

Article

Balancing Long–Short-Term User Preferences via Multilevel Sequential Patterns for Review-Aware Recommendation

Li Jin ^{1,†}, Xinzhe Li ^{1,†}, Suji Kim ¹ and Jaekyeong Kim ^{1,2,*}

¹ Department of Big Data Analytics, Kyung Hee University, 26, Kyungheedaero-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea; kimlyeo@khu.ac.kr (L.J.); lixz@khu.ac.kr (X.L.); waterland@khu.ac.kr (S.K.)

² School of Management, Kyung Hee University, 26, Kyungheedaero-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea

* Correspondence: jaek@khu.ac.kr

† These authors contributed equally to this study.

Abstract

Personalized recommender systems play an essential role in enhancing user experience by accurately predicting user preferences. Previous approaches mainly focus on modeling long-term preferences or capturing short-term dynamics through sequential patterns, while few achieve an effective balance between the two. This study proposes Rec-SSP, a novel review-aware recommendation model that integrates long-term and short-term preferences through a gated fusion mechanism. Long-term preferences are extracted from aggregated user reviews, whereas short-term preferences are modeled by identifying sequential patterns from recent interactions at both the review and category levels. This multilevel design captures fine-grained opinions across items, ensuring a more accurate understanding of the evolving user intent. This study conducted various experiments on real-world datasets, showing that Rec-SSP outperforms baseline models. These findings demonstrate that balancing long-term and short-term preferences with multilevel sequence modeling can significantly improve recommendation accuracy across diverse domains.

Keywords: recommender system; review text; long–short-term preferences; sequential patterns; gated fusion mechanism

1. Introduction

The rapid development of online platforms has made an overwhelming amount of information available to users, leading to information overload during item exploration [1]. This challenge is more evident in e-commerce platforms and digital content services, where users often struggle to find optimal items that suit their preferences. To address this challenge, recommender systems have been designed to recommend relevant items based on user feedback [2]. Recommender systems play a significant role in platforms such as Amazon, Netflix, and Spotify, enhancing user satisfaction and driving profitability [3].

Early recommender systems relied primarily on collaborative filtering (CF), which analyzed users' purchase histories, such as rating data, to recommend suitable items [4,5]. Matrix factorization (MF) methods were widely used in recommender systems to learn the latent factors of users and items to provide personalized recommendations [6].

Academic Editor: Arkaitz Zubiaga

Received: 22 January 2026

Revised: 8 February 2026

Accepted: 10 February 2026

Published: 10 February 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

However, these methods often suffer from data sparsity issues. To address this limitation, review texts have emerged as a valuable resource as they provide insights into user preferences, emotions, and specific characteristics of items. Some studies have argued that fully integrating review texts can enhance the representation of user and item characteristics by leveraging diverse perspectives in review sets for recommendations. These studies demonstrated that leveraging review sets can extract overall characteristics and provide more accurate recommendations [7]. For users, these characteristics reflect detailed insights from various experiences, offering a comprehensive representation of long-term preferences.

While most studies have primarily focused on representing long-term user preferences based on their overall interaction histories, they have often failed to effectively capture users' short-term preference changes. Short-term preferences are not a single homogeneous signal. In contrast, they are manifested through heterogeneous semantic cues embedded in recent user interactions, especially in review texts, such as emotional emphasis, expressive needs, and contextual descriptions. For example, a user's historical reviews may consistently emphasize functional attributes such as durability or value for money, reflecting stable long-term preferences. However, in a recent short period, the user's reviews begin to highlight different semantic aspects, such as design details or specific usage scenarios. Although the underlying intent expressed in review texts changes noticeably over a short period, such short-term semantic shifts cannot be easily captured when recent interactions are treated as homogeneous signals. Fine-grained changes in recent reviews can strongly affect users' purchasing decisions, and this impact can be easily diluted if all recent interactions are modeled the same way. These temporal changes in preferences significantly affect purchasing behavior and should be carefully considered when designing recommender systems. To address this, recent research has increasingly explored methods for modeling short-term preferences by extracting meaningful sequential patterns from user interaction histories and proposing various approaches to capture such preference changes [8–10].

However, some sequence-based models tend to emphasize recent interactions to model short-term preference changes but often overlook long-term preferences, which are equally important [11]. In other words, research efforts to balance short-term and long-term preferences remain relatively limited. Because both types of preferences jointly influence a user's final decision, failing to integrate them properly may lead to an incomplete understanding of user intent and degraded recommendation performance. Moreover, recent interactions do not always reflect a single and unified short-term preference. Many existing models represent recent interactions as one short-term sequence, mainly based on surface-level sequential patterns derived from interaction histories, which limits their ability to capture different types of short-term signals [12,13]. In practice, recent preference changes are often reflected at multiple semantic levels, such as fine-grained opinions and category transitions expressed in review texts. Modeling these heterogeneous signals in the same way may obscure their distinct roles, whereas distinguishing them enables a clearer representation of short-term user intent than relying solely on ID-based sequences.

