Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 25, Issues 5, pp. 3531–3539, DOI 10.12694/scpe.v25i5.3072

# OPTIMIZATION OF E-COMMERCE PRODUCT RECOMMENDATION ALGORITHM BASED ON USER BEHAVIOR

YIFAN JI[*], LAN CHEN[†], AND RUI XIONG[‡]

**Abstract.** In order to implement personalized recommendation algorithms for e-commerce, the author proposes a genetic fuzzy algorithm based on user behavior to improve the sales, personalized recommendation, user satisfaction, and purchase matching performance of e-commerce. Collect data based on e-commerce personalized preference recommendation information, extract the associated feature quantities of personalized data for clustering processing, and then combine fuzzy B-means clustering method to achieve e-commerce personalized recommendation. According to the individual preferences of e-commerce, the collected data samples are fitted with differences and restructured, and a genetic evolution method is adopted for global optimization. The experimental results show that the optimized genetic fuzzy algorithm used in this method has improved stability and accuracy compared to the PSO method, with an accuracy increase of 4%. This proves that the algorithm can provide the services needed by users more quickly and is an effective means.

**Key words:** Genetic algorithm, Personalized recommendations, Fuzzy clustering, User behavior

**1. Introduction.** The internet generates a massive amount of information every day, some generated during purchases and some generated during page browsing. Behind this complex information lies the unique behavioral characteristics of each internet user. If this information is well utilized, better services can be provided more accurately for specific users [1]. With the skyrocketing amount of information on the Internet, methods for analyzing this information face severe challenges. Therefore, it is of great commercial value to analyze and predict user behavior through data mining and other related technologies, and directly recommend the predicted results to users, providing personalized information recommendation services. In recent years, after successful e-commerce models such as B2B and B2C, the O2O model has for the first time connected online virtual operations with offline physical stores, indicating that internet technology has further spread to people's daily lives. Multiple internet giants are actively participating in the O2O business. In July 2015, Baidu announced an investment of 20 billion yuan in the O2O field within three years, indicating its level of importance. In this context, the daily transaction volume of the Internet will increase, and e-commerce will be the most important market in the future [2].

However, the increasingly serious problem of information overload still troubles internet users. Selecting information that users are interested in from the ocean of data is like finding a needle in a haystack. At the same time, it is difficult to distinguish the accuracy and authenticity of information. Due to limitations in one's own abilities and concerns about expenses, a large amount of information actually brings confusion and confusion to users. In fact, what consumers want is not the amount of information, but the information tailored to their personal interests and hobbies. In recent years, internet companies have shifted their approach from passively providing information to actively guessing what information users need. Nowadays, whether using apps to listen to music or online shopping, there is a "guess what you like" section for personalized recommendations. The rapid development of e-commerce has prompted e-commerce websites to provide higher quality services and more professional information. Personalized services are an important research direction in the field of e-commerce. Currently, e-commerce platforms that only passively provide information services are no longer able to establish themselves. Only by actively attracting users and providing the goods and services they need can

[*]Sichuan Vocational College of Finance and Economics, Chengdu, Sichuan, 610101, China
[†]Sichuan Vocational College of Finance and Economics, Chengdu, Sichuan, 610101, China (Corresponding author's e-mail: chenlanncncnc@163.com)
[‡]Sichuan Vocational College of Finance and Economics, Chengdu, Sichuan, 610101, China

they attract users. Driven by this emerging business model with billions of users, new science and technology have been continuously developed, and recommendation systems have played a crucial role in it [3].

From a research perspective, the massive amount of data brought by the Internet has led to the emergence of a new research model, which has shifted from studying actual experimental objects to analyzing virtual data. In this model, the dependence on theory is not very sensitive, and there are some problems that cannot be accurately explained in theory, but have been proven effective in practical applications. This research method is a breakthrough attempt, it not only prompts the academic community to re-examine research methodology, but more importantly, it itself promotes the progress of current science and technology [4].

From a business perspective, accurate recommendation algorithms can better personalize marketing for users. When placing advertisements on a page, if recommendations are made based on the user's recent browsing data, it can improve the accuracy of advertising placement. On shopping websites, corresponding recommended products can also be provided based on users' recent purchasing behavior to increase sales. It can be seen that high-performance recommendation algorithms can generate significant economic benefits for both advertising and e-commerce businesses.

