

## Retraction

# Retracted: Using the Mathematical Model on Precision Marketing with Online Transaction Data Computing

### Mathematical Problems in Engineering

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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

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

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

### References

- [1] J. Wu, L. Shi, G. Xu et al., "Using the Mathematical Model on Precision Marketing with Online Transaction Data Computing," *Mathematical Problems in Engineering*, vol. 2021, Article ID 5539300, 7 pages, 2021.

## Research Article

# Using the Mathematical Model on Precision Marketing with Online Transaction Data Computing

Jun Wu,<sup>1</sup> Li Shi,<sup>2</sup> Guangshu Xu,<sup>3</sup> Yu-Hsi Yuan ,<sup>4</sup> Sang-Bing Tsai ,<sup>5</sup> Weiyi Hao,<sup>1</sup> and Jiesong Jiang<sup>6</sup>

<sup>1</sup>School of Economic and Management, Beijing University of Chemical Technology, Beijing 100029, China

<sup>2</sup>College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

<sup>3</sup>School of Logistics, Beijing Wuzi University, Beijing 100029, China

<sup>4</sup>Department of Labor & Human Resources, Chinese Culture University, Taipei 11114, Taiwan

<sup>5</sup>Regional Green Economy Development Research Center, School of Business, Wuyi University, Wuyishan 354300, China

<sup>6</sup>Datang Carera (Beijing) Investment Co., Ltd, Beijing 100191, China

Correspondence should be addressed to Yu-Hsi Yuan; [yuanyh@gm.ypu.edu.tw](mailto:yuanyh@gm.ypu.edu.tw) and Sang-Bing Tsai; [sangbing@hotmail.com](mailto:sangbing@hotmail.com)

Received 25 February 2021; Revised 26 February 2021; Accepted 30 March 2021; Published 14 April 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Jun Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a decision-making system for precision marketing is presented to deal with real-world problems based on real e-business data collected in a company in Beijing. During the data preprocessing, the authors conducted a cleaning course to make sure the data to be analyzed in the latter part of the paper were credential. Based on the processed data, the authors analyzed consumer purchasing behaviors using three classic recommendation algorithms and made a performance comparison of the three algorithms. At the end of this paper, the authors proposed a series of precision marketing strategies which had been adopted by the data source company and had been proved to be effective in improving the performance.

## 1. Introduction

With the boom of e-commerce, shopping online is becoming more and more popular around the world. In China, for instance, the annual volume of online retail sales reached RMB 10,632.4 billion in 2019. While shopping online, consumers often find themselves overwhelmed by massive information of product and service. It is a challenge to deliver each consumer with best-matched options of products and services out of hundreds of thousands of candidates. To recommend desired products and services to consumers, people developed recommendation systems [1–4].

For e-commerce companies, matching consumers with the most appropriate products and services is the key to improve users' satisfaction and loyalty and consequently corporates' competitiveness. That is the reason why personalized recommendation systems are developed. Such systems enable quick and accurate matching between items

and potential users. With convincing results from practice, they are widely adopted by online business operators, including Internet giants like Amazon, Google, and Alibaba.

A personalized recommendation system is not independent, but a module attached to a main website as an IT tool. Based on analysis of demographic information, behaviors, and preferences of website visitors, the system is able to identify what attracts them the most within minutes. The principle of personalized recommendation is to present consumers with the products they are most likely to purchase. Thus, it can save time for consumers by navigating them to the desired information more efficiently.

Personalized recommendation systems have been widely adopted by e-commerce, social network, and personalized music and video services providing entities and by other platforms that interact with users. Personalized recommendation systems developed for e-commerce are able to analyze and generate forecasts with big data. They also enable efficient use of data derived from customers to

determine their personal preferences. With such a system in place, performance indicators of enterprises such as profit margin and number of active customers of a given time window can be greatly improved.

In this paper, in order to demonstrate the role of personalized recommendation systems, we base our research on a real-world scenario of an enterprise located in Beijing. We developed a decision-making system to support its precision marketing with the aid of online transaction data. After preprocessing of original data, three different algorithms were used, and their performance in the recommended scenario was compared. On this basis, we suggested a series of marketing strategies geared to the needs of customers. The proposed algorithm and strategies have been adopted by the enterprise which led to good performance.

## 2. Relevant Research Studies

Personalized recommendation systems can be roughly categorized into collaborative filtering-based type and content-based type. The first type of system mainly uses similarity within users or among products to make recommendations. Such collaborative filtering can be model-based [5] or neighbor-based [6, 7]. The most common model-based collaborative filtering algorithm and neighbor-based collaborative filtering algorithm are SVD and KNN, respectively.

