Research on Explainable Personalized Recommendation Method for Balancing User Long- and Short-term Preferences (Postprint)

Authors: Li Weiqing, Chi Maomao, Wang Weijun

Date: 2023-04-01T16:02:52+00:00

Abstract

[Purpose/Significance] To address issues such as increasingly complex recommendation models, growing data inputs, low interpretability of traditional recommendation models, and "over-specialization" of recommendation results, this paper proposes an interpretable personalized recommendation method oriented toward user long-short preference adjustment. [Method/Process] From the two dimensions of users' recent product needs and their long-term lifestyle, a user long-short preference model is constructed. By drawing on user rating bias and attention mechanisms, users' long-short preferences are combined with their ratings for rating prediction, thereby forming Top-N recommendations. [Results/Conclusion] Experimental results on two datasets demonstrate that the proposed method exhibits good performance across different user behaviors (explicit feedback or implicit feedback), different numbers of recommended items, and different recommendation algorithms. Without requiring significant modifications to various recommendation models, it improves the accuracy, recall, and diversity of recommendation results. Additionally, by adjusting the long-short preference coefficients, it achieves adjustment of both recommendation result diversity and accuracy, and generates corresponding recommendation explanations.

Full Text

An Explainable Personalized Recommendation Method Based on Adjustment of Users' Long- and Short-Term Preferences

Li Weiqing[1,2], Chi Maomao[1], Wang Weijun[2]

[1]School of Information Management, Central China Normal University, Wuhan 430079

[2]Key Laboratory of Adolescent Cyber Psychology and Behavior, Ministry of Education, Central China Normal University, Wuhan 430079

Abstract: [Purpose/Significance] To address the current problems of increasingly complex recommendation models, growing data inputs, low interpretability of traditional recommendation models, and "over-specialization" of recommendation results, this paper proposes an explainable personalized recommendation method based on adjustment of users' long- and short-term preferences. [Method/Process] We construct a user preference model from two dimensions: users' recent product needs and their long-term lifestyles. Drawing on user rating bias and attention mechanisms, we combine users' long- and short-term preferences with their ratings for score prediction, thereby forming Top-N recommendations. [Result/Conclusion] Experimental results on two datasets demonstrate that our method performs well across different user behaviors (explicit or implicit feedback), different numbers of recommended items, and different recommendation algorithms. Without requiring substantial changes to various recommendation models, it improves the accuracy, recall, and diversity of recommendation results. Additionally, by adjusting the coefficients of long- and short-term preferences, it achieves tunable diversity and accuracy of recommendation results while generating corresponding recommendation explanations.

Keywords: personalized recommendation; user preference; explainability; long- and short-term preference adjustment

Classification Number: G203
DOI: 10.13266/j.issn.0252-3116.2021.12.010

Personalized recommendation has long been a research focus in both academia and industry. The success of a personalized recommendation system largely depends on its ability to identify and represent user preferences. How to comprehensively and deeply understand user behavior and preferences, and how to effectively integrate user behavior and product feature data into recommendation models, are key to improving recommendation system performance [1]. Many scholars have conducted extensive explorations and research in this area.

For instance, H. Naji [2] first applied Bayesian classification to users' psychological characteristics and demographic information, then combined it with collaborative filtering algorithms. Y. Huang et al. [4] enriched user preference attributes from lifestyle and emotional dimensions, constructing a seven-dimensional online lifestyle dictionary based on user network review texts and incorporating it into a classic collaborative filtering recommendation system to improve user preference prediction performance. Hou Yinxiu et al. [5] considered the emotional characteristics in user preferences and used sentiment analysis to mine book attributes from user book review information, proposing a personalized book recommendation method based on matching user emotional preferences with product attribute sentiments. In recent years, deep learning has attracted widespread attention and research in the recommendation field due to its excellent feature extraction and representation capabilities. J. Bi et al. [6] used user average ratings, basic user data (gender, age, occupation, user ID), basic item data (name, category, item ID), and item average ratings to build a deep neural network-based user rating prediction model for generating recommendations. Experimental results showed that this algorithm not only outperformed state-of-the-art collaborative filtering algorithms but also effectively addressed data sparsity and cold start problems. Liang Changyong et al. [7] proposed a recommendation model based on Convolutional Neural Networks and Latent Factor Model (CNN-LFM), which used LFM to mine rating data for user and item latent features, employed CNN to extract visual content from images to learn item latent features, and effectively combined CNN with LFM to leverage both item content information and user-item interaction information, thereby improving user preference prediction. X. Zhao et al. [3] incorporated Weibo users' demographic information and product features extracted from product reviews into a similarity ranking algorithm to form recommendations.

In summary, existing personalized recommendation research has achieved significant progress in feature extraction and representation, accurate identification of user preferences, recommendation prediction accuracy, and alleviation of data sparsity and cold start problems. However, this has also led to several issues: (1) To improve recommendation performance, newly proposed models have become increasingly complex and redundant, raising algorithmic complexity, and frequent model modifications may hinder practical application and iteration in production environments; (2) The "black box" nature of deep learning results in insufficient explainability of recommendation results, leaving users unclear about the basis for recommendations and potentially reducing their satisfaction and trust in the system; (3) The "over-specialization" problem in recommendation results—excessive pursuit of accuracy may lead to overly homogeneous recommendations and repeated suggestions, decreasing user satisfaction.