This study proposes a novel recommender system, Rec-SSP (recommender system emphasizing short-term preference using sequential patterns), which effectively models both long-term and short-term user preferences, and extracts sequential patterns from recent user interactions. Unlike existing approaches that model recent interactions as a single short-term sequence, Rec-SSP explicitly distinguishes heterogeneous short-term signals reflected in recent reviews and category transitions. Specifically, this study defines short-term preferences as user preference signals captured within a limited window of recent interactions, reflecting rapidly evolving semantic intents expressed in review texts.

First, we aggregated the reviews written by each user to capture their long-term preferences. We also compiled all reviews received for each item to represent its characteristics. Second, we extracted sequential patterns in recent user interactions by separately modeling review-level and category-level sequences, capturing different semantic aspects of dynamically changing short-term preferences. Finally, we dynamically balance and integrate long-term and short-term preferences through a gated fusion mechanism, effectively capturing meaningful transitions in user preferences to provide accurate recommendations. The main contributions of this study are as follows:

- This study proposes a novel review-based recommendation model, which models users' long-term preferences from aggregated review texts and further enhances recommendation accuracy by incorporating short-term preference dynamics from recent interactions.
- To capture recent preference changes more effectively, this study introduces a review-aware short-term preference modeling framework that separately considers review-level and category-level interaction sequences. This design complements long-term review representations by capturing heterogeneous short-term signals that cannot be fully reflected by aggregated reviews alone.
- Extensive experiments on real-world Amazon datasets demonstrate the effectiveness of the proposed Rec-SSP model. It outperforms baseline models, and ablation studies further validate the contribution of each model component.

The structure of this paper is as follows. Section 2 reviews the related work. Section 3 provides a detailed explanation of the proposed model. Section 4 describes the experimental design, while Section 5 reports the experimental results. Finally, Section 6 concludes the study by summarizing key findings, identifying limitations, and suggesting future research directions.

2. Related Works

2.1. Review-Based Recommender System

With the growth in online information, users increasingly encounter the challenge of information overload when searching for relevant items. To solve this problem, recommender systems have been developed to efficiently assist users in finding the desired items [14]. CF is the most widely used recommender system, showing high accuracy by recommending items that a user has previously preferred over other users with similar tastes, based on user–item interaction histories [15]. MF has been regarded as the most successful technique among the CF methods owing to its efficiency and accuracy. However, early CF models that relied on simple interaction histories faced challenges such as data sparsity and difficulty in capturing users' specific purchase motivations [16].

To overcome these limitations, researchers have actively explored the use of review texts in recommender systems. Reviews contain users' detailed opinions and sentiments regarding items, allowing for a deeper understanding of item preferences and enhancing the recommendation accuracy [17]. Early studies attempted to extract latent topics from review texts using topic modeling methods such as latent Dirichlet allocation (LDA) to capture user preferences [18]. However, because LDA is based on a bag-of-words, it fails to effectively incorporate contextual information [19].

Recent advancements in natural language processing and deep learning have significantly improved recommender system performance through a better contextual understanding of textual data [20]. For example, Zheng et al. [21] employed convolutional neural networks (CNNs) to process review sets containing aggregated user and item reviews, effectively capturing contextual information and learning user and item characteristics. In addition, Chen et al. [22] utilized an attention mechanism to assign different weights to

reviews within a review set, enabling the system to emphasize more useful reviews and achieve higher recommendation accuracy. The introduction of transformer-based models such as bidirectional encoder representations from transformers (BERT) has significantly improved contextual understanding by capturing bidirectional dependencies within sentences. For example, Pugoy and Kao [23] leveraged BERT to capture deep semantic representations of review texts, improving recommendation accuracy and model interpretability.

However, most studies have focused on understanding contextual information to reflect users' long-term preferences and have not sufficiently reflected how users' interests change over time. User preferences are not fixed but can change over time, and if interests or trends at a specific point in time are not reflected, recommendation performance may decrease [24]. Recent studies introduced methods for incorporating short-term user preferences by analyzing sequential patterns in recent interactions, thereby emphasizing the importance of sequence-aware recommendation models [25]. Therefore, this study aims to comprehensively model both long-term and short-term user preferences to overcome the limitations of conventional long-term preference dependency, enabling a more accurate recommender system by employing a multilevel sequence modeling approach.

2.2. Recommendation Based on Sequential Patterns

With the advancement of e-commerce platforms, research has been conducted to model user interactions as sequence data to capture interaction patterns and recommend suitable items. Existing sequence-based recommender systems primarily focus on user–item interaction histories and employ sequence pattern mining to predict the next likely interaction [26]. These methods have evolved around Markov chain-based models [12]. These models utilize L-order Markov chains, recommending based on the last L interactions and effectively capturing short-term item transitions [27]. However, early methods relied heavily on ID embeddings, making it difficult to account for user intent or essential factors, and limiting the depth of contextual understanding [28].

More sophisticated sequence-based recommendation methods have been proposed to address these limitations by incorporating more meaningful information. Specifically, studies have explored leveraging review sequences to reflect transitions in user interactions, or using item attribute sequences to model changes in user preferences. For example, Li et al. [29] utilized review sequences left by users for interacting items to effectively model user interaction patterns for recommendations. Similarly, Hao, Zhang, Zhao, Liu, Sheng, Xu, Liu and Zhou [12] analyzed sequences of item characteristics that interact with users to capture meaningful transitions in preferences at the feature level. These studies demonstrated that sequential data composed of diverse information types beyond mere ID sequences can reveal significant preference transitions.

