

Article

Capturing Dynamic User Preferences: A Recommendation System Model with Non-Linear Forgetting and Evolving Topics

Hao Ding ^{1,2}, Weiwei Zhu ^{1,*}, Guangwei Hu ^{3,*} and Zhan Bu ⁴¹ School of Management, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China² Research Center for Education Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing, 210023, China³ School of Information Management, Nanjing University, Nanjing, 210023, China⁴ School of Computer Science, Nanjing Audit University, Nanjing, 210023, China

* Correspondence: zhuww@njupt.edu.cn (W.Z.); hugw@nju.edu.cn (G.H.)

Abstract

Though recommendation systems can help users save time while shopping online, their performance is significantly limited by sparse user data and the inability to capture temporal dynamics of user preferences, such as interest forgetting and topic evolution in reviews. Existing studies primarily focus on static user–item interactions or partial temporal signals (e.g., rating timestamps) but fail to comprehensively model two critical aspects: the non-linear decay of user interests over time, where users gradually forget historical preferences, and the semantic evolution of review topics, which reflects implicit shifts in user preferences across different periods. To address these limitations, we propose a Temporal Dynamic Latent Review-aware Preference Model with Matrix Factorization. Our model integrates an adaptive forgetting-weight function to simulate users’ interest decay and a multi-interval latent topic model to extract evolving preference features from review semantics. Specifically, we design a joint optimization framework that dynamically weights user ratings based on temporal forgetting patterns and decomposes review texts into latent topic factors to alleviate data sparsity. Finally, the experiments employ five baseline methods on six datasets to test the recommendation performance, validating its effectiveness in tracking users’ temporal interest drift and improving recommendation accuracy.



Academic Editor: Vladimír Bureš

Received: 31 August 2025

Revised: 1 November 2025

Accepted: 14 November 2025

Published: 19 November 2025

Citation: Ding, H.; Zhu, W.; Hu, G.; Bu, Z. Capturing Dynamic User Preferences: A Recommendation System Model with Non-Linear Forgetting and Evolving Topics. *Systems* **2025**, *13*, 1034. <https://doi.org/10.3390/systems13111034>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: time series; interest fluctuation; recommendation system; latent factors; topic evolution

1. Introduction

With the rapid growth of e-commerce platforms, temporal recommendation systems face two critical challenges. The first is sparsity of user–item interactions, in that users’ explicit feedback (e.g., ratings) is often limited, leading to inaccurate preference modeling [1,2]. The second is dynamic preference evolution, which means users’ interests exhibit non-linear decay (e.g., forgetting historical preferences) and semantic shifts (e.g., changes in review topics over time). The existing studies suffer from three key limitations. First, most temporal recommendation methods primarily rely on rating timestamps but ignore non-linear interest forgetting [3,4]. For instance, they assume linear decay of user preferences, which contradicts psychological studies on human memory. Second, review-based approaches extract static topic features from reviews but fail to capture temporal topic dynamics [5]. For example, a user’s focus on “battery life” in electronics reviews may shift to “design aesthetics” over months, yet existing models treat all reviews as equally relevant.

Third, hybrid methods attempt to combine ratings and reviews but lack a unified framework to jointly optimize temporal decay and topic evolution. This results in suboptimal performance under data sparsity [6,7].

Although previous studies have separately explored temporal decay or comment topic modeling, no work has simultaneously modeled nonlinear interest decay and the semantic evolution of multi-period comment topics within a unified optimization framework. To overcome these limitations, the proposed TDLRP-MF model not only integrates these two types of dynamic signals but also achieves fine-grained perception of user interest drift through joint matrix factorization and latent topic word confidence modeling. Our model introduces three contributions:

1. We proposed a novel multi-feature adaptive time weighting approach. It comprehensively considers the impact of user interest drift features on recommendation generation to enhance the accuracy of user preference prediction.
2. We constructed a novel latent topic model for multi-interval individual topic reviews. The model can more accurately perceive hot topics in user reviews at different periods and capture the dynamic evolution of user preferences in time steps.
3. We built a feature fusion model of multi-dimensional latent factors. It can learn user ratings' synergistic temporal evolution features and multiple latent preference features of the topics to achieve high-accuracy user preference perception in data-sparse scenarios.

2. Related Work

2.1. Recommendation Based on Temporal Matrix Factorization

Traditional recommendation methods can be recommended by collaborative filtering [8], but with the rapid increase in the variety and number of products, the sparsity of data problem produces trouble, so people consider the time factor of the matrix decomposition of the matrix dimensionality reduction to improve the data sparsity problem. TimeSVD++ captures the evolving nature of user preferences by probability matrix [9]. This approach introduces time-evolving discrepancies for people and goods within the latent space. However, it overlooks the interconnections among shifts in user preferences at specific times, potentially missing crucial time-specific preference aspects. Moreover, it lacks consideration for auxiliary data, reducing its effectiveness in scenarios with sparse data. BPTF employs Bayesian probabilistic tensor decomposition to estimate ratings [10]. It offers insights into the evolution of latent features and ratings by utilizing the latent vectors. Nonetheless, it might not be sufficiently sensitive to detect subtle, localized variations in user preferences. TCMF is designed to address data sparsity by incorporating temporal dynamics and multimodal data [11]. However, it heavily relies on implicit feedback from user reviews and overlooks the implicit implications present in supplementary data. In TM-RevCo, temporal dynamics, and auxiliary data are considered to improve performance [12]. This is achieved by using item relevance measures in the auxiliary factors and correlating them with the matrix decomposition item factors. While this method considers user preferences across different periods, it does not consider the potential effects of varying time stages on users. This could result in significant deviations between predicted and actual user behavior, ultimately reducing the effectiveness of recommendations.

2.2. Recommendation Based on Deep Learning with Interest Drift

In addition to matrix decomposition methods, deep learning is also an important branch of research in the field of recommendation, and in recent years, a large number of deep learning-based methods have emerged to mine the evolutionary characteristics of user preferences. HITUCF is a hybrid recommendation algorithm that integrates similarity

and forgetting curves [13]. While the method partially emulates the forgetting patterns of human interests, it does not account for the temporal evolution of user review information. FSTS aims to address user interest drift, low accuracy, and long tail issues by extracting item features using stability variables and time-sensitive factors [14]. It also uses feature vectors and time-sensitive factors in the prediction process to enhance recommendation performance with minimal time complexity. However, its time-sensitive factor may not be sufficient to capture dynamic shifts in user interest accurately. TSCIA is a temporal collaborative interest awareness model that explicitly considers similar interests across adjacent sessions to model overall collaborative interests [15]. It addresses the issue of user interest drift and captures patterns of user preference changes more accurately. However, it does not take into account the time-varying characteristics of items. ADARIFT uses personalized learning rates to handle user interest drift and incorporates item features in adaptive recommender systems [16]. However, it does not utilize auxiliary information. ALSTM is a sentiment rating model utilizing long and short-term memory that addresses data sparsity and imbalance by combining sentiment bias in user ratings and reviews. While it effectively models the drift in user and item features, it does not delve into mining latent semantics embedded in review texts [17]. Additionally, some scholars have considered the influence of forgetting laws on item selection by studying the forgetting characteristics of users' interests. TPNE is a module with temporal positive and negative incentive modeling built through the static-dynamic interest learning framework for more accurate sequential recommendations [18]. TDADLFM is a time decay adaptive latent factor model, which assumes that people's preferences usually change over time [19]. It adds an attention mechanism to adjust the user's embedding vectors dynamically and combines a time decay function with an attention model for item score prediction. Although these methods take into account user interest drift, it does not consider the interaction between user comments and ratings, which may affect the prediction of user preference evolution.

2.3. Recommendation Based on Review Mining

Some studies have been conducted to identify user preference features for items by inferring user interest changes from semantic features of textual reviews. These reviews serve as a reference information source for recommender systems. ADR is a recommendation approach focusing on deep reviews, supplementing item representation with auxiliary vectors, and performing user-side modeling based on a user attention mechanism [20]. MRMRP is a multivariate review-based rating prediction model. It can extract valuable features from the supplemental reviews to enhance recommendation performance using deep learning methods [21]. ERP is an enhanced review-based rating prediction model that distinguishes the impact of reviews generated by different users through user influence. It also uses auxiliary and review information to learn item-perceived user preferences [22]. LUAR is a method that employs a neural attention mechanism to automatically identify and prioritize crucial auxiliary reviews in the dataset [23]. PSAR is a persona-driven sentiment-focused recommendation model that aims to learn segment-level and sequence-level personalized sentiment representations from user reviews [24]. DARMH is a multi-task dual-attention recommendation model that leverages reviews and their usefulness [25]. The method uses interactive and local attention mechanisms separately to capture personalized preferences for specific items by specific users. DeepCGSR is a deep recommendation model that combines textual review sentiment and rating matrices to conduct cross-granularity sentiment analysis by treating the set of user reviews of items as a corpus [26]. This allows the method to gain sentiment feature vectors of both users and items by associating fine-grained and coarse-grained level information. ERUR is a review-based user representation model based on learning social graphs, which enhances user modeling by integrating user reviews and

social relationships to achieve better recommendation results [27]. Prior research on recommendation systems that employ review topics has shown that considering various factors, such as semantics and context of user reviews, can enhance recommendation effectiveness. However, these studies do not explore the temporal dynamics of the recommendation process, which may lead to the loss of temporal interest drift characteristics of users. MTUPD is a method that incorporates multiple leaps of user interest drift [28]. The process utilizes multiple leap factors and adaptive temporal weights to gauge the relevance of user preferences across various periods using a forgetting curve function. However, this method only considers the topic model with non-negative matrix decomposition and does not explain the user review topic evolution law at a deeper level.