From the perspective of internet users, with the emergence of recommendation systems, we don't have to waste a lot of time finding things we are interested in from search engines or merchant lists. Instead, data providers can directly push recommendation results to us through recommendation systems on the page. This is not only an efficient way for users to obtain the required information, but also a pleasant user experience.

## 2. Methods.

### 2.1. E-commerce personalized recommendation information model and feature extraction.

*(1) Personalized information transmission.* Extract user personalized consumption data based on e-commerce correlation features, obtain fuzzy decision functions, construct non-linear mapping $\varnothing : n \in R^n \to Q$ to represent user personalized guidance space, combine data information with decision functions, and use intelligent algorithms to map to feature space F[5]. Assuming the e-commerce recommendation sample set is $\{(a_1, m_1), (a_2, m_2), ..., (a_n, m_n)\}$, the personalized feature quantity $a_i \in R^n$ represents the model input vector, $m_i \in R^n$ represents the target test value, and n represents the quantity, the personalized recommendation objective function is calculated:

$$\begin{aligned} \min \text{imize} \frac{1}{2} ||w||^2 + \underset{i=1}{\overset{n}{B}} (J_i + J_i^*) \\ \text{subject to } m_1 - (w' \varnothing(a_i) + b) \leqslant \varepsilon - J_i \\ (w' \varnothing(a_i) + b) - m_i \leqslant \varepsilon - J_i \\ J_i J_i^* \geqslant 0, i = 1, 2, ..., n; B > 0 \end{aligned} \tag{2.1}$$

In Equation 2.1, $J_i$ and $J_i^*$ are ontology attributes and association rule variables, and the cost factor is represented by B. The difference function is obtained by using control optimization method as shown in Equation 2.2.

$$Q(a) = \sum_{i=(e_i - e_i^*)}^{n} F(a_i, a_j) + b \tag{2.2}$$

In Equation 2.2, $e_i$ and $e_i^*$ represent personalized attribute values and the number of Tem plate categories, while the symmetric kernel function $F(a_i, a_j)$ represents the recommendation threshold. Based on the preference information in the database Web cloud, adaptive optimization is performed to extract the information multipath gradient map and obtain the conduction model $a_1, a_2, a_3, a_4$, which is represented as:

$$\begin{cases} a_1 = L_1 - u \\ a_2 = L_2 - u \\ a_3 = L_3 - u \\ a_4 = u \end{cases} \tag{2.3}$$

In Equation 2.3, u represents the genetic fuzzy clustering correlation attribute, and the domain of information subspace dimension is $\Omega$. Extract associated data and feature quantities based on e-commerce information, establish labels, and analyze group interaction relationships.

*(2) Extraction of associated feature quantities.* The personalized preference data is fitted with sample differences and restructured to extract associated feature quantities, $x(s), s = 0, 1, ..., n-1$, representing the time series of feature quantities in the user area. Under the constraints of relevant rules, the structural distribution function obtained is:

$$\begin{aligned} E^{bu}(b_1, b_2) =& T \cdot Length(b) + m \cdot Area(inside(b)) + P_1 \int inside(b)|i - b_1|^2 \\ &+ P_2 \int outside(b)|i - b_2|^2 dadm \end{aligned} \tag{2.4}$$

In Equation 2.4, $b_1$ and $b_2$ are the adaptive feature coefficients, Length (b) is the length coefficient, and Area (inside (b)) is the size of the feature region; Construct a feature vector set and control decision function, with a mixed kernel function as shown in Equation 2.5.

$$F_{min} = \mu F_{poly} + (1 - \mu)F_{rbf}, \mu \in (0, 1) \tag{2.5}$$

In Equation 2.5 $F_{poly} = [(a \cdot a_i + 1)]^2$, the preference trust kernel function $F_{RBQ} = exp(-a||a - a_i||^2)$ is RBQ, and the adaptive clustering process is recommended under the rule, resulting in:

$$Q_{lg-c}(o) = (Q_{lg}(o), Q_{lg-a}(o), Q_{lg-m}(o)) = (Q_{lg}(o), h_a^*(o)) \tag{2.6}$$

In Equation 2.6, $Q_{lg}(o)$ is the user's project rating value, and a quadruple is obtained based on the associated features:

$$max \begin{Bmatrix} |bh(z) - bh(z) \cap bh(z_2)| + |bh(z) \cap bh(z_2)| \\ |bh(z_2) - bh(z) \cap bh(z_2)| + |bh(z) \cap bh(z_2)| \end{Bmatrix} = max \begin{Bmatrix} bh(z) \\ bh(z_2) \end{Bmatrix} \leqslant \Delta \tag{2.7}$$

In Equation 2.7, bh(z) is a regular coefficient with correlation, which facilitates the extraction of data correlation features.