With advancements in deep learning techniques, substantial efforts have been made to enhance the processing of sequential data in recommender systems. Initial approaches adopted recurrent neural networks (RNNs), which are effective in modeling sequential dependencies by preserving historical information [30]. However, RNN faces limitations such as long-term dependency issues and training inefficiencies owing to their inherently sequential processing nature [31]. Long short-term memory (LSTM) networks and gated recurrent unit (GRU) were introduced to overcome this problem. LSTM employs a gating mechanism to selectively retain and discard information, enabling the effective modeling of long-term user preferences [32]. Because of their conceptual similarity to LSTM and more efficient architecture with comparable performance, GRU has become a widely adopted choice in recommender systems [33]. GRU leverages reset and update gates to manage sequential states, ensuring that essential preference information is retained while filtering out less relevant information [32]. By employing these advanced architectures,

recommendation models can better capture evolving user preferences to improve the recommendation accuracy [24].

Existing sequence-based recommendation models primarily capture short-term preferences from recent user interactions, while overlooking long-term interests. However, both short-term and long-term preferences play crucial roles in shaping user decisions, and neglecting either can reduce the recommendation performance. To address this issue, this study proposes a method that captures and effectively balances both preference types to enhance recommendation accuracy.

3. The Proposed Rec-SSP Model

This section presents a detailed description of the proposed model. We introduce the overall architecture of Rec-SSP and describe each module in detail.

3.1. Network Architecture of Rec-SSP

The design of Rec-SSP is motivated by the observation that users' short-term preferences are not homogeneous signals but are reflected through heterogeneous semantic cues in recent interactions, such as review texts and category transitions. Modeling these signals using a single sequence representation (e.g., simple concatenation or attentive pooling) may obscure their distinct roles in capturing dynamic user intent. The Rec-SSP model proposed in this study consists of three key components: a long-term preference extraction module, a short-term preference extraction module, and a final preference prediction module. The long-term preference extraction module extracts review text features and feeds them to the multilayer perceptron (MLP) to learn the accumulated long-term user preferences. The short-term preference extraction module analyzes review-level and category-level sequences to detect changes in user interaction patterns. Additionally, a multi-head attention mechanism is applied, assigning higher importance scores to information that is significantly related to item characteristics to reflect short-term preferences. In the final preference prediction module, the previously learned long-term and short-term preferences are fused in a balanced manner using a gated fusion mechanism, and then interact with the item to predict a rating that represents the user's final preference. The model aims to capture users' short-term preferences from diverse sequences composed of review texts and categories, integrate them with long-term preferences in a balanced manner, and facilitate user-item interactions to provide more accurate recommendations. The details of each module are discussed in the following sections, as illustrated in Figure 1.

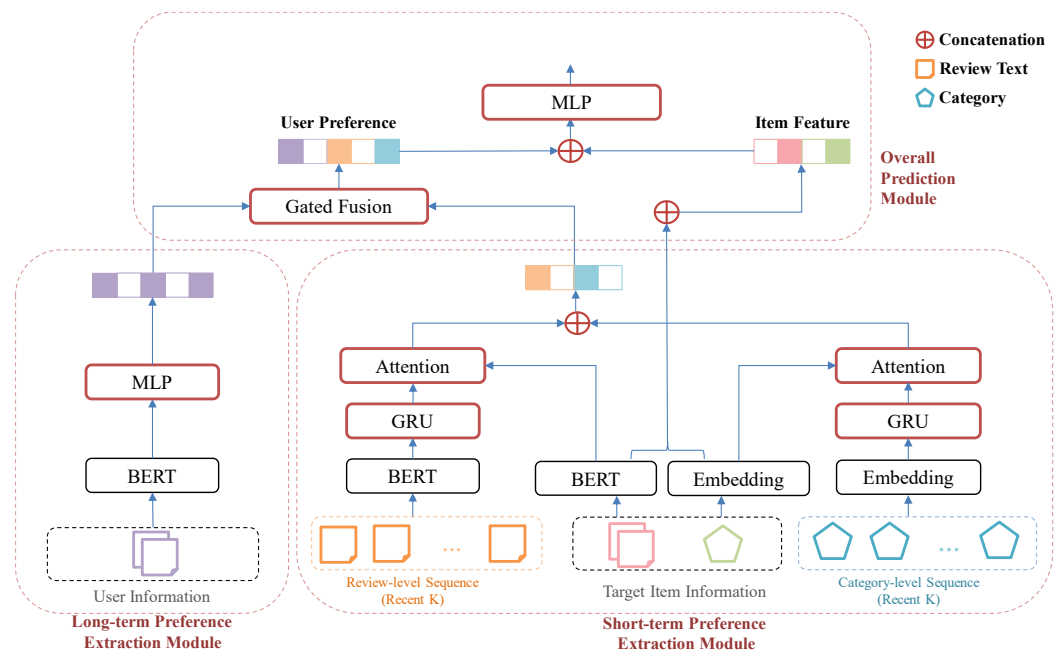


Figure 1. The architecture of Rec-SSP.

