get the number of recall modules N and timeout settings T
6: **for** i, N **do**
7:     Pass UF in parallel with concurrent threads to get ad candidates, and the acquired candidates will be passed to AD
8: **end for**
9: **for** T No timeout **do**
10:     Waiting
11: **end for**
12: **for** all $A \in AD$ **do**
13:     **if** A need to filter **then**
14:         AD $\leftarrow$ AD - A
15:     **end if**
16: **end for**
17: **if** $AD \neq \emptyset$ **then**
18:     return AD
19: **end if**

---

candidate ads within the given time limit. These recall modules are designed to operate in parallel, utilizing multi-threading technology to maximize the utilization of server multi-core performance, and generate more candidate ads for recommendation. The algorithm then uses coarse ranking to eliminate poor quality candidates and selects high-quality ads based on the ad candidate recall results. Finally, fine ranking is applied to select the most interesting and likely-to-interact ads for the user, based on the user's personalized data and preferences. The algorithm also incorporates various business rules, such as anti-fraud and anti-cheating rules, during the sorting and distribution process to ensure ad content quality and user satisfaction. In summary, the algorithm 1 is a comprehensive ad recommendation system that utilizes parallel multiplex recall, uniform sorting distribution, and various user traffic characteristics to effectively select and recommend relevant and high-quality ads to users. The system also considers business rules and user feedback to continuously improve the recommendation effect.

4 Experiment and analysis

4.1 Experimental setup

The training set comprised of 79,637 positive samples and 21,925,769 negative samples, yielding a positive to negative ratio of 1:275. The positive samples denote users who clicked on an advertisement, while the negative samples indicate those who did not. The sample set consisted of 84,327 positive samples and 22,467,358 negative samples, resulting in a positive to negative ratio of 1:266. Another sample set, unlabelled, was solely utilized for actual prediction detection, comprising of 20,525,716 data pieces. The samples underwent processing through merging and de-weighting with A, whereby all positive samples were selected, and negative samples were sampled at a 1:10 ratio. Subsequently, the results were mixed, shuffled, and divided into nine equal parts. Of these, 8/9 constituted the new training set (about 1.6 million), while 1/9 constituted the new validation set (about 200,000).

For the analysis, we set the number of interest mining to 50 and the number of iterations to 200, while the hyperparameters α and β were set to 0.5 and 0.01, respectively. By examining the identity document (ID) of the keywords corresponding to different interests, we can deduce that each

keyword holds varying probabilities in different interests. By calculating the corresponding conditional probabilities of keywords k545 and k670, for instance, we can determine their interest preferences and generate recommendation lists based on these preferences.

4.2   Analysis of results

In this paper, the proposed algorithm is compared with User-based-CF, Tag-based-CF and Tag-item-based-CF in terms of Precision, Recall and F1, and the experimental results are shown in Figure 3.
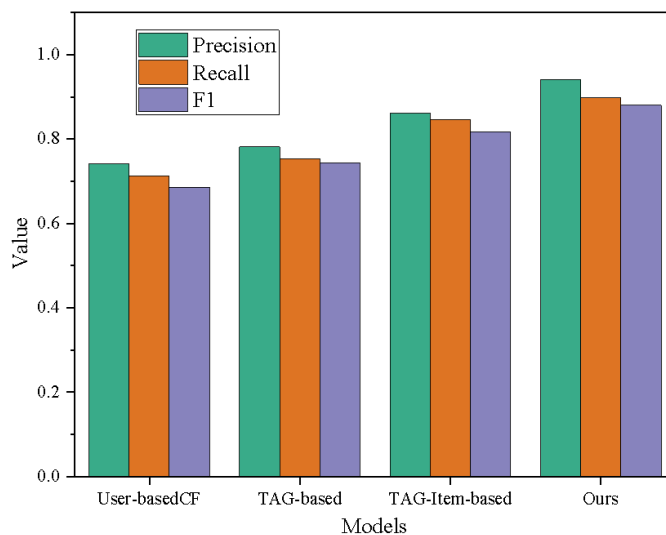