**2.2. Optimization of personalized recommendation algorithms for e-commerce.**

*(1) Genetic evolution optimization method.* Personalized data association feature extraction is based on sample difference fitting and structural restructuring. The author uses genetic fuzzy clustering method to extract association features based on kernel function construction. Utilizing the advantages of genetic algorithms, global optimization control is performed on potential user preference variables. The following Equation 2.8 expresses the meaning of the genetic evolutionary control function:

$$m_i = Tm_i(1 - m_i) \tag{2.8}$$

In Equation 2.8, T is a personalized recommendation control parameter, and the construction of $m_i \in [0, 1]$ random numbers is completed. Assuming that in a fuzzy clustering space W, u is the mutated individual, $t = \{L_1, L_2, ..., L_w\}$ is the population, $L_i^d(s)(i - 1, 2, ..., w)$ represents the user's search for individual i using latent features in the dimension space W, $V_i^w(s)(i = 1, 2, ..., t)$ is the optimization speed of individual i, $L_{best}^w(s)$ represents the best position of individual i, and $G_{best}^w(s)$ represents the optimal solution, in the process of genetic evolution, the expression for optimizing individual extremum and global extremum at each iteration is:

$$\begin{cases} N_i^w(s+1) = A \cdot N_i^w(s) + B_1 \cdot R_1(L_{best}^w(s)) \\ -L_i^w(s) + B_2 \cdot R_2 \cdot (G_{best}^w(s) - L_i^w(s)) \\ L_i^w(s+1) = L_i^w(s) + N_i^w(s+1) \end{cases} \tag{2.9}$$

In Equation 2.9, the conduction coefficient and correlation eigenvectors of particle i at the current and next time points are $N_i^w(s), N_i^w(s+1)$, and $L_i^w(s), L_i^w(s+1)$; In the formula, $B_1$ and $B_2$ represent learning factors, with values between -25 and 25; The search radius and threshold of genetic fuzzy clustering are represented by $R_1$ and $R_2$, respectively, and are randomly selected[6]. Combining genetic fuzzy clustering optimization with A, a recommended adjustment formula is obtained in interval $[A_{min}, A_{max}]$:

$$A(s+1) = 4.0w(s)(1 - A(s)) \tag{2.10}$$

$$A(s) = A_{min} + (A_{max} - A_{min})A(s) \tag{2.11}$$

The personal universality analysis of e-commerce users is based on the value of the inertia factor in Equation 2.11 and the convergence control of the pattern neural network recommendation process.

*(2) Data information mining and clustering processing.* Cluster the obtained personalized preference data and combine it with B-means to extract quadruples represented by $\{E_1, E_2, ..., E_p\}$. Extract the conduction control model under the premise of controlling constraint variables:

$$\gamma'_{desira} = \gamma_1 \cdot \frac{Density_i}{\sum_i Density_i} + \gamma_2 \frac{YL_i}{YL_{init}} \tag{2.12}$$

$$\begin{cases} \gamma_1 + \gamma_2 = 1, \gamma_1, \gamma_2 \in [0,1] \\ \gamma_2 = \frac{max_i(YL_i) - min_i(YL_i)}{YL_{init}} \end{cases} \tag{2.13}$$

Introducing radii $R_1$ and $R_2$ into cluster learning, the recommendation process update Equation 2.14 is obtained, with $R_i(s) \in (0,1), i = 1,2$ in Equation 2.14. The phenomenon of using fuzzy clustering method to jump out of the optimal value and merge with user rating measurement is described by Equation 2.15:

$$R_i(s+1) = 4.0R_i(s)(1 - R_i(s)) \tag{2.14}$$

$$N_i^w(s+1) = 4.0N_i^w(s)(1 - N_i^w(s)) \tag{2.15}$$

$$N_i^w(s) = N_{min} + (N_{max} - N_{min})N_i^w(s) \tag{2.16}$$

$$\rho_i(f,l) = \frac{\gamma(f)\sigma_{fl}d_l(o_{s+1})\mu_{s+1}(l)}{\sum_{f+1}^{V}\sum_{l+1}^{V}\gamma_s(f)\sigma_{fl}d_l(o_{s+1})\mu_{s+1}(l)} \tag{2.17}$$