Therefore, this paper proposes an explainable personalized recommendation method based on adjustment of users' long- and short-term preferences. We construct a long- and short-term preference model based on users' recent product categories of interest and their long-term lifestyles. Drawing on user rating bias and attention mechanisms, we combine users' long- and short-term preferences with their ratings for score prediction to generate Top-N recommendations. Specifically, we first obtain users' direct ratings for individual products based on explicit and implicit feedback. We then construct a user long- and short-term preference model according to users' recent needs, long-term lifestyles, and behavioral time cycles. This preference information is incorporated into the recommendation system to build a User-Item preference matrix, with adjustable preference weight coefficients to achieve an explainable recommendation model oriented toward user long- and short-term preference adjustment. The innovations of this method are: (1) Based on lifestyle theory, we categorize product features and construct a user long- and short-term preference model from the dimensions of recent needs and long-term lifestyles, enabling more comprehensive and detailed user preference prediction and improving recommendation accuracy; (2) We enhance the explainability of recommendation results from both users' recent product needs and long-term lifestyle perspectives; (3) We explore an adjustable mechanism for personalized recommendation system evaluation metrics, enabling tuning of recommendation accuracy and diversity through changes in long- and short-term preference coefficients, providing insights for addressing the "over-specialization" problem.

Related Research

Regarding the identification of users' long- and short-term preferences and explainable personalized recommendation, scholars have conducted extensive research. We introduce relevant achievements in these two aspects below.

2.1 User Long- and Short-Term Preference Identification

How to comprehensively and deeply understand users' interests and preferences is key to the success of recommendation systems. Preference generally refers to an individual's selective attitude toward objective things, manifested as intention to recognize, explore, approach, or acquire certain objects, and is the most obvious expression of personality [1]. Personal preferences can be manifested through behavior, so user browsing, clicking, purchasing, and reviewing behaviors can be used to infer their preferences and concerns within a certain period [8]. User preferences have stable aspects but also change over time, age, and experience. Therefore, user preferences include two main components: short-term preferences reflecting users' recent attention and interest changes, and long-term preferences reflecting users' persistent traits and psychological factors. The combination of both can more accurately characterize complete user preference information [9]. Users' online behaviors, purchase decisions, and preferences are influenced by many factors such as emotions, functions, and contexts, exhibiting complex, diverse, and variable characteristics. User behavior is explicit but cannot fully represent their intrinsic psychological features, and external behavior changes rapidly and is unstable. Therefore, how to extract users' intrinsic, stable factors influencing purchase behavior from their external behavior as a basis for recommendation is the key to this research. Purchase behavior is determined by external factors such as culture and society, as well as personal factors such as lifestyle and values [10]. Lifestyle is how people arrange their life patterns according to a central goal, expressed through language, behavior, and interests. Lifestyle is also understood as "the way individuals spend time and money." It is the external manifestation of opinions and attitudes formed by individuals' needs, values, and experiences, significantly influencing purchase behavior. Lifestyle is a consistent pattern of what individuals buy and don't buy, what they do and don't do, and what attracts or fails to attract them [11]. I. Sarki et al. [12] found that individuals' values and lifestyles are significantly correlated with their brand preferences. W. Swinyard [13] discovered that individuals' lifestyles are significantly correlated with purchase frequency and monetary and time costs. Pan Yu et al. [14] proposed that users' lifestyles significantly influence product functional value, perceived cost, and image value, with product price affecting perceived cost and product brand significantly correlating with image value. Luo Lijuan et al. [15] found that online lifestyles directly affect individual users' functional perception, brand perception, service perception, and purchase behavior. These studies demonstrate that lifestyle significantly influences users' choices regarding product price, brand, function, and other factors, thereby affecting purchase behavior. However, due to the abstractness and difficulty in identifying users' intrinsic characteristics, existing recommendation systems seldom consider how to incorporate users' lifestyle features into recommendations.