3.2. Long-Term Preference Extraction Module

The long-term preference extraction module was designed to capture users' long-term preferences based on user review sets. To achieve this, a pretrained BERT model was employed as an encoder to represent review texts as feature vectors, which were then processed through an MLP to learn the users' long-term preferences. BERT is a large-scale pretrained language model that enhances text representation by considering the bidirectional contextual relationships between words and sentences [34]. The input text to BERT undergoes three processing steps: token, segment, and position embedding. This study utilizes only the [CLS] token from the BERT's final output [35]. The semantic information within review texts was captured by applying the model, enabling the extraction of long-term preferences. The process is as follows.

First, an individual review U_{review} written by a user about an item is input into the pretrained BERT model to generate an embedding vector E_{user_review} , which can be represented by Equation (1):

$$E_{user_review} = BERT(U_{review}), \quad (1)$$

where the pretrained BERT model transforms the review text U_{review} left by a user on an item into a high-dimensional vector E_{user_review} that captures contextual information. Subsequently, the user review vectors are aggregated for each user and averaged to generate a vector E_{long_term} that represents the user's long-term preferences. This process is expressed by Equation (2):

$$E_{long_term} = \frac{1}{N} \sum_{j=1}^N E_{user_review}, \quad (2)$$

where N is the total number of reviews written by the user. The long-term preference vector E_{long_term} was obtained by averaging all user-generated review vectors. This process effectively captures the semantic characteristics inherent in each review text and comprehensively represents the user's accumulated preferences.

The previously obtained long-term preference vector was fed into the MLP, which learns complex interactions through a nonlinear optimization process. The refined

representation was then fed into the rating prediction module to predict the ratings. The MLP used in this study is defined in Equation (3) and the detailed process is described in Equation (4).

$$X_L = MLP(X_0), \quad (3)$$

$$\begin{aligned} X_1 &= \delta_1(W_1^T X_0 + b_1), \\ &\dots \\ X_L &= \delta_L(W_L^T X_{L-1} + b_L), \end{aligned} \quad (4)$$

where X_0 denotes the input vector to the MLP, and X_L denotes the output vector. Additionally, W_L^T , b_L , and δ_L correspond to the weight matrix, bias, and activation function of the L -th layer, respectively, where T indicates the transpose operation. In this study, a rectified linear unit (ReLU) was adopted as the activation function, which is formulated in Equation (5).

$$ReLU(x) = \max(0, x), \quad (5)$$

Consequently, after passing through the MLP, the final long-term user preference vector P_{long_term} was obtained, as expressed in Equation (6).

$$P_{long_term} = MLP(E_{long_term}), \quad (6)$$

3.3. Short-Term Preference Extraction Module

The short-term preference extraction module is designed to capture heterogeneous short-term preference signals reflected in users' recent interactions. Compared with modeling recent interactions as a single short-term sequence, this module separately learns review-level and category-level sequential patterns, as these two sources convey different semantic aspects of short-term user intent. Therefore, this module uses review and category sequences to generate the final short-term preference representation. The specific process is divided into two parts: embedding and short-term preference learning.

3.3.1. Embedding Process of the Short-Term Preference

The embedding process in the short-term preference extraction module includes the review text left by the user for the items they recently interacted with and the embedding for the corresponding category.

First, the review texts of items with which users have recently interacted were processed using a pretrained BERT model to obtain embedding vectors. These vectors were then arranged in temporal order to construct the review-level sequence. Specifically, because the embedding process is part of the long-term preference extraction module, it is not repeated here. Instead, the embedding vectors corresponding to the most recent K reviews are reused and organized in temporal order to construct a review-level interaction sequence. This sequence captures fine-grained preference transitions reflected in users' recent review behaviors, which is denoted as S_{review} .

Next, similar to the long-term preference extraction module, each review in an item review set is transformed into a high-dimensional feature vector using BERT. The reviews were then aggregated per item and their average was computed to generate an item vector. This process is mathematically expressed in Equations (7) and (8):

$$E_{item_review} = BERT(I_{review}), \quad (7)$$

$$E_{item} = \frac{1}{M} \sum_{i=1}^M E_{item_review}, \quad (8)$$

where I_{review} represents the review set received by an item. M denotes the total number of reviews that the item received, while E_{item_review} represents the embedding vector of each review related to item. Finally, E_{item} is a vector that captures the aggregated features extracted from the review set of items.

Meanwhile, the category sequence of items with which the user interacted, along with the target item category, underwent an embedding process. Because category information is more straightforward and intuitive than review texts, this study applies a multi-hot embedding method to each item, rather than a complex processing approach. To construct the category-level sequence, the category embedding vectors of the items that the user has recently interacted with were arranged in temporal order and grouped into a sequence. The final category-level sequence, denoted $S_{category}$, was obtained through this process.

In summary, we obtain the feature representation E_{item} and related category vector representation $E_{category}$ captured from the review text aggregated by item concerning item attributes. We also obtained two sequences ordered by time: user review-level sequence S_{review} and category-level sequence $S_{category}$.