The range of pheromone values for mutated individuals in Equation 2.16 is represented by $[N_{min} + N_{max}]$. Under the constraint of association rules, Equation 2.17 represents the clustering center of the association feature quantity, where $t+1(j)$ is the number of nodes, $d_l(o_{s+1})$ is the coefficient point set, and $y_{fl}$ is the score measurement information. The genetic evolution algorithm is used to calculate the variance $\delta^2$ and determine whether $\delta^2 < H$ is valid. In order to achieve personalized recommendation in e-commerce, genetic evolution and optimization processing are combined with clustering algorithms to determine whether the convergence criteria are met [7].

**2.3. Introduction to User Group Behavior Analysis Platform.** The process of the product recommendation function platform, as shown in Figure 2.1, is mainly to provide users with personalized product recommendation services [8,9]. Users log in to the platform to browse products (the products they browse are basically of interest to the user), and obtain the parameters and prices of the products, at the same time, understanding the type of product, regardless of whether the end user has successfully purchased, the user transaction database will record information such as product number, product type, and matching. The platform is divided into the following three modules to present the user behavior data analysis process:

1) User behavior tracking module: By tracking the user's clicking, browsing, bookmarking, storing and other behaviors on the website, the product data and browsing customer data are transformed into user purchasing behavior operation data, and users are grouped for the first time through log data;

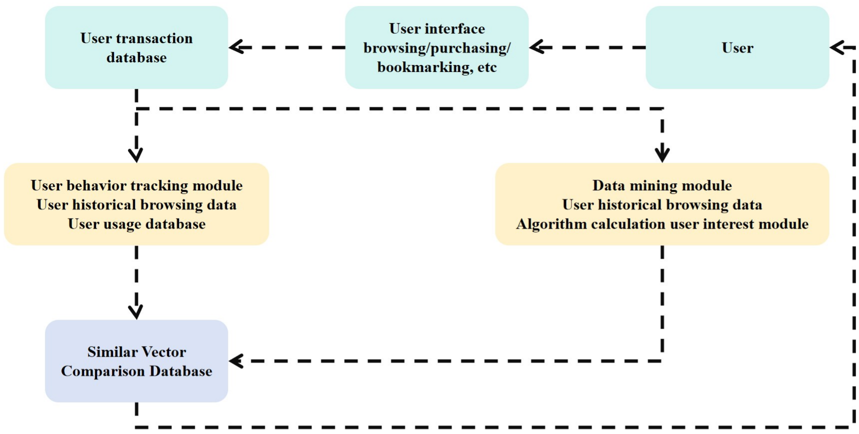


Fig. 2.1: Process of Product Recommendation Function Platform

2) Data mining module: Analyze user log data using traditional algorithms and improved fuzzy algorithms based on genetic algorithms, and obtain association rules for frequent items of users respectively. This can uncover the rule habits of users in their recent behavior;
3) Result display module: By using the user behavior tracking module to group user data, as well as the association rules of products analyzed by the data mining module, based on similarity vector comparison of user similarity, gather similar rule users, and make Top-N recommendations based on the similarity comparison results [10,11].

**3. Experiments and Results.** Using Matlab7 to design matrix simulation experiments, the data used in the experiments were all sourced from publicly available online data from JD.com and Alibaba, mainly to verify the accuracy and convergence of personalized e-commerce recommendations[12]. According to the simulation experiment parameters, Q=41 is set as the clustering center, the fuzzy control parameters are set to $b_1 = 124, b_2 = 364$, the e-commerce simulation time is 2 minutes (120 seconds), and 3120 iterations are run. In the W individual dimension space, the number is set to 1040, and the population size is 10. The experimental parameters are set to start the simulation analysis.

Test the sample information collected from e-commerce users and compare the convergence of traditional methods with the experimental process to obtain a comparison curve, as shown in Figure 3.1.

The comparison results of Figure 3.1 show that the author's use of e-commerce personalized recommendation has a very good response in terms of user recommendation satisfaction and information accuracy, which fully reflects the good performance of e-commerce personalized recommendation[13]. Analyzing and comparing personalized e-commerce recommendation methods with traditional methods, and comparing PCA algorithm with PSO, the results are shown in Figure 3.2.