Figure 3: Top10 recommendations for different models

Upon comparison, it was discovered that the collaborative filtering ad recommendation algorithm proposed in this paper, which incorporates labels, exhibited a 27% improvement in accuracy, a 25% improvement in recall, and a 28% increase in F-measure as compared to the traditional collaborative filtering algorithm. This improvement can be attributed to the consideration of CTR, ad keywords, and the relationship between ad keywords and ads when calculating the similarity between Query pages, allowing for the integrated similarity calculation method to effectively capture the preference information of Query pages for ads, the relevance of ad keywords to Query pages and ads, and the characteristics of the ads themselves, thereby providing a more comprehensive description.

Moreover, the model training results of LightGBM, XGBoost, and Wide&Deep were compared, as illustrated in Figure 4.

The experimental results indicate that LightGBM has slightly lower F1 and Recall scores compared to XGBoost. This is because LightGBM uses the histogram algorithm, which provides faster computing speed and lower memory usage at the expense of sacrificing some accuracy. In contrast, XGBoost incorporates a regularization term in the objective function to prevent overfitting and uses both first and second-order derivatives, resulting in more accurate and customizable losses. Meanwhile, the Wide&Deep model combines the memory capabilities of its generalized linear model with the generalization capabilities of its deep feedforward neural network model. It utilizes the generalized linear model to learn feature relevance from the underlying data of ad users and the deep feedforward neural network to uncover the interaction between ad interactions among implicit features of users, resulting in a high accuracy rate of recommendations. In this experiment, our model constructs the interest model of Query pages by analyzing ad keywords, ads, and their relationship, and calculates the integrated overall similarity. This approach correctly expresses the interest model of Query pages in the search ad recommendation system relative to the method that uses CTR as preference information and ad keywords as implicit preference information, ensuring the accuracy of calculating the neighborhood and improving the accuracy of the recommendation system.

Additionally, attention is paid to the overfitting of the two algorithms on the validation set as the number of rounds increases. We observed the change in AUC values on the validation set, and the
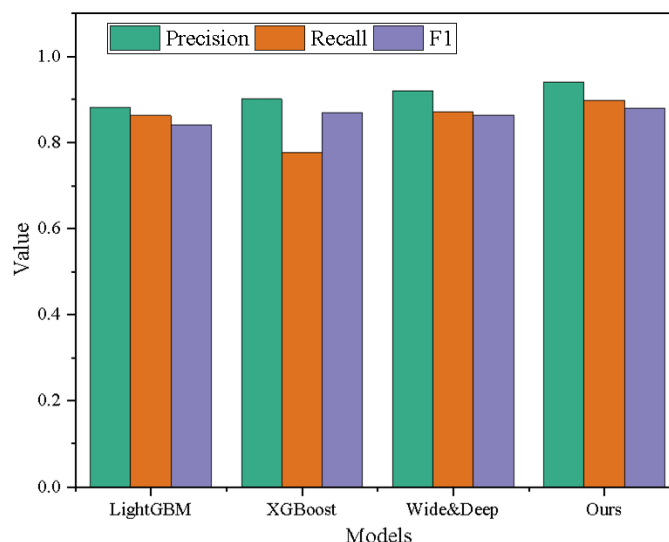
Figure 4: Comparison of model training results
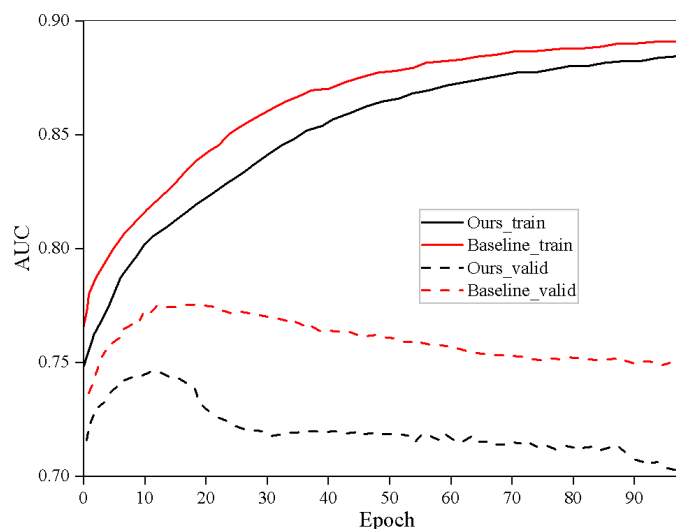
results are displayed in Figure 5.