3. Preliminary

The paper introduces a temporal recommendation model with a latent factor that utilizes users' ratings and reviews. In this section, we will discuss the initial steps and define and describe the fundamental concepts of the model. Additionally, the main symbols and parameters used in this paper are listed in Table 1.

Table 1. Main symbols and parameters.

Notation	Description
T	Time periods.
τ	Time decomposition dimension.
$R^{(t)} \in \mathbb{R}^{i \times j}$	The ratings matrix at time t .
$R^{(t-T)} \in \mathbb{R}^{i \times j}$	The ratings matrix at time $t - T$.
$U_R^{(t)} \in \mathbb{R}^{i \times \tau}$	User latent factor matrix.
$V_R^{(t)} \in \mathbb{R}^{\tau \times j}$	Item latent factor matrix.
$D_R^{(t-T)} \in \mathbb{R}^{\tau \times \tau}$	Drift transition matrix.
W^{t-T}	The enhanced time weighting.
l	Number of review topics.
c	Number of review topic words.
$G^{(t)} \in \mathbb{R}^{l \times c}$	Review topic words weight matrix.
$Z^{(t)} \in \mathbb{R}^{l \times c}$	Review topic words probability matrix.
$P^{(t)} \in \mathbb{R}^{l \times c}$	Review topic words preference weight matrix.
$SD(\bullet)$	Standard deviation.
$\ \bullet\ _p$	The p -norm.
$\ \bullet\ _F$	The Frobenius norm.
α	Time weighting balance parameter.
β	Topic words feature weight adjustment coefficient.
γ	Temporal decomposition adjustment parameter.
σ_1, σ_2	Regularization parameters.

3.1. Problem Promotion

User rating preference features change dynamically over time and show forgetfulness or indifference to specific item interests over time, while users also do reviews based on their interest points, which can contain richer topic preference features than considering only rating information. In contrast, conventional collaborative filtering recommendation methods struggle to thoroughly gauge how changes in users' various interests affect the recommendation system. Hence, this paper aims to closely mimic the temporal changes in users' multiple features, striving to enhance the system's recommendation accuracy. The proposed recommendation model is illustrated in Figure 1.

Let t denote the current period requiring the recommendation, while $t - T$ represents a historical period relative to t , determined by the frequency of recommendation updates on the social media platform. To illustrate, in a total span of 16 months with $T = 6$, if t

corresponds to the 16th month, then $t - T$ pertains to the timeframe spanning from the 10th to the 15th month before t . We give the function definitions as follows:

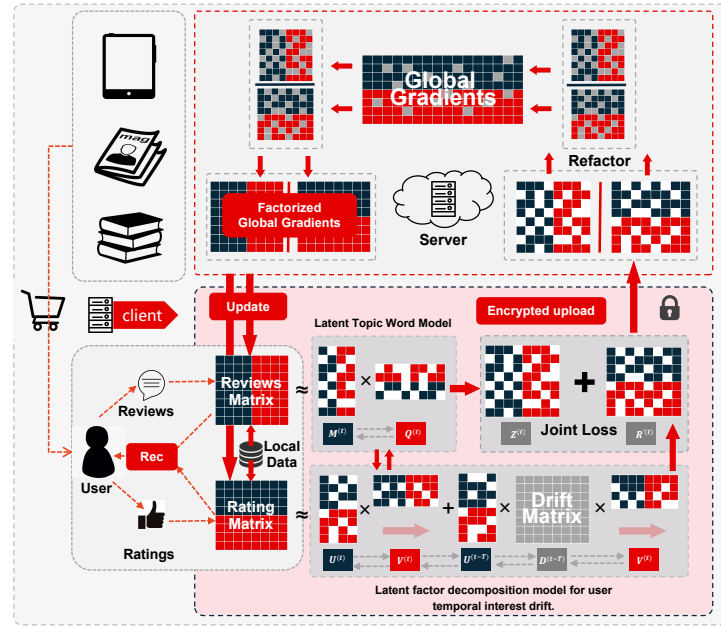


Figure 1. The proposed TDLRP-MF model diagram.

3.1.1. Adaptive Temporal Forgetting-Weight

Ebbinghaus demonstrates that human forgetting does not decrease linearly, and the rate of forgetting is typically faster at first and then slows down. While various methods are used in previous studies to simulate human forgetting patterns [29,30], in reality, people are often affected by multiple factors that influence their perception and forgetting process. Thus, this paper aims to develop new forgetting weights based on memory and forgetting laws, considering the impact of real-life environments on individuals' forgetting and simulating the relationship between users' interest in items over time. From this, we propose a new time-weighted formulation that is defined as shown in Equation (1).

$$H^{t-T}(rl_{ij}) = \mu e^{-\delta T} + \varepsilon \quad (1)$$

The rl_{ij} denotes the viewing relationship of the i th user's j th item, the t be the time at which user i rates item j , and the t -moment rating time weight of user i on item j . The fitting parameters μ , δ , and ε are used to calculate the memory residual rate $H^{t-T}(rl_{ij})$ for an item from the $t - T$ moment to the t moment. A higher value of the H function indicates that users have a stronger memory for the item's content. The main focus of the forgetting function in this paper is to accurately reflect the change in users' interest in the item. To simulate the forgetting characteristics of users' interest, the forgetting curve tracing points of Ebbinghaus were selected as (0, 100), (1/3, 58), (1, 44), (9, 36), (24, 33), (48, 28), (144, 25), and (744, 21). The horizontal coordinates of these data points represent time in hours, while the vertical coordinates represent the memory retention rate in percentage. Using MATLAB 2024a's Curve Fitting Tool, the data was fitted according to the definition of the H function. The most recent item rating is then used to represent the user's current interest. After normalizing the fitting results, the final interest forgetting weighting function is defined as shown in Equation (2):

$$H^{t-T}(rl_{ij}) = \mu e^{\delta \left[1 - \frac{T}{|t_{max} - t_0|} \right]} + \varepsilon \quad (2)$$

where $\mu = 0.6509, \delta = 0.8023, \varepsilon = 0.3481$, t_0 denotes the most recent evaluation time of the item, t_{max} denotes the longest evaluation time of the item by the user, and T denotes the time interval between the current moment and the initial viewing moment.

Users often retain a degree of memory regarding items due to their enduring preferences [31]. For instance, if users show consistent interest in items with literary qualities, they are likely to continuously seek out such items over time. However, solely relying on the weight of interest and forgetting an item might diminish the lasting impact of specific features on sustained interest. Therefore, this article integrates the temporal progression of user preferences for item reviews across various time frames, as depicted in Equation (3):

$$s_i^{t-T} = \frac{\sum_{p=1}^{|P_i^{(t)}|} \sum_{q=1}^{|P_j^{(t-T)}|} \text{Sim}(wv_{rlp}, wv_{rlq})}{|P_i^{(t)} \ominus P_j^{(t-T)}|} \quad (3)$$

where $P_i^{(t)}$ and $P_j^{(t-T)}$ represent the sets of items preferred by users at moments t and $t - T$, respectively. The symbol $\ominus$ denotes the set containing elements unique to each set and does not overlap. $\text{Sim}(wv_{rlp}, wv_{rlq})$ indicates the similarity between the word vectors of user reviews wv_{rlp} and wv_{rlq} . To eliminate the influence of deactivated words, the initial step involves utilizing the deactivated word lexicon from the text. Subsequently, the words are tokenized using Jieba. Finally, Word2vec is employed to train the word vectors and determine their similarity, as demonstrated in Equations (4) and (5):

$$\theta = \frac{wv_{rlp} \bullet wv_{rlq}}{\|wv_{rlp}\| \times \|wv_{rlq}\|} \quad (4)$$

$$\text{Sim}(V_{rlp}, V_{rlq}) = \varphi(a^{wv}\theta + b^{wv}) \quad (5)$$

The word vector cosine similarity is denoted by θ , the logistic sigmoid function is denoted by φ , and the weight matrix and bias vector are denoted by a^{wv} and b^{wv} , respectively. The review weights s_i^{t-T} reflect the stability of the evolution of user interest in item review topics over time for user i . Larger values of s_i^{t-T} indicate more stable changes in topic interest over time.

Through the integration of forgetting weights and the evolution of review weights, the ultimate model can be depicted using the temporal weights as indicated in Equation (6):

$$W^{t-T} = \alpha * H^{t-T} + (1 - \alpha) * s^{t-T} \quad (6)$$

Here, α signifies the weight balance coefficient employed for self-adjustment to modulate the influence of various weights.

3.1.2. User-Item Eigen Decomposition

The original NMF algorithm only considers the decomposition of static data, but in reality, the user's preference may keep changing over time, so this paper considers each user's latent factor dynamic change characteristics for improvement.