The above experimental results indicate that in the personalized dimension space W of e-commerce personalized recommendation, the collection and setting of all experimental data, and the experimental process are compared with traditional methods. E-commerce personalized recommendation provides accurate user information recommendation, promotes transaction completion, and greatly improves efficiency[14]. It is confirmed that genetic evolution method is combined with B-means in global optimization, and preference data is clustered to achieve personalized recommendations for e-commerce. User satisfaction and accuracy are also reflected in Figures 3.1 and 3.2.

By collecting and analyzing a large amount of test data, three constructed methods were used for product recommendation, and the experimental data results shown in Figures 3.3 and 3.4 were obtained.

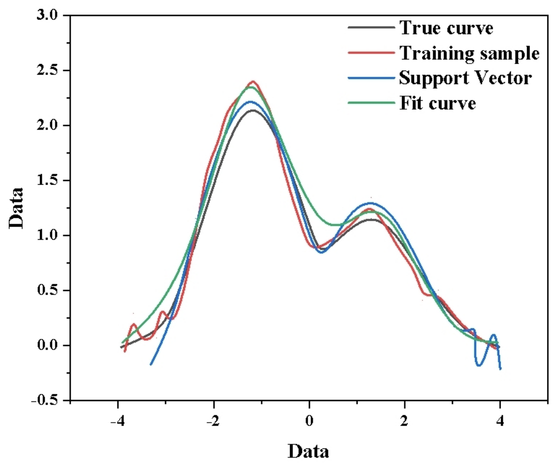


Fig. 3.1: Comparison of personalized recommendation curves for e-commerce

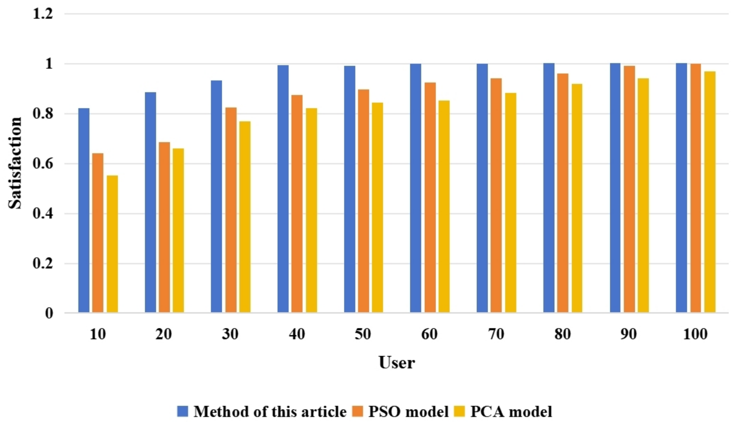


Fig. 3.2: Comparison of satisfaction with personalized e-commerce recommendations

Figure 3.3 shows the accuracy distribution of some users. It can be seen that in a small number of users, the accuracy difference between the PSO method and this method is not significant. However, as the number of users increases, this method demonstrates its advantages of high efficiency and accuracy [15].

Figure 3.4 shows the distribution of recall rates for some users. Similarly, it can be seen from the figure that when the number of users is relatively low, both accuracy and recall rates do not show the advantages of the algorithm. As the number of users increases, the improvement in system performance after algorithm optimization can gradually be seen. This also fully demonstrates that the results of big data analysis are more accurate and effective.

Finally, the author performed simple average calculations on a large amount of accuracy and recall data, and obtained experimental results as shown in Table 3.1 by calculating the time complexity of the three methods during operation.

Due to the fact that in the PCA model, the entire system does not display the function of product recommendation, but only conducts data analysis on the use and purchase of products, and presents the analysis