Existing research has also extensively explored how to identify and represent users' long- and short-term preferences in recommendation models: For example, L. Xiong et al. [16] proposed using a decay function f(x) = e^(-αt) to quantify the impact of users' past ratings on their preferences, where t is the time when the rating was given and α is the decay rate controlling time. Wang Weijun et al. [17-18] found that users have screen visual hot zones during browsing and extracted short text information in real-time from these hot zones as reference for users' immediate attention preferences, then combined users' instant preferences, long- and short-term preferences, and contextual factors to form an integrated user preference model for interactive convergent personalized recommendation. L. Hu et al. [19] proposed a graph neural network news recommendation model based on users' long- and short-term interests, extracting users' recent and long-term interaction items and news content features to predict users' long- and short-term information preferences. D. Chen et al. [20] proposed a sequence-aware model based on long- and short-term attention memory networks, which vectorizes user sessions within each period (embedding) and uses a memory neural network with attention mechanism to learn the weights of long- and short-term preferences. L. Xiang et al. [21] proposed a session-based temporal graph algorithm to simultaneously model users' long- and short-term preferences, combining this preference model with random walk algorithms to form a time-aware personalized recommendation. Wang Weiwei et al. [22] defined user interests expressed in current update cycles as short-term preferences and user interests expressed in continuous browsing history as long-term preferences, proposing an interest model update algorithm based on user behavior feedback that analyzes users' purchase records and browsing behaviors to achieve real-time preference updates and generate personalized recommendation lists for dynamic preferences. Overall, existing user long- and short-term preference models mostly operate at the algorithmic level, assigning weights to user behaviors such as sessions, browsing, purchasing, and rating across different time periods to adjust the importance of behavioral information in different time windows, or automatically learning to extract user long- and short-term preferences through long short-term memory networks in deep learning. These approaches lack deep mining of the differences between user long- and short-term preferences from the perspective of users' intrinsic traits.

2.2 Explainable Personalized Recommendation

When users must decide among different alternatives or evaluate recommendation quality, displaying only recommendation lists without explanatory information makes it difficult for them to determine the usefulness of recommended items [23]. The explainability of recommendation results is crucial for the development and improvement of recommendation systems. Explainable recommendation is a method that explains why an item is recommended. In recent years, it has become a widely studied topic in many web applications such as social media, e-commerce, and content sharing sites. Appropriate display of recommendation reasons can improve user acceptance of recommendation results. By explaining how the system works or why an item is recommended, the system becomes more transparent, thereby enhancing user trust and potentially helping users identify when the system makes mistakes, enabling better and faster decisions and improving user satisfaction [24-25]. Yu Yisheng et al. [26] categorized current explainable recommendations into three types: (1) Using the most contributory factor as the explanation, requiring each factor to correspond to an element that can serve as a recommendation reason, such as user or item features and tags; (2) Using predefined rule phrases, such as Amazon's "Customers who bought this item also bought…" or social media's "Your friends also followed/viewed this content," which are easy to deploy but offer generic reasons with limited effect on improving user trust and promoting purchases; (3) Introducing text review features as recommendation reasons, such as Y. Zhang et al. [27] extracting explicit user opinions on product aspects from reviews to understand which details users care more about, thereby proposing explainable suggestions. Additionally, deep learning attention mechanisms can help algorithms capture data features that require focused attention. For example, interpretable deep learning rating prediction models proposed in literature [29-30] can display which parts of reviews are more important for predicting user-item ratings based on learned attention weights, highlighting important words in reviews as explanations to help users understand recommendations.

In summary, recommendation explanations are derived from the algorithms and data used by corresponding systems, with differences manifested in generation methods and presentation forms. The main logic is to present the reasons for recommendation results in specific forms, such as user or product similarity, or user preferences for certain products or features. Based on this, our study combines the definition of online lifestyle and its influencing factors on user purchase decisions, using users' price and brand preferences as their long-term lifestyle preference features, and product categories as users' recent need preferences. Drawing on the idea of rating bias [26,31], we integrate users' long- and short-term feature preferences into their direct ratings, and combine attention mechanisms [28-30] to adjust feature weights, making recommendation results focus on user preference features with higher weights and forming corresponding explanations.

Research Method

3.1 Product Direct Rating Based on User Behavior

User behavior includes five aspects: rating, purchasing, collecting (adding to cart), clicking, and browsing time. Among these, rating is explicit feedback, while the others are implicit feedback. The importance hierarchy of these five factors is: browsing < clicking < collecting < purchasing < explicit rating [32]. Based on this, this study adopts a 5-point rating method to define the above behaviors, recording each successful user-product interaction as a direct rating p_{ui}.

3.2 User Long- and Short-Term Preference Model

User interests expressed in the current cycle are called short-term preferences. If within cycle T, a user's purchase, collection, and browsing frequency for products with certain features increases, their preference for such features is relatively high, or their preference for products related to such features may be high. The short-term preference for a certain feature in the current cycle is recorded as:

$$p_{now} = \alpha h_{ui} + \beta p_{ui}$$

where h_{ui} represents the attention frequency to products with this feature in the current cycle T, recorded as $h_{ui} = \Sigma_{j=1} tf_{uj}$ [22]. The corresponding long-term preference for this feature is expressed as:

$$p_{per} = \Sigma_{i=1} e^{-i} \times p_{now}$$

Combining users' short-term and long-term preferences, the overall user preference is expressed as:

$$p = x * p_{now} + y * p_{per}$$

where x and y are coefficients, and x + y = 1 [22].

Based on this, we apply one-hot encoding to product categories and brands, and use the decile principle to divide all prices of products in the same category into intervals. We represent users' long- and short-term preferences for product features as the following triple: $(p_{ij}G, p_{ij}P, p_{ij}B)$.

$$p_{ij}G = x((n+1-i)/10 + \beta * p_{ui})$$

$$p_{ij}P = y1(\Sigma_{i=1} e^{-i} \times (h_{ui} * 0.75)/(n*0.75 + 1))$$

$$p_{ij}B = y2(\Sigma_{i=1} e^{-i} \times (h_{ui} * 0.75)/(n*0.75 + 2))$$

where p_{ij} represents product j in cycle i, $p_{ij}G$ represents the user's short-term preference for product category, $p_{ij}P$ represents the user's long-term preference for perceived price, $p_{ij}B$ represents the user's long-term preference for brand, n represents the number of time cycles, and x, y1, y2 are coefficients with x + y1 + y2 = 1.