Initially, an auxiliary matrix is formulated to establish connections between the latent factors of users who have transitioned across various periods. Subsequently, temporal weights are introduced to accommodate the diminishing impact of factor matrix weights over time. This aims to capture the temporal changes in user preferences more accurately and enhance the precision of predicting evolving dynamic preferences. Moreover, we explore integrating an influence factor that signifies the temporal effect on a user's characteristic factor matrix at different time points, contingent upon the duration of the time interval T up to moment t . Building upon the concept above and integrating it with the

decomposition paradigm by Lee [32], the formula for temporal drift matrix decomposition is defined as presented in Equation (7):

$$\begin{cases} R^{(t)} \approx U_R^{(t)} V_R^{(t)} \\ R^{(t)} \approx U_R^{(t-T)} D_R^{(t-T)} V_R^{(t)} \end{cases} \quad (7)$$

The dimension τ is set to 2, meaning that time is divided into two parts: the current moment t and the past moment $t - T$. The transformation matrix D represents the relationship between these two moments, showing how the current matrix U_t is transformed into the past matrix U_{t-T} .

4. Methodology

4.1. Temporal Latent Rating Matrix Prediction Model

The objective function is framed as a minimization problem to minimize the variance between rating prediction matrices and $R^{(t)}$. This approach assesses the efficacy of the dimensionality reduction transformation applied to the rating matrix, as depicted in Equations (8) and (9):

$$\arg \min_{U_R^{(t)}, V_R^{(t)}} \|R^{(t)} - U_R^{(t)} V_R^{(t)}\|_F^2 \quad (8)$$

$$\arg \min_{U_R^{(t-T)}, V_R^{(t)}} \|R^{(t-T)} - U_R^{(t-T)} V_R^{(t)}\|_F^2 \quad (9)$$

Through the correlation of Equations (7)–(9), while considering the duration of the time interval up to time T , we can formulate the objective minimization relationship as expressed in Equation (10):

$$\begin{aligned} \arg \min_{U, Q, V} & \|R^{(t)} - \gamma U_R^{(t)} V_R^{(t)}\|_F^2 + \|R^{(t)} - (1 - \gamma) U_R^{(t-T)} D_R^{(t-T)} V_R^{(t)}\|_F^2 \\ \text{s.t. } & U_R^{(t)}, U_R^{(t-T)}, D_R^{(t-T)}, V_R^{(t)} \geq 0 \end{aligned} \quad (10)$$

The $\gamma (0 < \gamma < 1)$ represents the time regularisation factor, accounting for varying time influences on user preferences.

4.2. The Latent Topic Model

To delve deeper into unexplored user preference topics using the rating model, this section introduces a latent topic model to describe the potential association between user reviews and topic words. This model will be used to construct a joint optimization model for rating and topic decomposition in the section.

4.2.1. Topic Corpus Processing

In this study, the topic words refer to all the terms that accurately reflect the content of the review topic. These words are derived from the review collection, differentiating them from general terms. Unlike narrative words, which are manually cited, these topic headings are extracted directly from the original review content. This approach is beneficial for automatically identifying the validity of the topic, making it particularly useful for processing large amounts of review data. Unlike keywords, which only include words that appear in the reviews, topic words also encompass other relevant terms that may not be explicitly mentioned in the reviews. Consequently, there has been a notable increase in the quantity of topic words, offering a more extensive coverage within the review. In addition,

to stop words such as “of,” “a,” “he,” etc., and some symbols, we use the NLTK tool to remove them from the text.

4.2.2. Definition of the Basic Concept

The chosen topic words may reflect a connection to the review topic, but it is complex. On the one hand, it is impossible to cover all relevant topics in a review due to the existence of synonyms and differences in users’ interests in word choice. On the other hand, topic words that do not appear in the reviews may also reflect the review topic. Therefore, we refer to the topic words that appear in the review content as explicit topic words and other words that do not appear in the current review content as latent topic words. Latent topic words related to the current review topic are called implicitly relevant topic words, while latent topic words unrelated to the review topic are called implicitly irrelevant topic words. To ensure that the keywords fully represent the content of the comments, then we adopted the following pre-process. First, we trained a Word2Vec model using the entire dataset of comment corpora to obtain semantic vectors for words. Next, we calculated the TF-ICF values for all nouns and adjectives, selecting the top 20 words with the highest TF-ICF values in each comment as candidate explicit keywords. Finally, we manually sample and validate high-frequency keywords across multiple categories to ensure their semantic relevance. This process aims to minimize lexical selection bias and guarantee that the chosen words represent the core content of the reviews.

Here, we propose the notion of relevance probability to ascertain the likelihood that a latent topic word is implicitly associated with the reviews. A higher relevance probability indicates a greater likelihood that the latent topic word is implicitly related to the reviews.

First, the class term frequency-inverse class frequency (CTF-ICF) is applied to construct the weight matrix $G^{(t)} = g_{jk}$ for the review data at time t . It can address the issue of a topic word appearing frequently in a shorter text but not having a high word frequency. Simply considering topic word frequency is insufficient in determining a term’s importance. It considers the user’s historical preferences for the same item category by aggregating them into a longer text and then calculating the frequency of the topic word. If the term appears frequently in all texts within the category, it is considered more important for that category. Here, g_{jk} represents the weight of topic word k in review j . Next, a probability matrix $Z^{(t)} = z_{jk}$ is constructed, where all terms that have already been identified are considered positive samples and z_{jk} is set to 1. On the other hand, all latent topic words are considered negative samples and assigned a correlation probability of 0, as shown in Equation (11):

$$z_{jk} = \begin{cases} 0, & g_{jk} = \emptyset \\ 1, & g_{jk} \neq \emptyset \end{cases} \quad (11)$$

We have preference weight moments $P^{(t)} = P_{jk}$ based on the richness of reviews and the popularity of topic words at moment t . The corresponding P_{jk} is set to 1 for review topic words that already have a weight value. For latent topic words, their weights represent the confidence level of the related topic words. A higher weight indicates a higher probability that it is a related topic word.

In determining the weight assigned to latent topic words, Zhao et al. examined a non-uniform weighting method for missing items based on item popularity. They observed that items with more incredible popularity are more likely to be viewed by users. However, the absence of user interaction with these items may suggest a lack of interest [33]. Consequently, we propose Hypothesis 1: the absence of reviews for trendy topic words better signifies the insignificance of both reviews and those particular topic words. Hypothesis 1 primarily considers the impact of topic word popularity and does not account for the symmetric factor of review richness. Zhang et al. contend that more active users tend to

be less interested in items that do not generate interactions, as indicated in [34]. From this perspective, we propose Hypothesis 2: topic words absent from wealthy reviews are more likely to be irrelevant to the topic discussed in the review. ‘Richness’ refers to the extent of topic information within a review, wherein a higher abundance of topic words signifies more extraordinary richness.

4.2.3. The Objective Latent Topic Model

Based on the basic concept, we proposed a latent factor factorization model that utilizes the richness and popularity weights of review topic words to enhance the topic model further. This model aims to infer the relationship between reviews and latent topic words through objective function optimization.

According to Hypothesis 1 and Hypothesis 2, we posit that when the popularity or richness of words associated with a latent topic increases, the probability of these words being implicitly related to the review decreases. Consequently, the preference weight P_{jk} for such latent topic words can be adjusted to a smaller value.

Then we define the richness of review j , denoted as X_j , as the total number of topic words present in that review, i.e., $X_j = \sum_{k=1}^n z_{jk}$. Similarly, the popularity of a topic word k is represented by Y_k , which is the total number of times that word appears in different comments, i.e., $Y_k = \sum_{j=1}^l z_{jk}$. To normalize the values of X_j and Y_k , we use the Z-score normalization method, expressed in the following Equations (12) and (13):

$$\hat{X}_j = \frac{|X_j - \bar{X}|}{SD(X)} \quad (12)$$

$$\hat{Y}_k = \frac{|Y_k - \bar{Y}|}{SD(Y)} \quad (13)$$

Linear weighting of Equations (12) and (13), the final weighting equation is Equation (14)

$$P_{jk} = \begin{cases} \beta \frac{1}{e^{\hat{X}_j}} + (1 - \beta) \frac{1}{e^{\hat{Y}_k}}, & z_{jk} = 0 \\ 1, & z_{jk} = 1 \end{cases} \quad (14)$$

$\beta \in [0, 1]$ is a weighting factor to adjust the share of review richness and topic word popularity.

As illustrated in Figure 2, we use Equation (14) to construct the preference coefficient matrix $P^{(t)}$. Then, we utilize the correlation probability matrix $Z^{(t)}$ and the preference weight matrix $P^{(t)}$ to perform a weighted matrix decomposition of the associated probability matrix $Z^{(t)}$. This decomposition results in two low-dimensional feature matrices, $M^{(t)}$ and $Q^{(t)}$, where $M^{(t)}$ represents the topic word correlation latent factor matrix and $Q^{(t)}$ represents the topic word correlation probability confidence latent factor matrix. Based on our previous hypotheses, we apply the preference weights $P^{(t)}$ to the word relevance latent factor matrices to ensure accuracy. We then calculate the loss function Θ using alternating least square error (ALS) as shown in Equation (15):

$$\begin{aligned} \operatorname{argmin} \Theta &= \sum_{jk} P_{jk}^{(t)} \left\| Z_{jk}^{(t)} - \hat{Z}_{jk}^{(t)} \right\|_F^2 \\ &= \sum_{jk} P_{jk}^{(t)} \left\| Z^{(t)} - M_j^{(t)} Q_k^{(t)T} \right\|_F^2 \end{aligned} \quad (15)$$

The $\tilde{P}^{(t)}$ denotes a diagonal matrix composed of the corresponding elements in $P^{(t)}$. To prevent overfitting, we add regularisation terms to the loss function Θ , and the final loss function Θ is defined as Equation (16)

$$\begin{aligned} \operatorname{argmin} \Theta = & \sum_{jk} P_{jk}^{(t)} \left\| Z_{j,k}^{(t)} - M_j^{(t)} Q_k^{(t)T} \right\|_F^2 \\ & + \sigma \left(\left\| M_j^{(t)} \right\|_F^2 + \left\| Q_k^{(t)} \right\|_F^2 \right) \end{aligned} \quad (16)$$

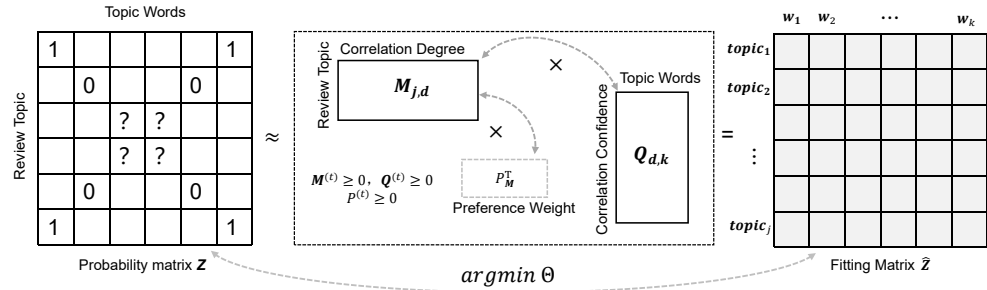


Figure 2. The latent topic words model diagram.