**2.1. Singular Value Decomposition (SVD).** SVD is a step in collaborative filtering. In the course of matrix decomposition, it decomposes high-dimensional data through dimensionality reduction to obtain main influential factors for user ratings. Alternatively, SVD can be used to predict hidden factors related to users or items through model training and to make corresponding recommendations. In order to address inadequate rating information and the lack of user or item characteristics in SVD rating models, a Biased-SVD model was brought forward by Koren et al. [8, 9]. Later, Koren and Bell integrated an implicit item feedback function into the Biased-SVD rating prediction model, introducing the notion of SVD++ (matrix decomposition based on implicit feedback of items) to reduce the error in rating prediction [10]. In order to address the problem that rating prediction results obtained from SVD++ may lie out of a recommended rating range, Tan et al. [11] proposed PBESVD++ (Proportion-based Baseline Estimate SVD++), a proportion-based rating prediction model dealing appropriately with rating differences among users. Saia et al. [12] improved the SVD++ algorithm from the perspective of item popularity, focusing on the impact of popular items on user purchase behaviors, and proposed PSVD++ based on item popularity (Popularity-Based SVD++). Panagiotis et al. [13] put forward a multidimensional matrix factorization model called XSVD++ and combined it with collaborative filtering algorithm to improve accuracy of rating prediction.

The core concept of the SVD++ model involves linking user interests to corresponding items through implicit feature relationships. It merges implicit feedback information about users and items and maps such information into a joint implicit semantic space with a factor dimension of  $f$ . The interaction between users and items is then modeled as the inner product of the space [14].

**2.2. K-Nearest Neighbors (KNN).** The KNN algorithm is a simple and efficient data classification method. Initially brought forward by Cover et al. in 1968, it is theoretically mature and is one of the simplest machine learning approaches [15]. The core idea of the algorithm involves the following steps. Training samples and samples to be classified in a given dataset are first identified. The similarity distances between samples to be classified and training samples of known categories in the dataset are then calculated. The distances thus obtained are sorted by value, and  $K$  samples of known categories with the shortest distances are selected. Next, the weights of the samples to be classified and these  $K$  neighboring samples are derived and compared. After that, the samples to be classified with the largest weight are included in the category, and the label of this category is given as an output [16].

**2.2.1. User KNN.** The similarity among users is the main factor evaluated by User KNN algorithm. Through score weighting of the nearest neighbor set, the rating for specified items by a target user can be predicted. On this basis, items followed by users with the same preferences as the target user but not yet followed by the target user are recommended to the latter.

In 1994, user-based collaborative filtering algorithm was first proposed by Group Lens [17]. The main principle of this algorithm is to obtain a current user's rating for an item based on those given by similar neighbors. In real life, this practice is easily accepted. Since similar users tend to have similar preferences, rating for an item by a current user can be generally inferred from ratings offered by similar users. This enables prediction of current user rating and generation of corresponding recommendations.

**2.2.2. Item KNN.** Item KNN algorithm and User KNN algorithm follow quite similar implementation principles. The difference is that the former evaluates similarity between items. In 2001, Sarwar et al. proposed a collaborative filtering approach based on items, and the approach was later applied by Amazon with good recommendation results [18]. Its underlying idea is that users tend to like items similar to their favorites. By finding such similar items, appropriate recommendations can be made to users.

Although Item KNN algorithm and User KNN algorithm share similar implementation steps, it should be noted that in the case of massive users, item recommendations are typically supported by the item-based collaborative filtering algorithm for two reasons: first, a large number of users relative to items would require more computing time;

second, due to the sparse nature of rating data from typical users, similarity calculation based on items is more accurate. Therefore, the collaborative filtering algorithm based on items is able to identify similar items more quickly and accurately and consequently leads to more targeted recommendations.

### 3. Precision Marketing Framework with a Recommendation System

This section details the framework employed in this study. The framework is presented in Figure 1, which mainly includes data preprocessing, statistical analysis, user portrait generation, and recommendation algorithm.

**3.1. Data Processing.** In the real-life case adopted in this research, all data were obtained from the e-business operated by a Beijing-based company. An App named T-App was developed by the company in June 2017. It is an integrated service platform including both administration and service functions for its employees and business partners. Users of T-App are completely different from those of large-scale e-commerce platforms such as TMall.com and JD.com. Since the users are mostly employees and business partners working in the same environment, their purchasing behaviors are relatively stable and repeatable. In view of these characteristics, T-App can be classified as a typical community-based e-commerce platform [19].