We use the entropy weight method [33] to calculate users' long-term preference weights y1 and y2. The entropy weight method is an objective weighting approach that determines indicator weights based on the concept of information entropy. Information entropy: $E = -\Sigma_{i=1}^n p_i \log p_i$, represents the average uncertainty of all possible situations in an information source. Larger entropy values indicate higher uncertainty. Based on this, users' long-term perceived price preference and brand preference are expressed as formulas (8) and (9):

$$E_{uP} = -\Sigma_{i=1}^n f_u(P_i) \log f_u(P_i)$$

$$E_{uB} = -\Sigma_i f_u(B_i) \log f_u(B_i)$$

where $f_u(P_i)$ is the probability that user u behaves toward products with perceived price P_i. More behavior instances result in higher probability. For each user's long-term preference weights, smaller entropy values E_{uP} and E_{uB} indicate more concentrated corresponding preferences, which should be given higher weights. Therefore:

$$y1 = (1-x) * (E_{uB})/(E_{uP} + E_{uB})$$

$$y2 = (1-x) * (E_{uP})/(E_{uP} + E_{uB})$$

3.3 Rating Bias

Recommendation models predict ratings by capturing interactions between users and items. However, ratings are related not only to users' direct ratings for items but also to users themselves or items themselves. The bias term considers users' rating strictness and item quality, enabling the model to better fit users' true preferences and significantly improving recommendation performance [27]. This is expressed as:

$$R_{ui} = r_{ui} + b_i + b_u$$

where r_{ui} is user u's rating for product i, b_i is product bias, b_u is user bias, and R_{ui} is the adjusted preference of user u for product i. Based on this, we incorporate users' long- and short-term preferences as bias terms to represent their comprehensive preferences, as in formula (13):

$$R_{ui} = p_{ui} + p_{ij}G + p_{ij}P + p_{ij}B$$

3.4 Long- and Short-Term Preference Weight Adjustment Based on Attention Mechanism

Attention mechanisms are primarily used in various deep learning domains. Their core design originates from human visual attention mechanisms. Human vision always quickly scans the global scene and immediately focuses on target areas requiring attention—the so-called attention focus—then allocates more attention resources to these areas to obtain needed detailed information while suppressing useless information. The selection of deep learning attention models is similar to visual attention mechanisms, with the core goal of selecting features more critical to current output results from numerous input features. This is widely used in natural language processing, image recognition, speech recognition, personalized recommendation, and other deep learning tasks.

In RNN Encoder-Decoder models, the attention mechanism can be briefly expressed as:

$X = \{x_1, x_2, x_3, \dots x_n\}$

$h_t = f(x_t, h_{t-1})$

$S_j = q(\{h_1, h_2, h_3, \dots h_n\})$

$\alpha_{ij} = \mathrm{align}(h_i, S_j)$

where X represents the input feature vector of the Encoder network, h_t represents the output state of each neuron in the Encoder network, and S_j represents the final output of the Encoder network after one forward propagation. When executing the Decoder process, the attention model is activated to calculate the weight relationship α_{ij} between S_j and each neuron output $\{h_1, h_2, \dots h_n\}$ through the align function, with $\Sigma_n \alpha_{ij} = 1$ [29]. Based on this idea, by assigning different values to x in formula (4), we adjust the relationship between the predicted target rating R_{ui} and users' long- and short-term preferences.

Experiments and Results

4.1 Research Data and Rating Calculation Process

We validated the proposed recommendation method using two datasets. One is historical data from Suning.com provided by a partner (Data_1); the other is the Amazon e-commerce dataset (Data_2: http://snap.stanford.edu/data/amazon/productGraph/categoryFil
Data_1 has no user ratings, so we calculated implicit feedback as users' direct product preferences using the method in Section 2.1. The Amazon dataset only has user ratings, which we used as direct product ratings. The statistical description of the cleaned datasets is shown in Table 1 .

We used the methods in Section 2 to calculate users' preference scores for each product in terms of category, brand, and price. We illustrate the user rating calculation process using behavioral data from one user in Data_1. The user's partial behavioral data is as follows:

00000066c84079975340f9ed898825551 | 000000000126823313:1509940376000:BROWSE
| 000000000109830038:1509940475000:BROWSE | 000000000109830038:1509940444000:BROWSE
| 000000000126507919:1508812552000:CART | 000000000137280398:1508025116000:BROWSE
| 000000000143689868:1509670166000:OTHER | 000000000148922124:1509936814000:BUY
| …

where 00000066c84079975340f9ed898825551 is the user ID. Each behavior is separated by "|". In each behavior, the first column (e.g., 000000000126823313) is product ID, the second column (e.g., 1509940376000) is timestamp, and the third column is behavior type, including BROWSE, OTHER, CART, COLLECT, and BUY. This user has 247 behaviors in total.