4.2.4. The Optimization of Topic Model

Firstly, the process of taking partial derivatives of $M^{(t)}$ is as follows: first, take the derivative of a single element as shown in Equation (17):

$$\frac{1}{2} \frac{\partial (M^{(t)}, Q^{(t)})}{\partial M_{jm}^{(t)}} = \sum_k P_{jk}^{(t)} \left(M_j^{(t)} Q_k^{(t)T} - Z_{j,k}^{(t)} \right) Q_{km}^{(t)} + \sigma \left(\sum_k P_{jk}^{(t)} \right) M_{jm}^{(t)} \quad (17)$$

The $(0 \leq m \leq d)$, then derive the vector element as Equation (18):

$$\begin{aligned} \frac{1}{2} \frac{\partial (M^{(t)}, Q^{(t)})}{\partial M_j^{(t)}} &= \frac{1}{2} \left(\frac{\partial (M^{(t)}, Q^{(t)})}{\partial M_{j1}^{(t)}}, \dots, \frac{\partial (M^{(t)}, Q^{(t)})}{\partial M_{jd}^{(t)}} \right) \\ &= M_j^{(t)} \left(Q^{(t)T} \tilde{P}_j^{(t)} Q^{(t)} + \sigma \left(\sum_k P_{jk}^{(t)} \right) I \right) - Z_j^{(t)} \tilde{P}_j^{(t)} Q^{(t)} \end{aligned} \quad (18)$$

The I is the unit matrix. Taking partial derivatives of $M^{(t)}$ and derive $\frac{\partial \Theta(M, Q)}{\partial M} = 0$ to obtain the updated $M^{(t)}$, which is calculated as Equation (19)

$$M^{(t)} \leftarrow M^{(t)} \odot \frac{Z_j^{(t)T} \tilde{P}_j^{(t)} Q^{(t)}}{Q^{(t)T} \tilde{P}_j^{(t)} Q^{(t)} + \sigma \left(\sum_k P_{jk}^{(t)} \right) I} \quad (19)$$

Likewise, taking partial derivatives of $Q^{(t)}$, and derive $\frac{\partial \Theta(M, Q)}{\partial Q} = 0$ to obtain the updated $Q^{(t)}$, which is calculated as Equation (20)

$$Q^{(t)} \leftarrow Q^{(t)} \odot \frac{Z_k^{(t)T} \tilde{P}_j^{(t)} M^{(t)}}{M^{(t)T} \tilde{P}_j^{(t)} M^{(t)} + \sigma \left(\sum_j P_{jk}^{(t)} \right) I} \quad (20)$$

The matrices M and Q , which represent latent factors, undergo iterative updates using Equations (19) and (20) until the loss function reaches a local optimum. At this stage, we define $\hat{Z} = MQ^T$, where each element $\hat{Z}_{jk}$ in the matrix $\hat{Z}$ signifies the probability that review j , when weighted, is associated with topic word k . Our objective is to approximate

$\hat{Z}$ to Z , wherein a higher $\hat{Z}_{jk}$ value indicates a greater likelihood of the topic word k being relevant to review j .

4.3. Joint Optimization Loss Function

The paper aims to transform the association between user rating and latent topics into a task of finding a joint optimal solution. Optimizing the regularisation factor Z allows the model to extract topic word features from users' various topics over time. This process enables the reflection of the evolving relationship between each latent review topic associated with user i and item j ratings. So we give the objective functions in Equations (10) and (16) as Equation (21):

$$\begin{aligned} \arg \min \mathcal{L} = & \left\| R^{(t)} - \gamma U_R^{(t)} V_R^{(t)} \right\|_F^2 \\ & + \left\| R^{(t)} - (1 - \gamma) U_R^{(t-T)} D_R^{(t-T)} V_R^{(t)} \right\|_F^2 \\ & + \sigma_1 \left\| D_R^{(t-T)} - I \right\|_F^2 \\ & + \sum_{jk} P_{jk}^{(t)} \left\| Z^{(t)} - M_j^{(t)} Q_k^{(t)T} \right\|_F^2 \\ & + \sigma_2 \left(\left\| U_R^{(t)} \right\|_1 + \left\| U_R^{(t-T)} \right\|_1 + \left\| V_R^{(t)} \right\|_1 + \left\| M_j^{(t)} \right\|_F^2 + \left\| Q_k^{(t)} \right\|_F^2 \right) \end{aligned} \quad (21)$$

In Equation (21), the function indicates that the error between the predicted and actual data is minimized. The bias level of $U^{(t-T)}$ is governed by the time regularization term $\sigma_1 \|D_R^{(t-T)} - I\|_F^2$. Here, I represents the identity matrix, while σ_1 and σ_2 are regularisation parameters. A higher σ_1 value implies that the model $U_R^{(t-T)}$ is more sensitive to the impact of the auxiliary matrix $D_R^{(t-T)}$ on its influence. The model's structure is depicted in Figure 3.

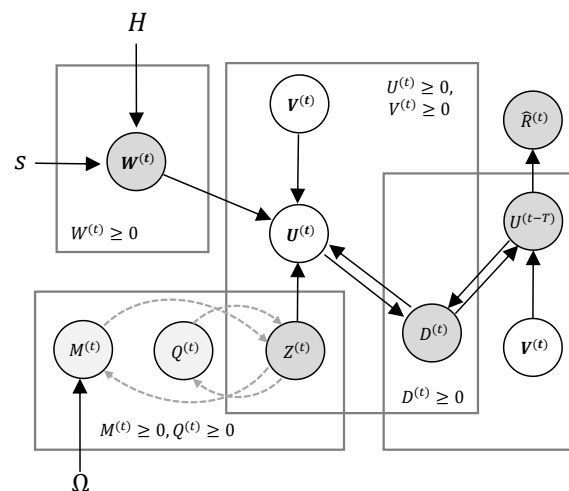


Figure 3. Graphical illustration of the proposed TDLRP-MF model.

For the objective function, the optimization problem is solved using the first-order necessary KKT condition equation constraints [35]. Especially, each rating matrix R_i^t of user i is updated according to the temporal weight of Equation (6) with Equation (22):

$$R_i^{(t)} \leftarrow W^{(t-T)} \odot R_i^{(t)} \quad (22)$$

The ultimate goal of this model is to ascertain the optimal solution for the parameters U , V , D , M , and Q , respectively, to link the user rating with the review information. The algorithm is presented in Algorithm 1.

Algorithm 1: The proposed algorithm

Input: User-item rating matrices $R^{(t)}, R^{(t-T)}$;
 Topic words initialized with NLTK;
 Regularization parameters $\{\sigma_1, \sigma_2\}$;
 $\maxIter, \maxSubIter$

Output: Predicted Matrix $\hat{R}^{(t)}$

Initialization:
 Give initial non-negative values to $U_R^{(t)} \in \mathbb{R}^{i \times \tau}, V_R^{(t)} \in \mathbb{R}^{\tau \times j}$,
 $D^{(t-T)} \leftarrow I, \forall T = 1, 2, \dots, T_{max}$;
 $G^{(t)} \leftarrow$ Review CTF-ICF Weight Matrix;
 Compute $U_R^{(t-T)}$ and $V_R^{(t)}$ by applying on $R^{(t-T)}$;
while $\neg \argmin \mathcal{L} \vee iter < \maxIter$ **do**
 Update $R_i^{(t)}$ with Equation (6);
 Update $U_R^{(t)}, V_R^{(t)}$ and $D^{(t-T)}$ with Multiplicative Update Rules,
 $\forall T = 1, 2, \dots, T_{max}$;
 if $Iter = \maxSubIter$ **then**
 Compute $Z^{(t)} \leftarrow \begin{cases} 0, & g_{jk} = \emptyset \\ 1, & g_{jk} \neq \emptyset \end{cases}$;
 Normalization $\hat{X}_k$ and $\hat{Y}_k$ with Z-score;
 Compute $P^{(t)}$ with Equation (14);
 while $\neg \argmin \Theta \vee SubIter < \maxSubIter$ **do**
 Update $M^{(t)}$ based on ALS with Equation (19);
 Update $Q^{(t)}$ based on ALS with Equation (20);
 $SubIter = SubIter + 1$
 end
 end
 Compute $\argmin \mathcal{L}$ with Equation (21);
 $Iter = Iter + 1$
end
 $\hat{R}^{(t)} = U_R^{(t)} V_R^{(t)}$

4.4. Computational Complexity

The model's primary considerations involve the computational complexity of the objective function and topic decomposition. Regarding the objective function $\mathcal{L}$, the algorithm's maximum complexity during iteration is $O(Iij\tau|T|)$, where I represents the iterations needed for matrix decomposition, and $|T|$ signifies the total time dimension. The maximum time complexity for the factorization iterations of the latent topics in Equation (14) is $O(Ijk)$. Consequently, the overall complexity is denoted as $T(n) = O(Iij\tau|T| + Ijk)$. As τ remains a constant term in this context, its complexity is $O(1)$. Thus, the maximum time complexity of the model algorithm stands at $O(Iij|T|)$.