3.3.2. Learning Process of the Sequence Pattern

Next, the sequence pattern for the short-term preference of the user was learned based on the previously obtained review-level sequence and category-level sequence. The GRU can effectively maintain information from a previous moment and remove unnecessary information using a gate mechanism. It has the advantage of being able to learn sequential data. Owing to these characteristics, the GRU can effectively model user preferences by understanding user interaction patterns. To capture the user's short-term preferences that change over time, S_{review} and $S_{category}$ were input into the GRU to learn the sequence pattern reflecting the temporal flow, which can be expressed by Equations (9) and (10):

$$G_{review} = GRU(S_{review}), \quad (9)$$

$$G_{category} = GRU(S_{category}), \quad (10)$$

where G_{review} and $G_{category}$ represent the sequence pattern vectors learned through the GRU from the review-level and category-level perspectives.

The multi-head attention mechanism incorporating item characteristics is applied to extract users' short-term preferences. Originally introduced in transformer models, multi-head attention is widely used in deep learning models to enhance their performance by assigning higher weights to more important information. This mechanism allows for a more precise understanding of user interactions and contributes to personalized recommendations. In this study, the feature vector of the target item is used as a query to emphasize and reflect information that is highly relevant to the target item within the sequential pattern of the user's interaction. Specifically, for the review-level sequence, E_{item} is used as the query, while for the G_{review} sequence, $E_{category}$ is used as the query and $G_{category}$ serves as the key and value for category-level multi-head attention. By applying multi-head attention, the importance of each item in user interaction related to the target item is effectively captured from multiple perspectives, allowing for more accurate learning of users' short-term preferences. This process is expressed in Equations (11)–(13):

$$Q_i = M \times W_i^Q, K_i = M \times W_i^K, V_i = M \times W_i^V, \quad (11)$$

$$O_i = \text{Attention}(Q_i, K_i, V_i) = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d}}\right) V_i, \quad (12)$$

$$O = \text{Multihead}(M) = \text{Concatenate}(O_1, O_2 \dots O_h) W^O, \quad (13)$$

where Q_i , K_i , V_i represent the query, key, and value vectors of the i -th head, respectively, and W_i^Q , W_i^K , W_i^V represent the weight matrices for query, key, and value of the i -th head, respectively. W^O represents a projection matrix, and O is the final output of multi-head attention that connects the outputs of h heads, emphasizing the essential nodes in the sequence. Here, M denotes the sequence representation learned by the GRU. Specifically, M corresponds to G_{review} for the review-level attention and G_{category} for the category-level attention. The query vectors are constructed using the target item representation, while the key and value vectors are derived from the corresponding sequence representations. The attention score is calculated using the inner product of the query and the key, and then transformed by applying the softmax function. The final expression O is obtained by connecting the outputs of each head using a combination process. Therefore, the review-level and category-level attention outputs are obtained as shown in Equations (14) and (15):

$$Z_{\text{review}} = \text{Multihead}(G_{\text{review}}), \quad (14)$$

$$Z_{\text{category}} = \text{Multihead}(G_{\text{category}}), \quad (15)$$

we then combine Z_{review} and Z_{category} to obtain the user's final short-term preference vector representation. This process is expressed in Equation (16):

$$P_{\text{short_term}} = \text{Concatenate}(Z_{\text{review}}, Z_{\text{category}}), \quad (16)$$

here $P_{\text{short_term}}$ reflects the relevance between the target item and the user's recent interaction. The user's short-term preference, captured in this manner, more precisely expresses the user's preference change over time. It is then used as an input to the rating prediction module to effectively represent a user's final preference.

3.4. Overall Preference Prediction Module

In this module, a fused vector is generated by balancing the importance of various user preferences and then interacting with an item to predict the final rating that represents the final preference for the target item. The process is as follows.

Users' final purchase decisions are comprehensively influenced by their long-term and short-term preferences, the importance of which varies depending on the specific situation [36]. Therefore, in this study, the users' long-term and short-term preferences are balanced and fused using the gated fusion method to obtain the final user preference representation. The gated fusion mechanism is characterized by dynamic learning, which appropriately weighs the contributions when combining different data sources, thereby minimizing information loss and obtaining an optimal fusion result [37]. Unlike other fusion strategies that treat long-term and short-term signals in a uniform manner, the gated fusion mechanism explicitly learns a data-dependent gate to regulate their relative contributions. This design enables the model to emphasize short-term preferences when recent interactions exhibit semantic shifts, while relying more on long-term preferences when

user behavior remains stable. The process of gated fusion can be expressed as Equations (17)–(20):

$$h_v = \tanh(W_v \square X_v), \quad (17)$$

$$h_t = \tanh(W_t \square X_t), \quad (18)$$

$$z = \text{sigmoid}(W_z \square \text{Concatenate}(X_v, X_t)), \quad (19)$$

$$h = z \square h_v + (1 - z) \square h_t, \quad (20)$$

where W_v and W_t are the weight matrices of the data sources X_v and X_t , to be fused, respectively. The calculated h_v and h_t are concatenated to calculate the gate vector z , and W_z is the weight matrix for calculating the gate vector. Subsequently, the contributions of the two data sources, X_v and X_t are integrated through a gating vector h , enabling the model to learn the vector representation. The final fused vector representation h is obtained, ensuring a comprehensive and balanced representation of the user preferences.