(1) We sorted each user behavior in reverse chronological order by timestamp, with the most recent behavior first. Time cycle selection and division have multiple methods and are not the focus of this study. For calculation convenience, we divided each user's online behaviors into 10 equal cycles by behavior count. For this user, after reverse sorting, we divided behaviors into cycles 1-25, 26-50, …, 201-225, and 226-247.

(2) We encoded each user behavior as described in Section 2.1: BROWSE = 1, OTHER = 2, CART and COLLECT = 3, BUY = 4.

(3) We calculated the adjusted rating for each product. Taking the product 000000000148922124 (Haier washing machine XQG80-HB14636, category: major appliances/washing machine, brand: Haier, price: 3299) purchased by this user as an example. The purchase behavior occurred in cycle 2, during which the user paid attention to washing machines 5 times. According to formulas (1) and (4):

p_{ij}G = x * ((10+1-2)/10 + β * 4) = 1.228x

To highlight the impact of product features on user preferences, we set $\alpha = 0.75$ and $\beta = 0.25$. We then calculated the user's brand preference and price preference. The Haier brand product appeared (7, 12, 9, 11, 3, 0, 0, 6, 4, 0) times in each cycle. The price decile for the 3299 washing machine is 0.4, as shown in Table 2 .

Washing machines with price decile 0.4 appeared (4, 6, 5, 6, 7, 9, 3, 4, 11, 7) times in each cycle. According to formulas (5) and (6):

p_{ij}P = y1(0.055(e^{-1} * (40.75)/(24.7+1) + e^{-2} (60.75)/(24.7+1) + … + e^{-9} (11*0.75)/(24.7+1))) = 0.726y1

p_{ij}B = y2(0.055(e^{-1} * (40.75)/(24.7+2) + e^{-2} (60.75)/(24.7+2) + … + e^{-10} (7*0.75)/(24.7+2))) = 0.664y2

(4) We calculated coefficients y1 and y2. This user paid attention to 9 brands and 7 price levels with frequencies (43, 21, 52, 9, 15, 12, 17, 49, 29) and (8, 13, 35, 62, 58, 53, 18) respectively. According to formulas (8), (9), (10), and (11):

E_{uB} = -(0.174log0.174 + 0.085log0.085 + … + 0.117log0.117) = -2.036

E_{uP} = -(0.032log0.032 + 0.053log0.053 + … + 0.073log0.073) = -1.751

y1 = (1-x) * (-2.036)/(-2.036 - 1.751) = 0.538(1-x)

y2 = (1-x) * (-1.751)/(-2.036 - 1.751) = 0.462(1-x)

(5) Finally, according to formula (13), we obtained the user's adjusted rating for this product:

R_{ui} = 4 + 1.228x + 0.391(1-x) + 0.307(1-x) = 4.698 + 0.53x

4.2 Evaluation Metrics

We adopted Top-N recommendation accuracy (Pre@N), recall (Rec@N), and diversity (Div@N) as evaluation metrics. Accuracy represents the proportion of correctly recommended items among all recommended items. Recall represents the proportion of correctly recommended items among actually purchased items. Generally, higher accuracy and recall indicate better recommendation algorithm performance, as shown in formulas (18) and (19):

Pre@N = |L_t   L_r| / |L_r|

Rec@N = |L_t   L_r| / |L_t|

where L_t is the set of items actually browsed by the user, and L_r is the recommendation result set.

Diversity of recommendation results refers to the degree to which they satisfy users' interests across different domains. Diversity describes the differences between items in a recommendation list. Assuming S(i,j) describes the similarity between items i and j, the diversity of user u's recommendation list R(u) is defined as:

Diversity_u = 1 - (Σ_{i,j R(u), i≠j} S(i,j)) / (0.5|R(u)|(|R(u)|-1))

The overall diversity of the recommendation system is the average of all users' recommendation lists:

Diversity_{all} = Σ_{u U} Diversity(R(u)) / |U|

4.3 Result Analysis

Based on the above recommendation algorithms, we used 10-fold cross-validation (Hold-Out Cross Validation) with recommendation list sizes N = (5, 10, 20), calculating and comparing recommendation accuracy, recall, and diversity across different long- and short-term preference weights x = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0).

4.3.1 Overall Performance on Different Datasets On both datasets, using all recommendation algorithms and different recommendation list sizes (Top-N), the overall average performance of each evaluation metric is shown in Figure 1 [Figure 1: see original paper]. The horizontal axis represents the value of long- and short-term preference weight x, the vertical axis represents the average value of each metric, with Data_1 results on the left and Data_2 results on the right.

Overall, on both datasets, the introduction of weighted long- and short-term preferences significantly improved the average accuracy, recall, and diversity of recommendation results. For Data_1, as x increased, accuracy and recall improved slowly, while diversity peaked around x = 0.4. For Data_2, as x increased, accuracy and recall improved significantly, while diversity peaked around x = 0.6.