Data preprocessing is a time-consuming task that can be divided into two parts: data acquisition and data cleaning [20–22]. While the acquisition part is intended to guarantee availability of data, the cleaning part is designed to ensure high data quality [23, 24]. Data acquisition mainly involves collecting useful information from original transaction data. This platform generates two types of transaction data from cooked food sales and fruit sales, respectively. The cooked food data are collected from 1,209 consumers with 14,912 transaction entries in total, and the fruit data are collected from 196 consumers with 985 transaction entries in total.

During data cleaning process, errors in original data were corrected in order to set up a standard dataset for further analysis. Typically, abnormal data may include the following:

- (1) Purchasing orders for testing purpose created by App developers.
- (2) Unpaid orders.
- (3) Unclear purchasing behaviors of departing employees.

In this research, 4,000 cooked food data entries and 430 fruit data entries were considered to be abnormal and hence removed. Following that, a basic dataset with 10,343 cooked food data entries and 541 fruit data entries was established.

**3.2. Dimensions for Statistical Analysis.** In order to describe the consumers' purchasing behaviors, we conducted statistical analyses in dimensions of item, time, and consumer. The item dimension was mainly evaluated in terms of total sales, monthly sales, and Top-K values of all items. The time dimension primarily referred to weekly, monthly, and annual cumulative sales. Daily peak purchasing hours were accounted for as well. In the user dimension, top 20% users with the highest purchasing amount were selected for analysis. Meanwhile, special case analysis was performed on top 1%, top 5%, and top 15% users with user portraits, respectively.

The statistical analysis clearly revealed different factors, including Top-K values of the most popular food/fruit varieties, Top-K values of the most valuable consumers, peak purchasing hours, etc.

**3.3. User Portrait.** User portrait is an intuitive way to present consumers' purchasing behaviors. In this paper, user portrait is described in three aspects: demographic information, behavior information, and demand information. Among them, demographic information includes basic information and social attributes; behavior information includes consumption and use behaviors; demand information includes preferences and potential demands. Jointly, they enable the creation of a user label to describe and classify consumers concisely.

A user portrait is perceived as a virtual representation of a real user. As one of the most powerful user research tools of big data analysis, it incorporates pertinent information such as users' psychology, behavior, real-time status, and scenario demand into an e-commerce service platform, thereby realizing marketing well targeted at users. Depending on different behaviors and individual characters of target users, user groups can be defined for separate feature extractions. Each user in a group can be assigned with a label. Such grouping and labeling process leads to user portraits with rich user information. Simply put, a user portrait serves as a label containing certain user characteristics such as basic properties, behavioral tendencies, interests, and preferences. It can be used to describe and categorize people in a simple way. In this paper, user portraits mainly contains two perspectives: basic information and behavior information. Basic information includes simple user profile data including name, gender, and age. Behavior information is mainly derived from purchasing data, including price, time, and consuming preferences. A user portrait example is given in Figure 2.

In this paper, consumers are sorted by total purchasing amount. In the following paragraphs, top 20% and top 50% consumers are taken as examples to present their user portraits in detail. In our research, user purchasing habits and preferences were derived from their behavior data, and

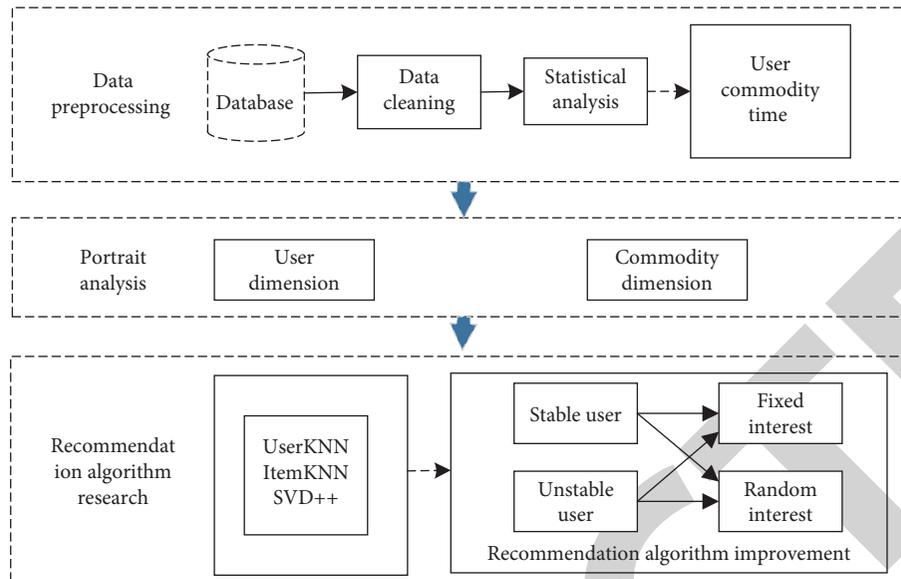


FIGURE 1: Decision framework for precision marketing.