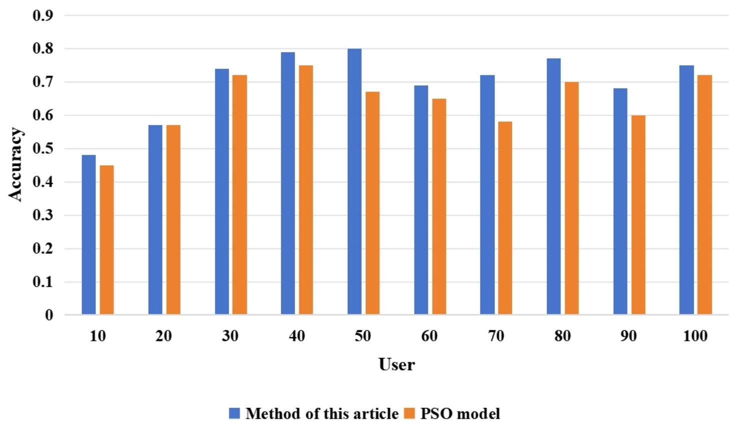


Fig. 3.3: Partial user usage accuracy chart

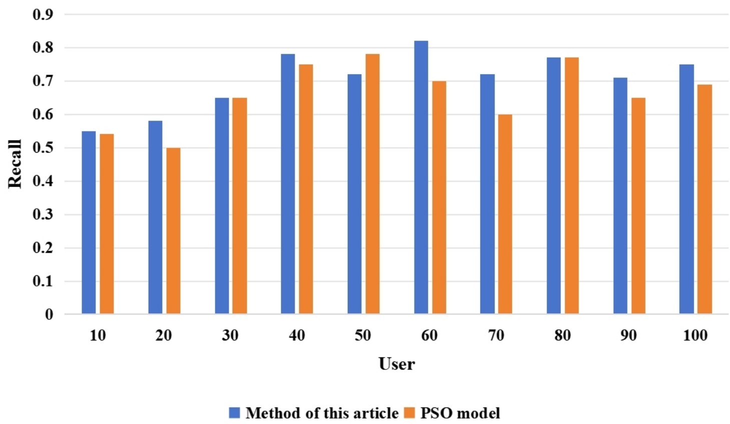


Fig. 3.4: Partial User Function Recall Rate Chart

results to users in the form of bar charts or pie charts, therefore, in the PCA model, users need to rely on these analysis data to determine the products that are suitable for their own configuration, thus unable to provide time complexity [16].

As can be seen, for the PSO method, the author uses traditional incremental mining algorithms to optimize the entire system, and the accuracy and recall of the product recommendation function have greatly increased. However, for the data platform Storm used in the system, traditional algorithms need to be compatible with its segmented data processing method and massive scanning times, which greatly increases the time and space required, the operating cost of the system has also generated tremendous pressure, therefore, the minimum time complexity of the matching algorithm is O(n). For this method, a half search is used, so the time complexity of the matching algorithm is set to 0 (1gn). Considering the cost, both types of time complexity are acceptable, but the system pursues higher efficiency. Moreover, Table 3.1 shows that the accuracy and recall of this method have also been improved to a certain extent compared to the PSO method. Therefore, considering all indicators comprehensively, this algorithm has indeed played an important role in improving product recommendation effectiveness [17,18,19,20].

From the above data analysis, it can be seen that the PSO method, using traditional incremental mining

Table 3.1: Comparison of evaluation indicators for three experimental methods

| | PCA model | The method of this paper | PSO model |
|---|---|---|---|
| Accuracy | 0.14 | 0.68 | 0.72 |
| Recall | 0.19 | 0.70 | 0.76 |
| Time complexity | Not have | O(n) | O(lgn) |

algorithms, has already achieved very good results in mining user preferences, and can more accurately and efficiently recommend products to users compared to before. And this method utilizes the optimized genetic fuzzy algorithm, which has improved stability and accuracy compared to the PSO method. More importantly, it greatly compresses the system's running time in terms of time cost, and can provide the services needed by users more quickly. It is an effective means.

**4. Conclusion.** The collection of personalized e-commerce data is sourced from publicly available data from JD.com and Alibaba. The author restructured the data and fitted the sample difference, then optimized and extracted the associated feature quantity, which was processed using the B-means clustering method to achieve personalized e-commerce recommendations. The author confirmed the effectiveness and stability of this recommendation method through experimental results. The effect of using weight increment mining is better than that of general mining, and it is also more efficient and fast. The accuracy of this experiment is as high as 72%, and the recall rate is as high as 76%, which is more accurate than other recommendation methods. Based on the above experimental methods, it can be proven that the algorithm optimized product recommendation method used by the author is an effective product recommendation strategy and has basically achieved the expected results.

**Acknowledgement.** This research was supported by the China West Normal University—Sichuan Provincial Department of Education Humanities and Social Sciences Research Base: Chengdu Chongqing Region Education and Economic and Social Collaborative Development Research Center (Grant No. CYJXF24061).