4.3.2 Performance with Different Numbers of Recommended Items (Top-N) Using Data_1, we compared the average performance of all recommendation algorithms across different recommendation list sizes (Top-N) and different x values. According to the definition of accuracy, for the same recommendation algorithm, larger N generally leads to lower accuracy, as shown in Figure 2 [Figure 2: see original paper], where accuracy is highest at N=5 and lowest at N=20. However, when $0.2 < x < 0.7$, Pre@N=10 > Pre@N=5, also demonstrating that our method improves recommendation accuracy.

For recall, larger N leads to higher recall, and recall shows an upward trend as x increases. For diversity, when N=5, recommendation diversity continues to rise as x increases. When N=(10, 20), diversity reaches its maximum when x (0.4-0.8). This indicates that when N is small, our method improves both accuracy/hit rate and diversity. When N is large, diversity decreases as accuracy improves.

4.3.3 Performance Analysis in Different Recommendation Algorithms Using Data_1 with N=10, we compared the performance of various recommendation algorithms at different x values, as shown in Figures 3 [Figure 3: see original paper], 4 [Figure 4: see original paper], and 5 [Figure 5: see original paper]. The horizontal axis represents the short-term preference weight x, and the vertical axes represent accuracy, recall, and diversity respectively.

Figures 3 and 4 show that for all baseline recommendation algorithms, using our method to recalculate consumer preferences significantly improved recommendation performance, demonstrating the effectiveness of our approach. In terms of accuracy and recall, UserCF performed worst, followed by ItemCF, likely because collaborative filtering algorithms have lower overall performance with sparse data. The relatively best performers were CNN-based and LSTM-based recommendation algorithms, possibly because these algorithms take item features and user behavioral time sequence information as input, and our method adds user long- and short-term feature preference scores to user preference scores, combining them more accurately represents user preferences. It may also be due to deep learning algorithms' superior feature extraction and representation capabilities and their deep, linear and non-linear combined model structures, which deliver superior performance on prediction tasks.

For diversity, various recommendation algorithms mostly reached maximum diversity when x (0.6-0.8).

These results demonstrate that our explainable personalized recommendation method based on user long- and short-term preference adjustment performs well across different datasets, different recommendation quantities (Top-N), different recommendation algorithms, and different user preference calculation methods (explicit or implicit feedback), significantly improving recommendation performance. Without requiring substantial changes to recommendation algorithm models, it improves recommendation accuracy and recall. Additionally, by adjusting long- and short-term preference weight coefficients, it can tune recommendation accuracy and diversity, and generate corresponding explanations based on preference features with higher weights.

Conclusion and Outlook

Based on online lifestyle theory, this study uses product category information as users' short-term preferences and product perceived price and brand information as users' long-term preferences. Drawing on rating bias and attention mechanism concepts, we incorporate users' long- and short-term preferences and direct ratings to construct a User-Item preference matrix, adjusting preference weight coefficients to form explainable recommendation lists oriented toward different user long- and short-term preferences, achieving good recommendation results. This research has the following innovations and values:

(1) **No major changes to original recommendation models required**: Only a new user long- and short-term preference calculation module is needed to significantly improve various recommendation models' performance, avoiding increased complexity of original models and facilitating rapid industrial application.

(2) **Adjustable accuracy and diversity of recommendation results**: Overall, when short-term preference coefficient $x < 0.6$, increasing x significantly improves accuracy, recall, and diversity. When $x \geq 0.6$, diversity decreases while accuracy changes insignificantly. Thus, our algorithm can adjust recommendation accuracy and diversity based on changes in long- and short-term preference coefficients, providing ideas for solving the "over-specialization" problem.

(3) **Explainability of recommendation results**: Explainability means recommendation results have clear and reasonable explanations based on algorithm logic [2]. Generally, in knowledge-based recommendation algorithms, explanations are based on encoded rules, while in collaborative filtering algorithms, explanations are based on how user preferences are obtained [25]. For UserCF, user relationships are calculated based on different users' ratings for the same product, with recommendations made between users with similar preferences, so its explanation is "Users who like product X also like Y." For ItemCF, relationships between products are obtained by calculating different users' ratings for different products, with similar items recommended to target users, so its explanation is "Because you like product X, you may also like product Y." For our method, users' long- and short-term preferences are added as rating bias to their direct ratings. When the short-term preference weight coefficient is large, the algorithm focuses more on product category information, and its expla-

nation is "You have recently paid attention to this category of products and may like product X." When the long-term preference weight coefficient is large, the explanation is "Based on your lifestyle (brand or price preference), we recommend product X to you." Our algorithm provides recommendation reasons from both users' recent product needs and long-term lifestyle preferences, offering higher rationality and richness of explanations. Compared with deep learning recommendation algorithms that can also combine user features and behavioral time sequence information, attention mechanisms can identify which factors are more relevant to recommendation results, but the massive computational load of attention models makes real-time recommendation difficult to achieve.

(4) **Changing target values in prediction tasks**: Our study changes user ratings according to certain rules to research the impact of product features on user preferences, providing new ideas and methods for feature engineering research.