Before gated fusion, the long-term preference vector P_{long_term} and the short-term preference vector P_{short_term} are projected into the same latent space to ensure dimensional compatibility. Therefore, W_v , W_t , and W_z are learnable projection matrices with compatible dimensions, enabling effective gating and adaptive fusion of heterogeneous preference representations. The process of fusion can be expressed as Equation (21):

$$U_{preference} = \text{Gate}(P_{long_term}, P_{short_term}), \quad (21)$$

where $U_{preference}$ represents the user's final preference and Gate represents the gated fusion mechanism. To interact with the user preference vector, the final item representation $I_{feature}$ was obtained by concatenating the vectors E_{item} and $E_{category}$ obtained from the review texts received for the item.

Subsequently, the user preference feature vector $U_{preference}$ and item vector $I_{feature}$ were combined and passed through the MLP to predict the user's final preference rating for the item. This process is expressed by Equation (22):

$$\hat{r}_{ui} = \text{MLP}(\text{Concatenate}(U_{preference}, I_{feature})), \quad (22)$$

where $\hat{r}_{ui}$ represents the predicted rating of user u for a specific item i .

3.5. Loss Function

The proposed model was trained using the mean squared error (MSE) as the loss function. This approach computes the difference between the predicted and actual values, squares it, and averages it across all data points [14,38]. The MSE is widely used in rating prediction tasks in recommender systems and can be formulated as Equation (23):

$$MSE = \frac{1}{N} \sum_{i=1}^N (r_{ui} - \hat{r}_{ui})^2, \quad (23)$$

where N represents the number of ratings, r_{ui} is the actual rating given by the user for item, and $\hat{r}_{ui}$ is the rating predicted by the model. The proposed model was trained using the adaptive moment estimation (Adam) optimizer [31].

4. Experiments

This study aims to analyze and answer the following four research questions (RQs).

RQ1: Does the proposed model outperform baseline models on real-world datasets?

RQ2: How does incorporating long-term and short-term user preferences improve the recommendation performance?

RQ3: How does analyzing sequence patterns at the review-level and category-level in the short-term preference extraction module affect the recommendation accuracy?

RQ4: What is the most effective method for fusing long-term and short-term user preferences?

4.1. Datasets

To evaluate the recommendation performance of Rec-SSP, we conducted extensive experiments on datasets collected from both Amazon and Yelp platforms. Specifically, we used three product categories from Amazon, including Video Games, CDs and Vinyl, and Musical Instruments. To further assess the generalizability of the proposed model, we also used a Yelp dataset, selecting restaurant data from the city with the most reviews. For all datasets, we filtered users with at least five reviews to ensure reliable preference sequences. Each dataset was split into training, validation, and test sets with a ratio of 7:1:2. The overall statistical characteristics of the datasets are summarized in Table 1.

Table 1. Descriptive statistics of experimental datasets.

Dataset	Number of Reviews	Number of Users	Number of Items	Sparsity
Video Games	114,872	82,834	1,002,380	99.990%
CDs and Vinyl	181,125	534,064	2,475,396	99.997%
Musical Instruments	82,878	122,086	786,939	99.992%
Yelp	274,651	19,446	3097	99.544%

Moreover, we further analyze the characteristics of user interaction sequences. Table 2 reports quartile-based sequence-level statistics, including the distribution of user interaction sequence length, the number of category transitions, and the review text length.

Table 2. Quartile-based sequence-level characteristics of experimental datasets.

Dataset	Ratio	Sequence Length	Category Change	Review Text Length
Video Games	25%	5	1	9
	50%	6	2	31
	75%	9	3	88
	100%	664	348	5860
CDs and Vinyl	25%	6	3	11
	50%	8	5	42
	75%	12	8	127
	100%	2602	1744	6515
Musical Instruments	25%	5	2	10
	50%	7	3	28
	75%	10	5	66
	100%	497	276	6060
Yelp	25%	6	3	52
	50%	8	4	92
	75%	14	7	157
	100%	610	253	1019

4.2. Evaluation Metrics

This study employs two widely used evaluation metrics in rating prediction research for recommender systems: the mean absolute error (MAE) and root mean squared error (RMSE). The MAE calculates the average absolute difference between the predicted and actual ratings, providing an intuitive measure of the prediction error of the model. The RMSE squares the differences between the predicted and actual ratings, averages them, and then takes the square root, assigning a greater weight to more significant errors. Both metrics indicate the deviation between the predicted and actual values, where smaller values correspond to a higher accuracy [39]. Equations (24) and (25) show the specific computational methods.

$$MAE = \frac{1}{N} \sum_{i=1}^N |r_{ui} - \hat{r}_{ui}|, \quad (24)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_{ui} - \hat{r}_{ui})^2}. \quad (25)$$

where N represents the number of data points, and r_{ui} and $\hat{r}_{ui}$ represent the actual rating values and the rating values predicted from the model, respectively.

4.3. Baseline Models

To verify the performance of the Rec-SSP proposed in this study, the following baseline models were selected and compared:

Neural collaborative filtering (NCF) [40]: A recommender system that overcomes the limitations of linear combination by incorporating nonlinear functions into the MF to learn interactions between users and items' latent features.

Dual attention-based CNN model (D-attn) [38]: A recommender system that applies local attention and global attention to user and item review sets, allowing it to focus on important words while deeply understanding the overall meaning of reviews, thereby improving preference prediction accuracy.