Figure 5: Training process of personalized model

As depicted in Figure 5, the Area Under the Curve (AUC) values of both algorithms exhibit an increasing trend on the training set as the models undergo continuous training. However, on the validation set, our proposed model exhibits superior predictive power at the 13th iteration, followed by a rapid decline in AUC values, which then asymptotes after 30 iterations. Similar behavior is observed in the benchmark algorithm, wherein the AUC values gradually diminish as the number of training iterations increases. Nonetheless, the decline rate in the benchmark algorithm is much lower than that of our model, which underscores the efficacy of incorporating feature vectors that encapsulate interaction information into the neural network layers, as opposed to relying on the original features. This feature engineering technique alleviates the overfitting problem by simplifying the mining of interaction information and lowering the training complexity.

To assess the scalability performance of our proposed algorithm, we randomly select and execute 20%, 40%, 60%, and 80% of the data set size and compare their execution times against the overall data set. The experimental outcomes are illustrated in Figure 6.

As evident from Figure 6, the algorithm's execution time exhibits a slow and gradual increment, followed by a sharp increase, and then a gradual and smooth increase with the escalation of data size. The growth of execution time of the collaborative filtering and recommendation algorithm equipped
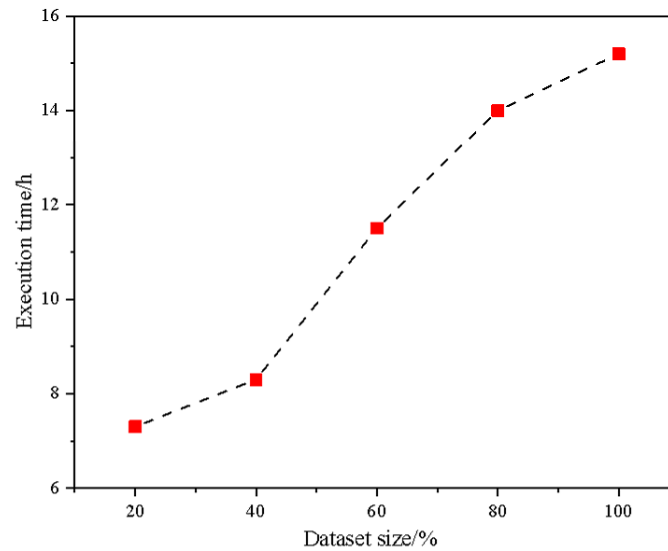
Figure 6: Scalability test results of Multiple recall System

Table 1: concurrency tests of Multiple recall System

	QPS/s	RT/ms	CPU/ %
Parallel recall	2100	5.3	62.0
Serial recall	520	23.4	24.6

with labels remains within an acceptable range as the data size increases, thus indicating its commendable scalability.

Drawing on the above experimental data, we perform a statistical analysis of the system's performance by comparing the parallel recall sorted distribution-based approach with the conventional serial sorted distribution-based system, based on Query Rate Per Second (QPS), Response Time (RT), and CPU usage. The evaluation results of the experimental data are presented in Table 1.