Due to space limitations, this paper only introduces how to combine users' long- and short-term preferences with their direct ratings using rating bias and attention model concepts, applied to various recommendation algorithms. Future work will explore how to use user long- and short-term preference features as direct inputs to recommendation models and construct new explainable recommendation models based on the characteristics of various recommendation algorithms. Additionally, how to generate personalized explanations based on users' long- and short-term preferences and different recommendation results, as well as personalized and adaptive adjustment mechanisms for weight parameters, will be future research priorities.

References

[1] LIU P, ZHANG L, GULLA J, et al. Dynamic attention-based explainable recommendation with textual and visual fusion[J]. Information processing & management, 2019, 57(6): 102099.

[2] NAJI H. Improving recommendation system based on homophily principle and demographic[J]. Research journal of applied sciences, 2016, 11(10): 1102-1106.

[3] ZHAO X, GUO Y, HE Y, et al. We know what you want to buy: a demographic-based system for product recommendation on microblogs[C]//ACM. KDD '14: proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM press, 2014: 1935-1944.

[4] HUANG Y, LIU H, LI W, et al. Lifestyles in amazon: evidence from online review enhanced recommender system[J]. International journal of market research, 2020, 62(6): 689-706.

[5] HOU YINXIU, LI WEIQING, WANG WEIJUN, et al. Research on personalized book recommendation based on matching user preferences and product

attribute sentiments[J]. Data analysis and knowledge discovery, 2017, 1(8): 9-17.

[6] BI J, LIU Y, FAN Z, et al. A deep neural networks based recommendation algorithm using user and item basic data[J]. International journal of machine learning and cybernetics, 2020, 11(4): 763-777.

[7] LIANG CHANGYONG, FAN RUXIN, LU WENXING, et al. Personalized recommendation based on CNN-LFM model[J]. Computer simulation, 2020, 37(3): 399-404.

[8] BEENEN G, LING K, WANG X, et al. Using social psychology to motivate contributions to online communities[C]//ACM. Proceedings of the 2004 ACM conference on computer supported cooperative work. New York: ACM press, 2004: 212-221.

[9] BENNETT P. Modeling the impact of short-term and long-term behavior on search personalization[C]//ACM. International ACM SIGIR conference on research & development in information retrieval. New York: ACM, 2012: 185-194.

[10] CHU Y, LEE J. The experiential preferences of the online consumers in different internet shopping lifestyles toward online shopping websites[C]//Springer. International Conference on Human-Computer Interaction. Lisbon: Springer, 2007.

[11] DIAZ A, GOMEZ M, MOLINA A, et al. A segmentation study of cinema consumers based on values and lifestyle[J]. Journal of retailing and consumer services, 2018, 3(41): 79-89.

[12] SARKI I, BHUTTO N, ARSHAD I, et al. Impact of Pakistani university student's cultural values and lifestyles on meaning of brands[J]. Interdisciplinary journal of contemporary research in business, 2012, 3(9): 643-654.

[13] SWINYARD W R. Shopping mall customer values: the national mall shopper and the list of values[J]. Journal of retailing and consumer services, 1998, 5(3): 167-172.

[14] PAN YU, GAO LI, WANG FANGHUA. Influence of lifestyle and customer perceived value on Chinese consumers' purchase behavior[J]. Journal of systems & management, 2009, 18(6): 601-607.

[15] PAN YU, LUO LIJUAN, LIU DAN, et al. Research on individual users' cloud service purchase intention based on online lifestyle[J]. Journal of systems & management, 2013, 22(4): 477-488.

[16] XIONG L, CHEN X, HUANG T, et al. Temporal collaborative filtering with Bayesian probabilistic tensor factorization[C]//SDM. Proceedings of the SIAM international conference on data mining. Philadelphia: SIAM, 2010: 211-222.

[17] ZHANG TINGTING, WANG WEIJUN, HUANG YINGHUI, et al. Real-time extraction method of Chinese short text keywords based on screen visual

hot zones[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(12): 1178-1188.

[18] LIU KAI. User Preference Extraction and Personalized Recommendation Based on Screen Visual Hot Zones[M]. Beijing: Science Press, 2016.

[19] HU L, LI C, SHI C, et al. Graph neural news recommendation with long-term and short-term interest modeling[J]. Information processing and management, 2020, 57(2): 1-10.

[20] CHEN D, ZHANG R, YUAN B, et al. Sequence-aware recommendation with long-term and short-term attention memory networks[C]//IEEE. 2019 20th IEEE international conference on mobile data management. IEEE, 2019: 437-442.

[21] XIANGL, YUAN Q, ZHAO S, et al. Temporal recommendation on graphs via long- and short-term preference fusion[C]//ACM. International ACM SIGIR conference on research and development in information retrieval. New York: ACM press, 2014: 83-92.

[22] WANG WEIWEI, XIA XIUFENG, LI XIAOMING. An interest model update algorithm based on user behavior feedback[J]. Journal of Liaoning University (Natural Sciences Edition), 2011, 38(1): 40-45.