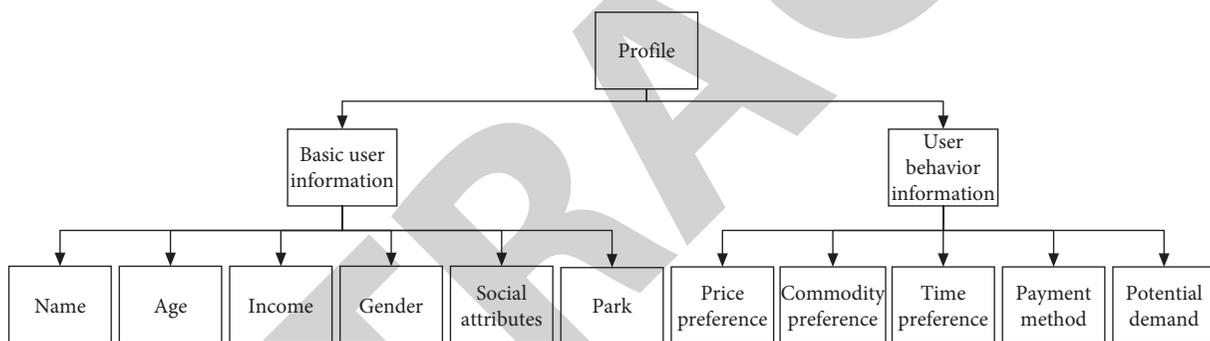


FIGURE 2: Different types of user data in a user portrait.

corresponding labels were then assigned to them, as shown in Figures 3 and 4.

Following user portrait generation, we employed the data of top 20% and 50% users to test our recommendation algorithms. In the next section, results of three classical algorithms are given and compared.

#### 4. Recommendation Algorithms

In this section, results of performance comparison of three classic recommendation algorithms are given. These algorithms were run with LibRec package in Java environment. Without loss of generality, the recommendation results of these algorithms were compared for top 5%, 20%, and 50% users, respectively. The comparison is intended to find the best recommendation algorithm that can reflect their personal preferences most accurately.

In our research, we took 99% of the total data produced by users during purchase with T-App as test data and the remaining 1% as training data. Based on actual circumstances, we used the following indices for comparison between the three recommendation algorithms.

- (1) Proportion of recommendations coinciding with past purchasing records: this reveals how many recommendations point to previously purchased items. The higher this proportion is, the more dependable the recommendations would be, and vice versa.
- (2) Maximum recommendation index: after calculation of similarity among users or that among items, an algorithm ranks the items recommended to a user based on their recommendation index. The one with the highest ranking corresponds to the maximum recommendation index.
- (3) Average recommendation index: for each of the algorithms, an average recommendation index of all items recommended to a user is calculated to reflect the level of suitability of the algorithm for the T-App dataset.

The recommendation results from the three algorithms are shown in Tables 1, 2, and 3, respectively.

Data above show that SVD++ provides better results than KNN and that User KNN provides better results than

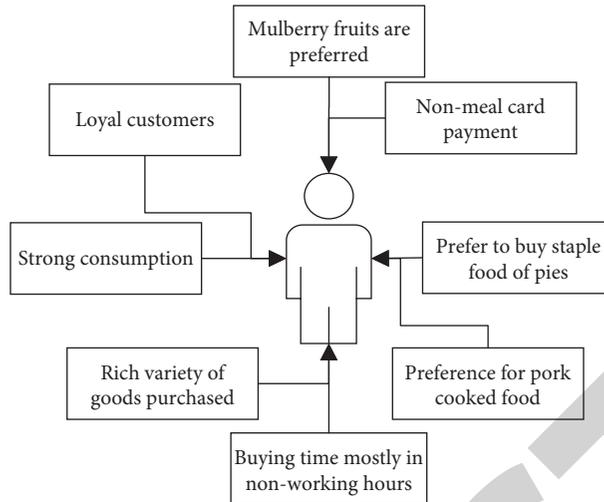


FIGURE 3: Portrait of top 20% users.