5. Experiment and Analysis

5.1. Datasets

The experiment uses six datasets to assess the model's performance. These datasets are sourced from the Amazon compiled by Stanford University [36]. Each dataset comprises

user ratings, item details, associated user reviews, temporal information on review behavior, and rating values ranging from integers 1 to 5. The Table 2 displays the sizes of each dataset.

Table 2. Datasets description.

Datasets	Users	Items	Ratings & Reviews	Sparsity	Time Range
Kindle Store	346,457	29,357	1,767,612	99.9826%	April 1999–October 2018
Books	578,858	61,379	4,159,480	99.9883%	December 1997–October 2018
Magazine Subscriptions	438,645	37,250	2,638,219	99.9839%	March 1999–October 2019
Electronics	9,838,676	786,868	20,994,353	99.9997%	December 1997–October 2018
Home&Kitchen	9,767,606	1,301,225	21,928,568	99.9998%	November 1999–October 2018
Toys&Games	4,204,994	634,414	8,201,231	99.9997%	October 1999–October 2018

5.2. Evaluation Indicators

In this paper, we utilized four evaluation metrics, precision, recall, NDCG, and F1, for our experiments. The objective of item recommendation involves generating Top-N item recommendation lists for the specified users in the test dataset. The computations for these metrics are depicted in Equations (23)–(27). The F1 metric's calculation can be performed using Equation (23).

$$F1 = \frac{2 \times \text{Precision (N)} \times \text{Recall (N)}}{\text{Precision (N)} + \text{Recall (N)}} \quad (23)$$

where precision and recall are defined as shown in Equations (24) and (25):

$$\text{Precision (N)} = \frac{|\text{Top (N)} \cap \text{BL}|}{N} \quad (24)$$

$$\text{Recall (N)} = \frac{|\text{Top (N)} \cap \text{BL}|}{\text{BL}} \quad (25)$$

where BL represents the user's purchase list, Top(N) represents the top N recommended item list, and NDCG is a measure used to evaluate the quality of recommendation results within the top N items. The formula for NDCG@N is expressed in Equation (26):

$$\text{NDCG@N} = \frac{1}{\text{IDCG@N}} \sum_{i=1}^N \frac{2^{k_i} - 1}{\log_2 i + 1} \quad (26)$$

Here, k_i represents whether item i is present in the recommended result set (1 if true, 0 if false). The formula for IDCG@N is defined as presented in Equation (27):

$$\text{IDCG@N} = \sum_{i=1}^{\text{RLT}} \frac{2^{k_i} - 1}{\log_2 i + 1} \quad (27)$$

The RLT in Equation (27) denotes the collection of the most pertinent outcomes within the suggested results under optimal conditions. It refers to the assortment comprising the top N results arranged in descending order of recommendation relevance. Hence, as the relevance of the recommended results increases, so does the NDCG value.

5.3. Experimental Parameter Setting

We compare the proposed TDLRP-MF with baseline models such as TPNE [18], TCMF [11], TMRevCo [12], TDADLFM [19], and MTUPD [28] which are described in Section 2.

In the experiment, we use the state-of-the-art parameters of the corresponding model for comparison. The TPNE model's head number of layers is set as 4 and the number of Transformer layers is set as 2. For the TCMF model experimental parameters, we set $\lambda = \beta = \gamma = 1$, $\epsilon = 10^{-4}$, and $\text{maxIter} = 1000$. The TMRevCo model parameter settings are $\lambda_c = 0.001$ and $\lambda_r = 0.05$. For the TDADLFM model, we set the history size, number of epochs, and batch size as 30, 20, and 1500, respectively. The parameter settings for the MTUPD model use $\alpha = 1$, $K = 20$, $\lambda_1 = 0.5$, and $\lambda_2 = 0.5$. To ensure training stability and convergence, all latent factor matrices of the proposed model were initialized using the Xavier method to prevent gradient vanishing or explosion. Then the optimization process employs Alternating Least Squares (ALS), which effectively enhances training stability by decomposing non-convex problems into a series of convex subproblems. All experiments were completed within 48 h on a server equipped with a 14,900 CPU and a 4090 GPU.

5.4. Parameter Analysis

In the parametric experiments, we configured the maximum number of iterations (maxIter) to 10^3 and the maximum number of sub-iterations (subMaxIter) to 10^2 . Additionally, we set the decomposition dimension to 20. We employed the random initialization method to initialize the weight matrix a^{wv} and the bias vector b^{wv} . We updated the weight and bias parameters using gradient descent. Based on this setup, we analysed the model's required parameters through experimental investigation, namely α , β , γ , σ_1 , and σ_2 .

This study examined the impact of the review topic hotness and richness regulation parameter β . We varied β between 0.05 and 0.95, with a time fragment of $T = 12$ months and a step size of 0.2. These experiments were performed across different datasets and evaluated using three metrics: Precision, F1, and NDCG. The results, depicted in Figure 4, showcase the relationship between the model's β values and changes in F1 performance. Notably, we observed that the model's performance improves as β increases, reaching its peak at $\beta = 0.55$. However, beyond $\beta = 0.55$, the accuracy starts to decline. This observation suggests that tremendous and small values of β can harm the accuracy of predicting user topic preferences. Therefore, we identify $\beta = 0.55$ as the optimal regulation parameter for review topic hotness and richness in this study.

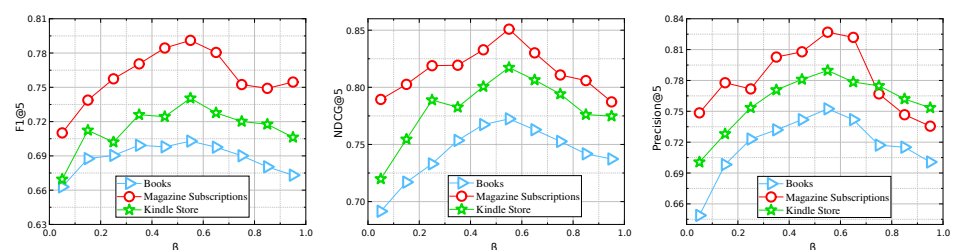


Figure 4. The impact of topic popularity and richness parameter β on model performance.

The impact of regularisation parameters σ_1 and σ_2 on model performance was investigated in our experiments. These parameters were varied in the range of $[0, +\infty)$, and the optional values are $\{0.01, 0.1, 0.2, 0.5, 1, 5, 10, 20\}$ for the Books dataset. In Figure 5, the horizontal axis represents the values of the regularisation parameters, and the vertical axis shows the changes in model NDCG performance. As the values of σ_1 and σ_2 increase, the model performance also improves. The highest overall performance is achieved when the values of σ_1 and σ_2 are set to 0.5, after which the accuracy decreases. This suggests

that a too-small value of the regularisation parameter leads to underfitting. At the same time, a too-large value leads to overfitting, resulting in a decrease in model accuracy and an inability to predict user preferences accurately. Therefore, we have chosen a value of 0.5 for the regularisation parameters σ_1 and σ_2 in this paper.

We investigated the impact of the time-weight balance factor, represented by alpha (α), and the time-evolution balance factor, represented by gamma (γ), on the model's performance. Sections 3 and 4 provided the definitions for α and γ to determine time segments ranging from 24 to 72 months. We adjusted the values of α and γ within the range of 0 to 1 in increments of 0.25 to assess their effects. The results, depicted in Figure 6, show the model's precision index across different values. Lighter colours in the graph denote a higher precision index, while darker shades indicate lower values. Our findings reveal that the model's performance improves with increasing α and γ values. The accuracy declines when α surpasses 0.5 and γ exceeds roughly 0.75. This suggests that extreme values of α or γ inadequately express the influence of forgetting weight and review evolution weight on users' preferences or ignore the crucial evolutionary relationship between users' current and past preferences, leading to reduced accuracy. Consequently, based on our experiments, we identified optimal balance factors of $\alpha = 0.5$ and $\beta = 0.75$ for this model. These outcomes underscore the significance of comprehensively incorporating the forgetting law and considering the combined impact of users' current and historical preferences on recommendations to enhance outcomes effectively.

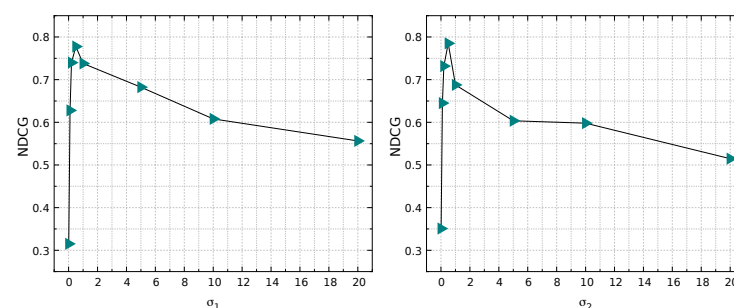


Figure 5. The impact of different values of regularization parameter σ_1, σ_2 on the model.