REFERENCES

[1] Li, Y., Li, Y., Sun, B., & Chen, Y. (2023). Zinc ore supplier evaluation and recommendation method based on nonlinear adaptive online transfer learning. Journal of Industrial and Management Optimization, 19(1), 472-490.
[2] El-Ashmawi, Walaa H.Ali, Ahmed F.Slowik, Adam. (2021). Hybrid crow search and uniform crossover algorithm-based clustering for top-n recommendation system. Neural computing & applications, 33(12),74-78.
[3] Zhang, Z. (2021). An optimization model for logistics distribution network of cross-border e-commerce based on personalized recommendation algorithm. Security and Communication Networks, 2021(4), 1-11.
[4] Ni, J. (2021). Predictive analysis of user behavior of e-commerce platform based on machine learning image algorithm in internet of things environment. Journal of Intelligent and Fuzzy Systems, 58(7),1-8.
[5] Hussien, F. T. A., Rahma, A. M. S., & Abdulwahab, H. B. (2021). An e-commerce recommendation system based on dynamic analysis of customer behavior. Sustainability, 13(19), 10786.
[6] Raja, D. R. K., Kumar, G. H., Basha, S. M., & Ahmed, S. T. (2022). Recommendations based on integrated matrix time decomposition and clustering optimization. International Journal of Performability Engineering,74(4), 18.
[7] Zou, F., Chen, D., Xu, Q., Jiang, Z., & Kang, J. (2021). A two-stage personalized recommendation based on multi-objective teaching–learning-based optimization with decomposition. Neurocomputing, 452(6),96-102.
[8] Peng, Z., Wan, D., Wang, A., Lu, X., & Pardalos, P. M. (2023). Deep learning-based recommendation method for top-k tasks in software crowdsourcing systems. Journal of Industrial and Management Optimization, 19(9), 6478-6499.
[9] Yang, F., Wang, H., & Fu, J. (2021). Improvement of recommendation algorithm based on collaborative deep learning and its parallelization on spark. Journal of Parallel and Distributed Computing, 148(2), 58-68.
[10] I. Glcük, & Ozsoydan, F. B. (2021). Q-learning and hyper-heuristic based algorithm recommendation for changing environments. Eng. Appl. Artif. Intell., 102(3), 104284.
[11] Fang, Z., & Chiao, C. (2021). Research on prediction and recommendation of financial stocks based on k-means clustering algorithm optimization. Journal of Computational Methods in Sciences and Engineering,147(2), 1-9.
[12] Zhong, D., Yang, G., Fan, J., Tian, B., & Zhang, Y. (2022). A service recommendation system based on rough multidimensional matrix in cloud-based environment. Computer Standards and Interfaces.(Aug), 82(34),65-69.
[13] Yan, H. C., Wang, Z. R., Niu, J. Y., & Xue, T. (2022). Application of covering rough granular computing model in collaborative filtering recommendation algorithm optimization. Advanced Engineering Informatics, 51(54), 101485.

[14] Ilker Glcük, & Ozsoydan, F. B. (2021). Q-learning and hyper-heuristic based algorithm recommendation for changing environments. Engineering Applications of Artificial Intelligence, 102(104284),96.
[15] Liu, S., Yang, Y., & Wang, Y. (2021). Integration of museum user behavior information based on wireless network. Mobile Information Systems, 2021(5), 1-8.
[16] Zhang, H., Li, X., Kan, Z., Zhang, X., & Li, Z. (2021). Research on optimization of assembly line based on product scheduling and just-in-time feeding of parts. Assembly Automation,71(5), 41.
[17] Sun, P., & Gu, L. (2021). Optimization of cross-border e-commerce logistics supervision system based on internet of things technology. Complexity,147(12),1-6.
[18] Zhong, H., Wang, Y., & Yue, W. (2022). An e-commerce product recommendation method based on visual search and customer satisfaction. International journal of knowledge and systems science,96(5),52-53.
[19] Sharma, S. N., & Sadagopan, P. (2021). Influence of conditional holoentropy-based feature selection on automatic recommendation system in e-commerce sector. Journal of King Saud University - Computer and Information Sciences,85(7),45-48.
[20] Gupta, S., & Dave, M. (2021). Product recommendation system using tunicate swarm magnetic optimization algorithm-based black hole renyi entropy fuzzy clustering and k-nearest neighbour. Journal of Information & Knowledge Management,65(2),35-38.

*Edited by:* Zhigao Zheng
*Special issue on:* Graph Powered Big Aerospace Data Processing
*Received:* Dec 25, 2023
*Accepted:* Jan 23, 2024