Deep cooperative neural networks (DeepCoNN) [21]: A recommender system that employs two parallel CNNs to extract user and item latent features from review sets, using factorization machines to model user and item characteristics jointly.

Neural attentional regression model with review-level explanations (NARRE) [22]: A recommendation model that applies an attention mechanism to review sets, recognizing the importance of each review and assigning higher weights to more critical reviews to enhance recommendation accuracy.

Review-driven neural sequential recommendation (RNS) [29]: A recommendation model that integrates long-term and short-term user preference learning using user review sets. It employs CNNs to extract aspects representing user interactions with items and applies a hierarchical attention mechanism to effectively highlight relevant items.

Cascaded cross attention (CCA) [41]: A recommendation model that models complex relationships between users and target items by aggregating user review texts and capturing interaction patterns across multiple sequences, enabling refined user-item interaction learning.

4.4. Experimental Design

One of the key challenges in this study is the analysis of user interaction sequence patterns and the learning of meaningful transitions. The experiments used two sequences: a sequence consisting of a single review left by the user on the most recently interacted items, and a sequence based on related categories. All sequences were sorted in temporal

order, and zero-padding was applied at the beginning of shorter sequences to maintain a consistent length. A pretrained BERT language model was applied to the textual reviews utilized in this study, and the output [CLS] token was used as an embedding vector. The embedding vectors were averaged for the aggregated review sets collected per user and item. Categorical data were represented using a multi-hot embedding technique.

To optimize the performance of the proposed model, key hyperparameters were tuned based on validation performance. Specifically, the batch size was selected from [64, 128, 256, 512], the learning rate from [0.0002, 0.0005, 0.002, 0.005], and the number of attention heads in the multi-head attention mechanism from [4, 6, 8, 10, 12]. Based on model tuning results, the final configuration adopted a batch size of 512, a learning rate of 0.0002, and 6 attention heads. For short-term preference modeling, the sequence length K was initially set to 5 following the previous study [41]. To further examine the impact of K , we conducted a sensitivity analysis in Section 5.5.

All models employed early stopping, meaning that training was stopped if the validation loss did not decrease for five consecutive epochs. Each model was run five times for all experiments, and the average results were used for the evaluation. All performance measurements were conducted using TensorFlow 2.9.0 and Python 3.9 to ensure experimental fairness.

5. Experimental Results

5.1. Performance Comparison with Baseline Models (RQ1)

To evaluate the performance of the Rec-SSP model proposed in this study, we conducted a performance comparison experiment with existing baseline models using user review data and metadata collected from Amazon and Yelp platforms. Table 3 presents the results of the study.

Table 3. Performance comparison between the proposed model and baseline models.

Model	Video Games		CDs and Vinyl		Musical Instruments		Yelp	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
NCF	0.957	1.214	0.747	1.014	0.838	1.103	0.866	1.129
D-attn	0.864	1.082	0.733	0.920	0.790	0.999	0.829	1.045
DeepCoNN	0.801	1.076	0.616	0.880	0.717	0.999	0.803	1.041
NARRE	0.798	1.070	0.594	0.851	0.705	0.998	0.802	1.037
RNS	0.777	1.065	0.597	0.863	0.702	0.989	0.793	1.034
CCA	0.711	1.008	0.572	0.833	0.674	0.970	0.767	0.988
Rec-SSP	0.679	0.979	0.537	0.824	0.621	0.924	0.753	0.977

The experimental results demonstrated that the proposed Rec-SSP model outperformed all baseline models across all datasets. These findings indicate that effectively modeling user preferences from both long-term and short-term perspectives and fusing these preferences in a balanced manner significantly improves the accuracy of predicting user preferences for items. In particular, such a consistent improvement observed on the Yelp dataset indicates that Rec-SSP remains effective even in recommendation scenarios with more heterogeneous user behaviors and item distributions.

Interaction-based model like NCF exhibited the lowest recommendation performance, indicating that relying solely on interaction data is insufficient to capture users' explicit purchasing motivations and limits their ability to fully understand user preferences.

Review Set-based models such as D-attn, DeepCoNN, and NARRE, which leverage review text, demonstrated improved performance over interaction-based models. This

result suggests that extracting user preferences from semantic information in review texts effectively enhances recommendation accuracy. NARRE provides more accurate recommendations than D-attn and DeepCoNN by assigning higher weights to critical reviews, reinforcing the idea that emphasizing essential information in user reviews leads to better preference predictions.

Sequence pattern-based models that analyze user interaction patterns such as RNS and CCA outperformed interaction- and review-set-based models. This indicates that incorporating short-term preference transitions enhances the recommendation performance. Additionally, the CCA surpassed the RNS, suggesting that analyzing multiple sequences rather than a single sequence improves the model's ability to capture users' evolving preferences more effectively.