[23] GEDIKLI F, JANNACH D, GE M, et al. How should I explain? A comparison of different explanation types for recommender systems[J]. International journal of human-computer studies, 2014, 72(4): 367-382.

[24] TINTAREV N, MASTHOFF J. Designing and evaluating explanations for recommender systems[M]. Boston: Recommender Systems Handbook, Springer, 2011: 479-510.

[25] ZHANG Y, CHEN X. Explainable recommendation: a survey and new perspectives[J]. ArXiv: information retrieval, 2018, 14(1): 1-101.

[26] YU YISHENG, WEI RUI, LIU XINYAN. Research on explainable real-time book information recommendation model[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(2): 209-216.

[27] ZHANG Y, LAI G, ZHANG M, et al. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis[C]//ACM. International ACM SIGIR conference on research and development in information retrieval. New York: ACM press, 2014: 83-92.

[28] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[EB/OL]. [2021-04-01]. https://www.researchgate.net/publication/265252627_{{{Neural}}*{{Machine}}*}{{{Translation}}*{{by}}*}{{{

[29] HU L, JIAN S, CAO L, et al. Interpretable recommendation via attraction modeling: learning multilevel attractiveness over multimodal movie contents[C]//AAAI. Twenty-seventh international joint conference on artificial intelligence IJCAI-18. Stockholm: AAAI Press, 2018, 7: 3400-3406.

[30] SEO S, HUANG J, YANG H, et al. Interpretable convolutional neural networks with dual local and global attention for review rating prediction[C]//ACM. RecSys'17: proceedings of the eleventh ACM conference on recommender systems. New York: ACM press, 2017, 8: 297-305.

[31] YU J, ZHU H, CHANG C, et al. Influence function for unbiased recommendation[C]//ACM. SIGIR'20: the 43rd international ACM SIGIR conference on research and development in information retrieval. New York: ACM press, 2020, 7: 1929-1932.

[32] CUI CHUNSHENG. User clustering analysis in e-commerce implicit browsing input[J]. Library and information service, 2011, 55(14): 130-134.

[33] WANG JUN, ZHANG ZIKE. Personalized recommendation algorithm based on social tagging information entropy[J]. Library and information service, 2013, 57(23): 31-35.

[34] LIU H, HU Z, MIAN A, et al. A new user similarity model to improve the accuracy of collaborative filtering[J]. Knowledge-based systems, 2014, 56(1): 156-166.

[35] DESHPANDE M, KARYPIS G. Item-based top-N recommendation algorithms[J]. ACM transactions on information systems, 2004, 22(1): 143-177.

[36] POLAT H, DU W. SVD-based collaborative filtering with privacy[C]//ACM. SAC'05: proceedings of the 2005 ACM symposium on applied computing. New York: ACM press, 2005, 3: 791-795.

[37] WU Y, LI Y, QIAN R. NE-UserCF: collaborative filtering recommendation model based on NMF and E2LSH[J]. International journal of performance engineering, 2017, 13(5): 610-619.

[38] BI J W, LIU Y, FAN Z P. A deep neural networks based recommendation algorithm using user and item basic data[J]. International journal of machine learning and cybernetics, 2020, 11(4): 763-777.

[39] DEVOOGHT R, BERSINI H. Long and short-term recommendations with recurrent neural networks[C]//ACM. UMAP'17: proceedings of the 25th conference on user modeling, adaptation and personalization. New York: ACM press, 2017, 7: 13-21.

Author Contributions

Li Weiqing: Responsible for data processing and paper writing;
Chi Maomao: Responsible for paper revision;
Wang Weijun: Responsible for paper conceptualization, framework design, and article revision.

Abstract (English)

Explainable Personalized Recommendation Method Based on Adjustment of Users' Long- and Short-Term Preferences

Li Weiqing[1,2], Chi Maomao[1], Wang Weijun[2]

[1]School of Information Management, Central China Normal University, Wuhan 430079
[2]Key Laboratory of Adolescent Cyber Psychology and Behavior, Ministry of Education, Central China Normal University, Wuhan 430079

Abstract: [Purpose/Significance] We put forward an explainable personalized recommendation method based on adjustment of users' long- and short-term preferences in view of the current problems that increasingly complexity and more feature data inputs of recommendation models, low interpretability of traditional recommendation models and over-specialization of recommendation results. [Method/Process] We constructed a user preference model from two dimensions of users' recent product needs and their long-term lifestyles, used the user's rating bias and attention mechanism for reference, combined the user's long- and short-term preferences with their direct score to predict the score of unknown items, and formed the Top-N recommendations. [Result/Conclusion] The experimental results on two datasets showed that our method had a good performance to different user behaviors (explicit feedback or implicit feedback), different number of Top-N recommended items, and in different recommendation algorithms. It improves the accuracy, recall and diversity of the recommendation results without making great changes to various recommendation models, and based on the change of long- and short-term preference coefficients, it realizes the adjustment to the diversity and accuracy of the recommendation results, and form the corresponding recommendation explanation.

Keywords: personalized recommendation; user preference; explainable recommendation; adjustment of long- and short-term preference

Note: Figure translations are in progress. See original paper for figures.