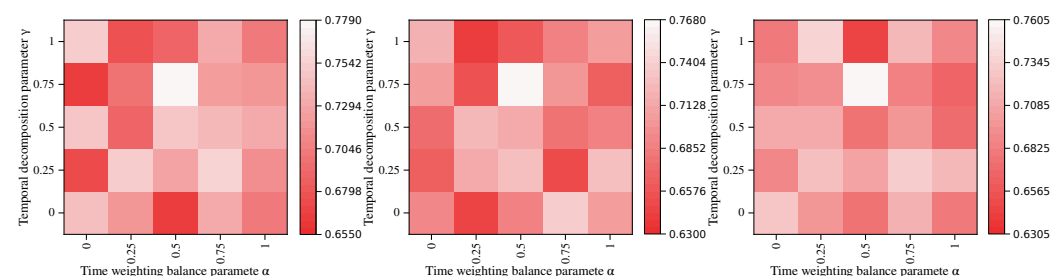


Figure 6. The impact of time weight adjustment parameter α, γ on the model. (Top-5 recommendation, $T = 24, 48, 72$ respectively).

To validate the robustness of the training process, we conducted 10 additional independent experiments on the Books dataset using different random seeds. The results show that the NDCG@5 mean metric on the test set is 0.7941 with a standard deviation of 0.0023 (relative error $< 0.3\%$). This indicates that despite different starting points, the model converges stably to locally optimal points with comparable performance. Additionally, we implemented an early stopping mechanism that terminates training if the validation set loss fails to decrease for 10 consecutive epochs. This effectively enhances training efficiency and prevents overfitting.

5.5. Ablation Results

To delve deeper into the impact of each component's enhancement on performance, we partitioned the ablation model into three categories: experiments for ablation of the topic model, experiments isolating the temporal evolution factor, and experiments altering topic weights. These ablation experiments were conducted across each dataset, and their results are detailed in Table 3.

Table 3. Ablation results.

Object	Books				Kindle Store				Magazine Subscriptions			
	NDCG		F1		NDCG		F1		NDCG		F1	
	@Top5	@Top10	@Top5	@Top10	@Top5	@Top10	@Top5	@Top10	@Top5	@Top10	@Top5	@Top10
+NMF	0.6878	0.6825	0.6392	0.6160	0.7353	0.7163	0.6936	0.6709	0.7168	0.7095	0.6508	0.6361
+LSA	0.6994	0.6873	0.6419	0.6313	0.7147	0.7237	0.6841	0.6450	0.7031	0.6931	0.6387	0.6345
+pLSA	0.7337	0.7190	0.6418	0.6255	0.7459	0.7179	0.6709	0.6598	0.7258	0.7163	0.6535	0.6435
+LDA	0.7158	0.7058	0.6461	0.6102	0.7327	0.7396	0.6794	0.6535	0.7237	0.7174	0.6567	0.6472
TDW1 _{ff=1,fl=0}	0.6979	0.6910	0.6609	0.6687	0.7549 †	0.7543	0.6883	0.6856	0.7201	0.7273	0.6540	0.6544
TDW2 _{ff=0,fl=1}	0.6931	0.6946 †	0.6524	0.6571	0.7417	0.7427	0.6741	0.6750	0.7132	0.7179	0.6582	0.6598
TDW3 _{ff=fl=0}	0.6678	0.6683	0.6123	0.6133	0.7380	0.7390	0.6841	0.6888	0.6989	0.7039	0.6371	0.6435
TDW4 _{ff=fl=1}	0.6762	0.6762	0.6208	0.6212	0.7496	0.7464	0.6709	0.6793	0.7168	0.7126	0.6514	0.6592
LRP1 _{fi=0}	0.7026	0.7039	0.6340	0.6392	0.7649 ‡	0.7637	0.6829	0.6877	0.7307	0.7315	0.6614	0.6621
LRP2 _{fi=1}	0.6989	0.6968	0.6376	0.6355	0.7607	0.7664	0.6989 ‡	0.6909	0.7205	0.7248	0.6530	0.6613
TDLRP _{opt}	0.7807	0.7833	0.7047	0.7083	0.8515	0.8509	0.8029	0.8036	0.8193	0.8197	0.7913	0.7942

Four topic models were compared in the initial experiments with those utilized in this paper, as depicted in Table 3. LSA and pLSA employ matrix decomposition to create matrices for document topics and topic words. A Bayesian statistics-based LDA generates distributions for document topics and topic words. NMF is another matrix decomposition-based method, that forms matrices for document and word topics. The approach adopted in this paper incorporates non-negative decomposition, accounting for temporal evolution, increased emphasis on topic popularity, and richness to enhance recommendation effectiveness. Bold results indicate that while combining these four ablated topic models attained the highest performance within their respective categories based on F1 and NDCG metrics across various datasets (identified as the best metric among these experiments), the overall performance still fell short compared to the final optimized model *TDLRP_{opt}* presented in this paper. The proposed model can mine the popular topics in user comments in different periods and capture the dynamic evolution law of user topic preferences through the potential features of popularity and abundance of topic terms to improve the effectiveness of recommendations.

In the second group of TDW ablation experiments, we analyzed four scenarios using the temporal weights defined by Equations (8) and (12). These scenarios were TDW1, TDW2, TDW3, and TDW4. In TDW1, the factor values were set to $\alpha = 1$ and $\gamma = 0$, which means that the temporal weight only considers the user forgetting weight and the interest preference of the past period $t - T$. In TDW2, the factor values were set to $\alpha = 0$ and $\gamma = 1$, which means that the temporal weight only considers the user review weight and the interest preference of the current period t . In TDW3, the factor values were set to $\alpha = 0$ and $\gamma = 0$, which means that the temporal weight only considers the user review weight and the interest preference of the period t . Finally, in TDW4, the factor values were set to $\alpha = 1$ and $\gamma = 1$, which means that the temporal weight only considers the user forgetting weight and the interest preference of the current period t . In Table 3, the symbol † indicates the best performance of the NDCG indicator, 0.7549, and the best performance

of the F1 indicator for TDW2, 0.6946. However, in each data test, the indicator results were lower than the final model $TDLRP_{opt}$, indicating that only by fully considering the temporal weights can we better reflect the temporal pattern of user interests and improve the recommendation effectiveness.

The third group conducted ablation experiments regarding weight LRP in their study. These experiments were split into two scenarios: LRP1, where $\beta = 0$ and the weight solely accounts for popularity, and LRP2, where $\beta = 1$ and the weight solely considers richness. The outcomes are detailed in Table 3. Across each dataset, the symbol ‡ highlights the highest performance of the NDCG indicator at 0.7649, and for TDW2, the best F1 indicator performance reached 0.6989. However, these dataset-specific results fell short compared to the indicator outcomes of the ultimate model $TDLRP_{opt}$. The result suggests that neglecting temporal topic weighting may lead to the loss of crucial user preference information, potentially impacting the accuracy of recommendation outcomes.

The enhanced model in this study comprehensively incorporates multiple users' temporal interests, showcasing the most vital capability in capturing the evolving aspects of users' diverse interests. This enhancement significantly contributes to improving the effectiveness of recommendations.

5.6. Performance Analysis and Case Study

Based on the optimized parameters mentioned earlier, we expanded the number of recommendations from 10 to 50 in increments of 10. The accuracy evaluation for Top@N recommendations across various datasets and recommendation methods is depicted in Figure 7. Our findings illustrate the superiority of the TDLRP-MF model over other compared temporal models. This underscores its exceptional capacity to accommodate users' evolving temporal preferences.

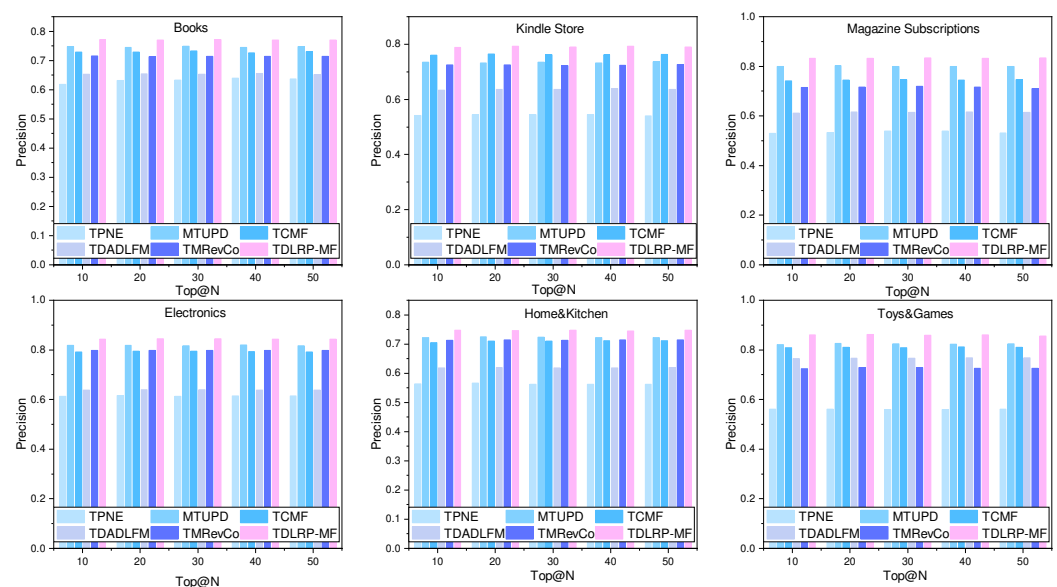


Figure 7. Results of precision comparison.