The experimental results indicate that the user-based personalized data advertisement recommendation system proposed in this thesis can achieve a Query Rate Per Second (QPS) of 2000 on a single 8-core 16g server by leveraging multiple parallel recall and unified sorting and distribution, which is significantly higher than the serial recall on the same machine. This parallel recall approach offers a promising solution to enhance the system's throughput and efficiency, compared to the traditional serial recall. Moreover, the recall content module can be updated seamlessly by simply adjusting the configuration and adding concurrent threads to invoke the new recall module, thereby enabling efficient module scalability and promoting high service availability. This finding can serve as a valuable guideline for improving the scalability and reliability of the system.

5 Conclusion

The advent of the Internet has revolutionized the way businesses connect with consumers, rendering online advertising more interactive and adaptable. However, it has also resulted in a surge of high-dimensional data that needs to be processed. In this context, this paper proposes a user-based personalized recommendation advertising promotion system and conducts a series of data experiments to validate its efficacy. In potential applications, users could be randomly recommended ads with different themes or content, and marketers would make a judgment based on user feedback. The basis of information flow recommendation is traffic, and the rapid growth of traffic brings about the problem of system throughput. Our research solves this problem by embedding a multipath parallel algorithm, which avoids the reduction of advertising revenue due to fewer advertisements, and further improves the accuracy of recommendation to target users.This research is of practical significance for imple-

menting digital media technology in real-world projects. Furthermore, the findings can serve as a valuable reference for advertisers in placing ads on ad delivery platforms. Nonetheless, the algorithm only accounts for a few factors influencing the click-through rate of ads, and factors like location and bidding price are overlooked. Consequently, future work will focus on incorporating machine learning algorithms to explore the characteristics of ads and extract feature information.

Funding

This research received no external funding

Author contributions

Qin Wu contributes to methodology, formal analysis, project administration, and writing original draft preparation.

Conflict of interest

The author declares no conflict of interest.

References

[1] Bara A, Affandi F, Farid A S. The Effectiveness of Advertising Marketing in Print Media during the Covid 19 Pandemic in the Mandailing Natal Region[J]. Budapest International Research and Critics Institute-Journal (BIRCI-Journal) Vol, 2021, 4(1): 879-886.

[2] Lee H, Cho C H. Digital advertising: present and future prospects [J]. International Journal of Advertising, 2020, 39(3): 332-341.

[3] Shen Q, Wen H, Tao W. Deep interest highlight network for click-through rate prediction in trigger-induced recommendation[C]//Proceedings of the ACM Web Conference 2022. 2022: 422-430.

[4] Abdollahpouri H, Adomavicius G, Burke R. Multistakeholder recommendation: survey and research directions[J]. User Modeling and User-Adapted Interaction, 2020, 30: 127-158.

[5] Wu T, Yang F, Zhang D. Research on Recommendation system based on user portrait[C]//2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS). IEEE, 2020: 462-465.

[6] Gorgoglione M, Panniello U, Tuzhilin A. Recommendation strategies in personalization applications [J]. Information & Management, 2019, 56(6): 103143.

[7] Aljunid M F, Huchaiah M D. IntegrateCF: Integrating explicit and implicit feedback based on deep learning collaborative filtering algorithm [J]. Expert Systems with Applications, 2022, 207: 117933.

[8] Chen H, Wang Z, Huang F. Generative adversarial framework for cold-start item recommendation[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022: 2565-2571.

[9] Hui B, Zhang L, Zhou X. Personalized recommendation system based on knowledge embedding and historical behavior [J]. Applied Intelligence, 2022: 1-13.

[10] Sallam R M, Hussein M, Mousa H M. Improving collaborative filtering using lexicon-based sentiment analysis [J]. International Journal of Electrical and Computer Engineering, 2022, 12(2): 1744.

[11] Ezugwu A E, Ikotun A M, Oyelade O O. A comprehensive survey of clustering algorithms: state-of-the-art machine learning applications, taxonomy. challenges, and future research prospects[J]. Engineering Applications of Artificial Intelligence, 2022, 110: 104743.