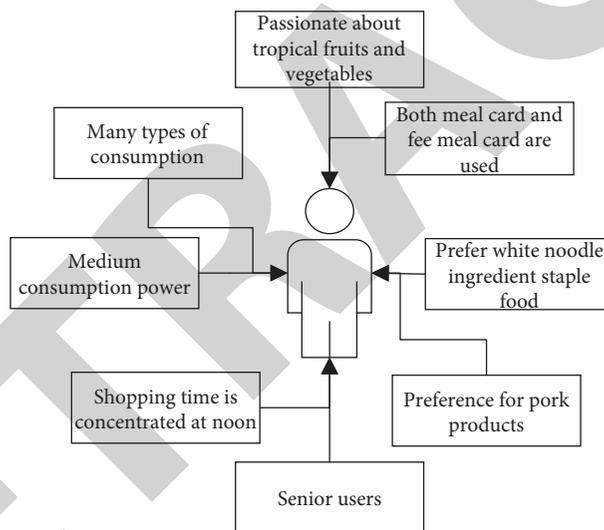


FIGURE 4: Portrait of top 50% users.

TABLE 1: Comparison of performance results for the top 5% users.

| Top 5% users  | SVD++ | User KNN | Item KNN |
|---|-------|----------|----------|
| Proportion of recommendations coinciding with past purchase records | 40%   | 30%      | 20%      |
| Maximum recommendation index  | 43.45 | 17.57    | 3.07     |
| Average recommendation index  | 19.92 | 10.78    | 2.32     |

TABLE 2: Comparison of performance results for the top 20% users.

| Top 20% users   | SVD++ | User KNN | Item KNN |
|---|-------|----------|----------|
| Proportion of recommendations coinciding with past purchase records | 30%   | 20%      | 10%      |
| Maximum recommendation index  | 38.60 | 17.40    | 3.43     |
| Average recommendation index  | 14.79 | 10.98    | 2.38     |

TABLE 3: Comparison of performance results for the top 50% users.

| Top 50% users   | SVD++ | User KNN | Item KNN |
|---|-------|----------|----------|
| Proportion of recommendations coinciding with past purchase records | 10%   | 10%      | 0        |
| Maximum recommendation index  | 18.12 | 16.52    | 2.27     |
| Average recommendation index  | 8.33  | 7.34     | 1.93     |

Item KNN. Furthermore, the following can also be concluded.

- (1) Compared with the KNN recommendation algorithm, the SVD++ recommendation algorithm fits the scenario presented in this paper much better.
- (2) The KNN algorithm often performs well on/with large e-commerce platforms. A small data size and multiple repurchasing behaviors in a relatively stable community may weaken its performance as well.
- (3) The recommendation results from Item KNN are not as accurate as reported in many references. It may be a result of the small scale of item dataset adopted in this research.

Based on the findings above, the Item KNN algorithm was recommended to the managers of the company under research. By using the results of our K-means clustering study, a completely new precision marketing strategy was developed and adopted by the company. It was proved to be very successful in practice.

## 5. Conclusions

In this paper, based on a real case, a decision-making system for precision marketing is proposed to address real problems faced by a Beijing-based company. In our research, data processing was first conducted, involving both data acquisition and data cleaning. A standard dataset was then obtained on this basis. Next, we performed statistical analyses in dimensions of item, time, and consumer to capture the consumers' purchase behaviors. With results of such analyses, we presented consumers' purchasing behaviors in an intuitive way—user portrait. Finally, we made a performance comparison among three classic recommendation algorithms on LibRec with Java language.

What we presented with this new research framework can effectively support the decision-making process in e-commerce companies and help to improve performance. In our study, customers are clustered into different groups based on their purchasing behaviors. Different CRM strategies are proposed accordingly to gain a high level of customer satisfaction. For instance, promotion activities can be held during peak purchasing hours. Top 3 popular products can be recommended to customers with high satisfaction permanently. Some new products can be widely promoted among active customers. The improvement results in some key performance indicators are given in brief as follows: the number of daily/monthly active customers has grown by 529, the total purchase volume has increased by 279%, and the total consumption amount has increased by 101.97%.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported, in part, by Funds for First-Class Discipline Construction (XK1802-5) and BUCT (G-JD202002).