Table 4 presents a more in-depth dataset analysis detailing each model's performance metrics (where Top@N = 5 or @N = 10). Table 4 compares these metrics across models, revealing that the TDLRP-MF model outperforms other baseline methods across all three metrics. To validate the statistical significance of the performance improvement achieved by the TDLRP-MF model, we conducted five independent experiments using different random seeds and performed paired t-tests on the NDCG@5 results against the strongest baseline model MTUPD. The results indicate that across all six datasets, the p-values

for improvements achieved by TDLRP-MF relative to MTUPD were consistently below 0.01. This strongly demonstrates that the performance gains obtained by our model are statistically significant. The result emphasizes our model’s effectiveness in capturing users’ dynamic preferences over time, considering user forgetting patterns and review evolution, thereby capturing temporal interest shifts. Consequently, our model significantly enhances recommendation accuracy compared to other temporal approaches.

Table 4. Performance evaluation on Top-N recommendation.

Comparison of Precision												
DataSets/Methods	TPNE		MTUPD		TCMF		TDADLFM		TMRevCo		TDLRP-MF	
	P@5	P@10	P@5	P@10	P@5	P@10	P@5	P@10	P@5	P@10	P@5	P@10
Books	0.5597	0.5397	0.7435	0.7218	0.7266	0.7098	0.6491	0.6167	0.7173	0.6946	0.7791	0.7652
Kindle Store	0.5471	0.5459	0.7371	0.7183	0.7627	0.7483	0.6373	0.6243	0.7134	0.6941	0.7986	0.7659
Magazine Subscriptions	0.5399	0.5334	0.7929	0.7834	0.7472	0.7574	0.6144	0.6038	0.7146	0.7374	0.8315	0.8191
Electronics	0.6126	0.5982	0.8172	0.7678	0.7824	0.7501	0.6385	0.6117	0.8081	0.7910	0.8437	0.8321
Home&Kitchen	0.5617	0.5636	0.7215	0.7156	0.7512	0.7390	0.6137	0.6105	0.7124	0.6917	0.7461	0.7435
Toys&Games	0.5631	0.5657	0.8261	0.8079	0.8021	0.7849	0.7631	0.7528	0.7224	0.6950	0.8594	0.8453
Comparison of F1												
DataSets/Methods	TPNE		MTUPD		TCMF		TDADLFM		TMRevCo		TDLRP-MF	
	F1@5	F1@10	F1@5	F1@10	F1@5	F1@10	F1@5	F1@10	F1@5	F1@10	F1@5	F1@10
Books	0.5811	0.5805	0.7058	0.6988	0.6960	0.6946	0.6581	0.6582	0.6698	0.6375	0.7232	0.7128
Kindle Store	0.5781	0.5766	0.7971	0.7754	0.7164	0.7052	0.7380	0.7553	0.6044	0.5972	0.8072	0.7984
Magazine Subscriptions	0.6744	0.6519	0.7175	0.7139	0.7113	0.7778	0.6537	0.6610	0.7024	0.6836	0.7361	0.7215
Electronics	0.6571	0.6581	0.7989	0.7992	0.7702	0.7667	0.7265	0.7341	0.7572	0.7336	0.8235	0.8196
Home&Kitchen	0.6309	0.6276	0.7057	0.6992	0.6987	0.6705	0.6726	0.6721	0.6421	0.6373	0.7238	0.7219
Toys&Games	0.6672	0.6513	0.7864	0.7727	0.7152	0.7079	0.7361	0.7174	0.5968	0.6008	0.8015	0.7952
Comparison of NDCG (N)												
DataSets/Methods	TPNE		MTUPD		TCMF		TDADLFM		TMRevCo		TDLRP-MF	
	N@5	N@10	N@5	N@10	N@5	N@10	N@5	N@10	N@5	N@10	N@5	N@10
Books	0.6536	0.6524	0.7721	0.7504	0.7428	0.7310	0.7467	0.7457	0.6865	0.6624	0.7984	0.7769
Kindle Store	0.6659	0.6238	0.8020	0.7946	0.8067	0.8329	0.6598	0.6679	0.7309	0.7185	0.8269	0.8139
Magazine Subscriptions	0.6369	0.6291	0.7317	0.7233	0.7544	0.7318	0.6545	0.6472	0.6928	0.6620	0.8376	0.8273
Electronics	0.6478	0.6393	0.8102	0.7983	0.8032	0.8056	0.6577	0.6416	0.7215	0.7089	0.8347	0.8131
Home&Kitchen	0.6196	0.6157	0.7008	0.7104	0.6972	0.6916	0.6212	0.6190	0.6554	0.6490	0.7492	0.7413
Toys&Games	0.6338	0.6315	0.8322	0.8249	0.7993	0.7805	0.6957	0.6931	0.6679	0.6505	0.8613	0.8504

Moreover, we compared NDCG@5 among different time-aware models at every period, as illustrated in Figure 8. Compared to the baseline model, our model can stay ahead of the curve at each time step, meanwhile, as the time step increases, our model can remain stable in terms of performance. This result demonstrates that leveraging user ratings and item review evolution not only effectively enhances recommendation performance, but also can learn the evolutionary pattern of user preferences through a history of longer time steps to improve the perception of sparse data so that the predicted items are close to the user’s real preferences.

We selected a real anonymized user with long-term active behavior in the Electronics category to qualitatively demonstrate how TDLRP-MF tracks their interest drift. Phase 1 (2016): The user exhibited strong long-term memory for items (slower decay of forgetting weights), with review keywords centered on “battery life” and “good price.” Phase 2 (2018): The user’s interests undergo a noticeable shift. The decay weight indicates accelerated forgetting of old interests, while the core comment keywords evolve to “camera quality” and “screen resolution.” Our model successfully captured this dynamic: In Phase 1, it recommended cost-effective devices with long battery life. By Phase 2, the recommendation list shifted to flagship models emphasizing camera and screen quality as core selling points. This case clearly demonstrates how the collaborative work of the forgetting mechanism

and topic evolution module enables the model to adapt to natural shifts in user interests, thereby delivering recommendations better aligned with their current preferences.

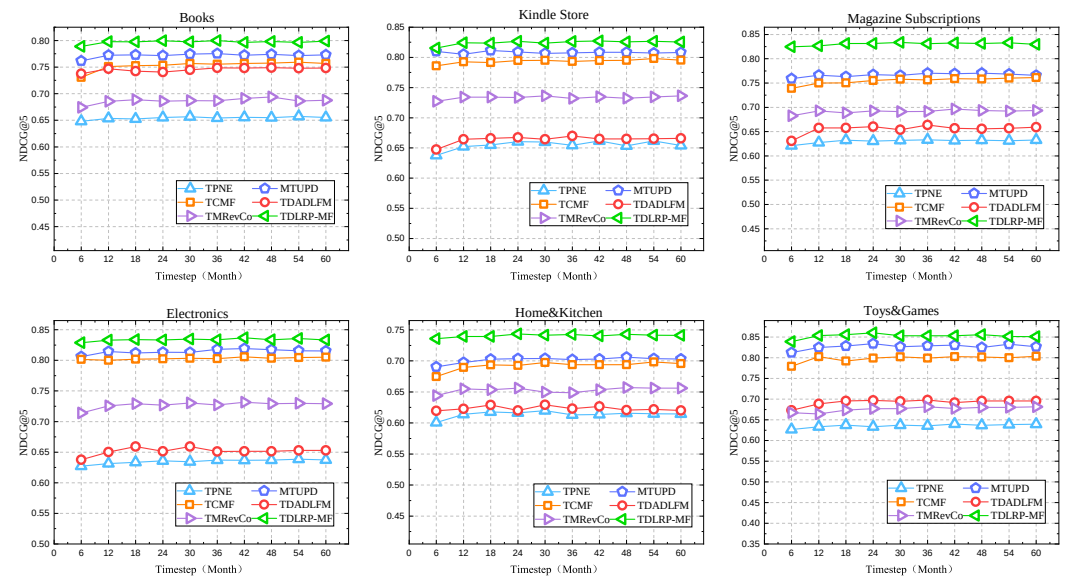


Figure 8. Comparison of NDCG@5 at each time-step.

6. Conclusions

The proposed model aims to seek the dynamic attributes of users' temporal preferences, encompassing aspects such as the fading of interests over time and the evolution of comments across various stages. This model not only tracks the shifting preferences of users throughout time but also accounts for the influence of comment evolution on user preferences. Consequently, it enables accurate, personalized recommendations from a substantial corpus of published information. The research reveals that current user preferences are shaped by their historical inclinations. Hence, considering popularity and depth, the model establishes a topic preference weight matrix for each user. Utilizing potential topic matrix decomposition, it captures the evolving interests of users in various subjects. Empirical findings demonstrate that this recommendation model surpasses existing baseline methods in effectively extracting user temporal characteristics. However, the study has several limitations that provide directions for future work, such as the forgetting function employed in the model using globally fixed parameters, failing to capture variations in interest decay rates across different users or item types.