[12] Sun C, Du M, Sun J. A three-way clustering method based on improved density peaks algorithm and boundary detection graph [J]. International Journal of Approximate Reasoning, 2023, 153: 239-257.

[13] Kliestik T, Zvarikova K, Lăzăroiu G. Data-driven machine learning and neural network algorithms in the retailing environment: consumer engagement. experience, and purchase behaviors[J]. Economics, Management and Financial Markets, 2022, 17(1): 57-69.

[14] Ahmad R, Wazirali R, Abu-Ain T. Machine learning for wireless sensor networks security: an overview of challenges and issues [J]. Sensors, 2022, 22(13): 4730.

[15] Yang Y, Zhai P. Click-through rate prediction in online advertising: A literature review [J]. Information Processing & Management, 2022, 59(2): 102853.

[16] Liu T, Li X, Ye Q. Research on Advertising Click Through Rate Prediction Based on Improved Xgboost Algorithm[J]. World Scientific Research Journal, 2022, 8(7): 226-233.

[17] Silva N, Werneck H, Silva T. Multi-armed bandits in recommendation systems: a survey of the state-of-the-art and future directions [J]. Expert Systems with Applications, 2022, 197: 116669.

[18] Silva N, Silva T, Werneck H. User cold-start problem in Multi-Armed Bandits: when the first recommendations guide the user's experience[J]. ACM Transactions on Recommender Systems, 2023, 1(1): 1-24.

[19] Aziz M, Anderton J, Jamieson K. Identifying New Podcasts with High General Appeal Using a Pure Exploration Infinitely-Armed Bandit Strategy[C]//Proceedings of the 16th ACM Conference on Recommender Systems. 2022: 134-144.

[20] Gan M, Kwon O C. A knowledge-enhanced contextual bandit approach for personalized recommendation in dynamic domains[J]. Knowledge-Based Systems, 2022, 251: 109158.

[21] Lee H, Hwang D, Min K. Towards Validating Long-Term User Feedbacks in Interactive Recommendation Systems[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022: 2607-2611.

[22] Gan M, Kwon O C. A knowledge-enhanced contextual bandit approach for personalized recommendation in dynamic domains [J]. Knowledge-Based Systems, 2022, 251: 109158.

[23] Ciucanu R, Soare M, Amer-Yahia S. Implementing Linear Bandits in Off-the-Shelf SQLite[C]//EDBT. 2022: 2:388-2:392.

[24] Erginbas Y E, Phade S, Ramchandran K. Interactive Recommendations for Optimal Allocations in Markets with Constraints[J]. arXiv preprint arXiv:2207.04143, 2022.

[25] Setiowati S, Adji T B, Ardiyanto I. Point of Interest (POI) Recommendation System using Implicit Feedback Based on K-Means+ Clustering and User-Based Collaborative Filtering[J]. Computer Engineering and Applications Journal, 2022, 11(2): 73-88.

[26] Lee Y, Lin C A. Exploring the serial position effects of online consumer reviews on heuristic vs. systematic information processing and consumer decision making[J]. Journal of Internet Commerce, 2022, 21(3): 297-319

[27] Chang, D., Wang*, Y. & Fan, R. (2022). Forecast of Large Earthquake Emergency Supplies Demand Based on PSO-BP Neural Network. Tehnički vjesnik, 29 (2), 561-571.

[28] Gao, M., Fang*, S., Wang, J., Zhang, X. & Cao, Y. (2022). A Dual Frequency Predistortion Adaptive Sparse Signal Reconstruction Algorithm. Tehniki vjesnik, 29 (2), 580-589.

[29] Zhang, L., Ma, J., Liu, X., Zhang, M., Duan, X. & Wang, Z. (2022). A Novel Support Vector Machine Model of Traffic State Identification of Urban Expressway Integrating Parallel Genetic and C-Means Clustering Algorithm. Tehnički vjesnik, 29 (3), 731-741.

C O P E

Member since 2012
JM08090

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control