## References

- [1] W. Zhang, Y. Du, Y. Yang, and T. Yoshida, "DeRec: a data-driven approach to accurate recommendation with deep learning and weighted loss function," *Electronic Commerce Research and Applications*, vol. 31, pp. 12–23, 2018.
- [2] W. Zhang, Y. Du, T. Yoshida, and Y. Yang, "DeepRec: a deep neural network approach to recommendation with item embedding and weighted loss function," *Information Sciences*, vol. 470, pp. 121–140, 2019.
- [3] W. Zhang, X. Li, J. Li, Y. Yang, and T. Yoshida, "A two-stage rating prediction approach based on matrix clustering on implicit information," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 517–535, 2020.
- [4] W. Zhang, Q. Wang, T. Yoshida, and J. Li, "RP-LGMC: rating prediction based on local and global information with matrix clustering," *Computers & Operations Research*, vol. 129, Article ID 105228, 2021.
- [5] Y. Koren, "Factor in the neighbors," *ACM Transactions on Knowledge Discovery from Data*, vol. 4, no. 1, pp. 1–24, 2010.
- [6] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, pp. 175–186, New York, NY, USA, October 1994.
- [7] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th International Conference on World Wide Web*, pp. 285–295, New York, NY, USA, May 2001.
- [8] C. C. Ma, "A guide to singular value decomposition for collaborative filtering," *Computer*, vol. 2, pp. 1–14, 2008.
- [9] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 426–434, New York, NY, USA, August 2008.
- [10] Y. Koren and R. Bell, "Advances in collaborative filtering," *Recommender Systems Handbook*, Springer, Boston, MA, 2015.
- [11] Z. Tan, L. He, H. Li, and X. Wang, "Rating personalization improves accuracy: a proportion-based baseline estimate

- model for collaborative recommendation,” in *Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pp. 104–114, Springer, Shanghai, China, October 2016.
- [12] R. Saia, L. Boratto, and S. Carta, “Exploiting the evaluation frequency of the items to enhance the recommendation accuracy,” in *Proceedings of the International Conference on Computer & Information Technology*, pp. 108–113, Libertad City, Ecuador, October 2016.
- [13] P. Symeonidis and D. Malakoudis, “Multi-modal matrix factorization with side information for recommending massive open online courses,” *Expert Systems with Applications*, vol. 118, pp. 261–271, 2019.
- [14] N. I. E. Peng, “An SVM-based multi-dimension factor decision-making model framework,” in *Proceedings of the 2016 International Conference on Applied Mechanics, Electronics and Mechatronics Engineering*, pp. 89–94, Beijing, China, May 2016.
- [15] Y. J. Su, Z. Y. Deng, D. B. Cheng, and M. Zong, “Fast KNN classification algorithm under big data,” *Application Research of Computers*, vol. 33, no. 4, pp. 1003–1006, 2016.
- [16] D. L. Lu, Q. Ning, and J. Zang, “Improved KNN algorithm based on BP neural network decision making,” *Journal of Computer Applications*, vol. 37, pp. 65–67, 2017.
- [17] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, “An algorithmic framework for performing collaborative filtering,” in *Proceedings of the International ACM SIGIR Conference on Research & Development in Information Retrieval*, July 1999.
- [18] A. B. Barragáns-Martínez, E. Costa-Montenegro, J. C. Burguillo, M. Rey-López, F. A. Mikic-Fonte, and A. Peleteiro, “A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition,” *Information Sciences*, vol. 180, no. 22, pp. 4290–4311, 2010.
- [19] N. Luo, Y. Wang, M. Zhang, T. Niu, and J. Tu, “Integrating community and e-commerce to build a trusted online second-hand platform: based on the perspective of social capital,” *Technological Forecasting and Social Change*, vol. 153, Article ID 119913, 2020.
- [20] L. Fabisiak, “Web service usability analysis based on user preferences,” *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 1–13, 2018.
- [21] A. Shahri, M. Hosseini, K. Phalp, J. Taylor, and R. Ali, “How to engineer gamification,” *Journal of Organizational and End User Computing*, vol. 31, no. 1, pp. 39–60, 2019.
- [22] T. Grublješič, P. S. Coelho, and J. Jaklič, “The shift to socio-organizational drivers of business intelligence and analytics acceptance,” *Journal of Organizational and End User Computing*, vol. 31, no. 2, pp. 37–64, 2019.
- [23] S. Namasudra and P. Roy, “PpBAC,” *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 14–31, 2018.
- [24] G. Khatwani and P. R. Srivastava, “Impact of information technology on information search channel selection for consumers,” *Journal of Organizational and End User Computing*, vol. 30, no. 3, pp. 63–80, 2018.