In future work, one potential improvement area is focusing on developing an adaptive mechanism that learns personalized forgetting laws from users' historical interaction data. In addition, the introduction of large language models (LLMs) for user topic representation learning is also a new research avenue for the future.

Author Contributions: Conceptualization, H.D.; methodology, H.D.; software, H.D.; validation, H.D.; formal analysis, H.D.; investigation, H.D.; resources, H.D.; data curation, H.D.; writing—original draft preparation, H.D.; writing—review and editing, H.D., W.Z., G.H. and Z.B.; visualization, H.D.; supervision, H.D.; project administration, H.D.; funding acquisition, H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by China Postdoctoral Science Foundation grant number 2025M773202.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The online survey data have been presented in this paper and the others are securely protected by the researchers.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, X.; Liu, C. Design of Personalized News Recommendation System Based on an Improved User Collaborative Filtering Algorithm. *Mob. Inf. Syst.* **2023**, *2023*, 9898337. [\[CrossRef\]](#)
2. Wang, Y.; Wu, J.; Wu, Z.; Adomavicius, G. Efficient and flexible long-tail recommendation using cosine patterns. *INFORMS J. Comput.* **2025**, *37*, 446–464. [\[CrossRef\]](#)
3. Jain, G.; Mahara, T.; Sharma, S.C. Effective time context based collaborative filtering recommender system inspired by Gower's coefficient. *Int. J. Syst. Assur. Eng. Manag.* **2023**, *14*, 429–447. [\[CrossRef\]](#)
4. Tanveer, J.; Lee, S.W.; Rahmani, A.M.; Aurangzeb, K.; Alam, M.; Zare, G.; Alamdari, P.M.; Hosseinzadeh, M. PGA-DRL: Progressive graph attention-based deep reinforcement learning for recommender systems. *Inf. Fusion* **2025**, *121*, 103167. [\[CrossRef\]](#)
5. Parmezan, A.R.S.; Souza, V.M.; Batista, G.E. Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. *Inf. Sci.* **2019**, *484*, 302–337. [\[CrossRef\]](#)
6. Pang, N. A Personalized Recommendation Algorithm for Semantic Classification of New Book Recommendation Services for University Libraries. *Math. Probl. Eng.* **2022**, *2022*, 8740207. [\[CrossRef\]](#)
7. Wang, X.J.; Liu, T.; Fan, W. TGVx: Dynamic personalized POI deep recommendation model. *INFORMS J. Comput.* **2023**, *35*, 786–796. [\[CrossRef\]](#)
8. Alabduljabbar, R.; Alshareef, M.; Alshareef, N. Time-Aware Recommender Systems: A Comprehensive Survey and Quantitative Assessment of Literature. *IEEE Access* **2023**, *11*, 45586–45604. [\[CrossRef\]](#)
9. Koren, Y. Collaborative filtering with temporal dynamics. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 447–456.
10. Xiong, L.; Chen, X.; Huang, T.K.; Schneider, J.; Carbonell, J.G. Temporal collaborative filtering with bayesian probabilistic tensor factorization. In Proceedings of the 2010 SIAM International Conference on Data Mining, Columbus, OH, USA, 29 April–1 May 2010; pp. 211–222.
11. Rafailidis, D.; Kefalas, P.; Manolopoulos, Y. Preference dynamics with multimodal user-item interactions in social media recommendation. *Expert Syst. Appl.* **2017**, *74*, 11–18. [\[CrossRef\]](#)
12. Wu, T.; Feng, Y.; Sang, J.; Qiang, B.; Wang, Y. A novel recommendation algorithm incorporating temporal dynamics, reviews and item correlation. *IEICE Trans. Inf. Syst.* **2018**, *101*, 2027–2034. [\[CrossRef\]](#)
13. Li, T.; Jin, L.; Wu, Z.; Chen, Y. Combined recommendation algorithm based on improved similarity and forgetting curve. *Information* **2019**, *10*, 130. [\[CrossRef\]](#)
14. Pang, S.; Yu, S.; Li, G.; Qiao, S.; Wang, M. Time-Sensitive Collaborative Filtering Algorithm with Feature Stability. *Comput. Inform.* **2020**, *39*, 141–155. [\[CrossRef\]](#)
15. Lv, Y.; Zhuang, L.; Luo, P.; Li, H.; Zha, Z. Time-sensitive collaborative interest aware model for session-based recommendation. In Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME), London, UK, 6–10 July 2020; pp. 1–6.
16. José, E.F.; Enembreck, F.; Barddal, J.P. Adadrift: An adaptive learning technique for long-history stream-based recommender systems. In Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 11–14 October 2020; pp. 2593–2600.
17. Rabiou, I.; Salim, N.; Da'u, A.; Nasser, M. Modeling sentimental bias and temporal dynamics for adaptive deep recommendation system. *Expert Syst. Appl.* **2022**, *191*, 116262. [\[CrossRef\]](#)
18. Huang, C.; Wang, S.; Wang, X.; Yao, L. Modeling temporal positive and negative excitation for sequential recommendation. In Proceedings of the ACM Web Conference 2023, Austin, TX, USA, 30 April–4 May 2023; pp. 1252–1263.
19. Wu, C.C.; Chen, Y.L.; Yeh, Y.H. A Deep Recommendation Model Considering the Impact of Time and Individual Diversity. *IEEE Trans. Comput. Soc. Syst.* **2023**, *11*, 2558–2569. [\[CrossRef\]](#)
20. Sun, Y.T.; Guo, G.B.; Chen, X.; Zhang, P.H.; Wang, X.W. Exploiting review embedding and user attention for item recommendation. *Knowl. Inf. Syst.* **2020**, *62*, 3015–3038. [\[CrossRef\]](#)
21. Wang, X.; Xiao, T.; Tan, J.; Ouyang, D.; Shao, J. MRMRP: Multi-source review-based model for rating prediction. In Proceedings of the Database Systems for Advanced Applications: 25th International Conference, DASFAA 2020, Jeju, Republic of Korea, 24–27 September 2020; Proceedings, Part II 25; Springer: Berlin/Heidelberg, Germany, 2020; pp. 20–35.
22. Wu, S.; Zhang, Y.; Zhang, W.; Bian, K.; Cui, B. Enhanced review-based rating prediction by exploiting aside information and user influence. *Knowl.-Based Syst.* **2021**, *222*, 107015. [\[CrossRef\]](#)

23. Hyun, D.; Park, C.; Cho, J.; Yu, H. Learning to utilize auxiliary reviews for recommendation. *Inf. Sci.* **2021**, *545*, 595–607. [\[CrossRef\]](#)
24. Wang, P.P.; Li, L.; Wang, R.; Zheng, X.H.; He, J.X.; Xu, G.D. Learning persona-driven personalized sentimental representation for review-based recommendation. *Expert Syst. Appl.* **2022**, *203*, 117317. [\[CrossRef\]](#)
25. Liu, Z.; Yuan, B.X.; Ma, Y. A multi-task dual attention deep recommendation model using ratings and review helpfulness. *Appl. Intell.* **2022**, *52*, 5595–5607. [\[CrossRef\]](#)
26. Cai, Y.; Ke, W.M.; Cui, E.; Yu, F. A deep recommendation model of cross-grained sentiments of user reviews and ratings. *Inf. Process. Manag.* **2022**, *59*, 102842. [\[CrossRef\]](#)
27. Liu, H.; Chen, Y.; Li, P.; Zhao, P.; Wu, X. Enhancing review-based user representation on learned social graph for recommendation. *Knowl.-Based Syst.* **2023**, *266*, 110438. [\[CrossRef\]](#)
28. Wangwacharakul, C.; Wongthanasu, S. A novel temporal recommender system based on multiple transitions in user preference drift and topic review evolution. *Expert Syst. Appl.* **2021**, *185*, 115626. [\[CrossRef\]](#)
29. Milovančević, N.S.; Gračanac, A. Time and ontology for resource recommendation system. *Phys. A Stat. Mech. Its Appl.* **2019**, *525*, 752–760. [\[CrossRef\]](#)
30. Cheng, Y.; Qiu, G.; Bu, J.; Liu, K.; Han, Y.; Wang, C.; Chen, C. Model bloggers' interests based on forgetting mechanism. In Proceedings of the 17th International Conference on World Wide Web, Beijing, China, 21–25 April 2008; pp. 1129–1130.
31. Yu, Z.; Lian, J.; Mahmood, A.; Liu, G.; Xie, X. Adaptive User Modeling with Long and Short-Term Preferences for Personalized Recommendation. *IJCAI* **2019**, *7*, 4213–4219.
32. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [\[CrossRef\]](#)
33. Zhao, Z.; Chen, J.; Zhou, S.; He, X.; Cao, X.; Zhang, F.; Wu, W. Popularity bias is not always evil: Disentangling benign and harmful bias for recommendation. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 9920–9931. [\[CrossRef\]](#)
34. Zhang, Y.; Feng, F.; He, X.; Wei, T.; Song, C.; Ling, G.; Zhang, Y. Causal intervention for leveraging popularity bias in recommendation. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual, 11–15 July 2021; pp. 11–20.
35. Kuhn, H.W.; Tucker, A.W. Nonlinear programming. In *Traces and Emergence of Nonlinear Programming*; Springer: Basel, Switzerland, 2013; pp. 247–258.
36. McAuley, J.; Leskovec, J. Hidden factors and hidden topics: Understanding rating dimensions with review text. In Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, China, 12–16 October 2013; pp. 165–172.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.