The proposed Rec-SSP model outperformed all the baseline models in terms of recommendation performance. These results can be interpreted as follows. Rec-SSP moves beyond traditional approaches that rely solely on review texts by capturing long-term user preferences from aggregated review sets, deeply analyzing short-term preferences through diverse interaction sequences at both the review and category-levels, and integrating these preferences in a balanced manner for recommendations. Unlike conventional methods that depend exclusively on interaction data, Rec-SSP effectively utilizes additional contextual information such as user review texts and item categories to derive deeper contextual insights from recently interacted items. This enables the model to capture evolving short-term user preferences dynamically. Moreover, rather than merging different preference signals, Rec-SSP employs a gated fusion mechanism that dynamically adjusts the weights of multiple preference components based on their impact on the final purchase decisions, allowing for a more precise prediction.

To assess the practical applicability of Rec-SSP, we report its model complexity and training efficiency. On the Video Games dataset, the proposed Rec-SSP model contains approximately 2.79 M trainable parameters and requires about 30 s per epoch during training. These results indicate that Rec-SSP can be trained efficiently while maintaining a compact model size, demonstrating that performance improvements are achieved without introducing prohibitive computational overhead.

5.2. Importance of Reflecting Long-Term and Short-Term Preferences (RQ2)

The proposed Rec-SSP can be divided into two primary modules: long-term and short-term preference extraction modules. To analyze the contribution of each module to the overall recommendation performance, the following ablation experiments were designed, and the results are presented in Table 4.

Long-term: The short-term preference extraction module is removed from Rec-SSP. The final user preference is predicted based on the interactions between the long-term preferences extracted from the user's aggregated review set and the item characteristics.

Short-term: The long-term preference extraction module is removed from the Rec-SSP. The final user preference is predicted based on interactions between short-term preferences, obtained by analyzing sequences of reviews, categories of recently interacting items, and item characteristics.

Table 4. Ablation study results of long-term and short-term preference extraction modules.

Module	Video Games		CDs and Vinyl		Musical Instruments		Yelp	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Long-term	0.693	0.993	0.563	0.829	0.649	0.930	0.757	0.981
Short-term	0.691	0.992	0.575	0.839	0.641	0.937	0.772	0.997
Rec-SSP	0.679	0.979	0.537	0.824	0.621	0.924	0.753	0.977

The experimental results demonstrate that the combined approach of capturing long-term and short-term user preferences in Rec-SSP effectively enhances recommendation performance. In contrast, using only long-term or short-term exhibited lower performance, likely owing to their limited representation of user preferences compared with Rec-SSP.

Rec-SSP overcomes the limitations of relying solely on interaction data or general review texts by extracting long-term preferences from aggregated review sets and capturing preference transitions using sequential patterns. In contrast, using only long-term or short-term failed to incorporate the essential characteristics of user preferences, resulting in less accurate predictions. Focusing exclusively on short-term preference transitions limits recommendation effectiveness because interactions are only considered within a given order window, making it difficult to account for long-term accumulated user preferences [42]. However, while long-term preference-based personalized recommendations can effectively suggest items that align with a user's general interests, they struggle to reflect recent changes in user preferences [27].

Moreover, Rec-SSP consistently demonstrates superior performance across datasets with varying characteristics, validating that the integrated approach of balancing long-term and short-term preferences is broadly applicable across different datasets. For example, the CDs and Vinyl and Yelp datasets may involve user interaction behaviors that emphasize accumulated historical preferences, while in other datasets, short-term preference transitions may be relatively more influential. The proposed model effectively adapts to these variations by jointly modeling long-term and short-term preferences, leading to robust recommendation performance across datasets with different characteristics.

5.3. Importance of Review-Level and Category-Level Sequence Pattern Analysis (RQ3)

The proposed Rec-SSP model enhances the recommendation performance by incorporating short-term user preferences through contextual patterns derived from reviews and category sequences of recently interacted items. This section explores the impact of each type of sequence feature and eliminates the possibility that the observed performance improvements were solely due to the inclusion of auxiliary information. The following ablation experiments were designed to achieve this goal, and the results are listed in Table 5.

ID sequence: Rec-SSP captures user preferences solely based on item ID sequences without incorporating textual or categorical information.

Review sequence: Rec-SSP captures short-term preferences using only a user review sequence.

Category sequence: Rec-SSP captures short-term preferences using only category sequences.

W/O GRU: Rec-SSP utilizes user reviews and category information but does not process them as sequences. Instead, it concatenates the features before they are used for training.

Table 5. Ablation study results for review-level and category-level sequence modeling.

Variant	Video Games		CDs and Vinyl		Musical Instruments		Yelp	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ID Sequence	0.708	0.996	0.646	0.935	0.685	0.966	0.774	0.992
Review Sequence	0.692	0.988	0.583	0.826	0.637	0.927	0.762	0.979
Category Sequence	0.702	0.991	0.587	0.827	0.642	0.930	0.765	0.980
W/O GRU	0.691	0.987	0.558	0.826	0.632	0.925	0.762	0.980
Rec-SSP	0.679	0.979	0.537	0.824	0.621	0.924	0.753	0.977

The experimental results demonstrate that the proposed Rec-SSP model consistently achieves stable performance and substantial results across multiple datasets. These findings highlight the importance of modeling user preferences from multiple perspectives and analyzing both review- and category-level sequence features to capture short-term preferences and enhance recommendation performance.

Specifically, using only the item ID sequence shows the lowest performance among all variants, indicating that ID-based interaction information alone is insufficient to capture users' fine-grained preferences. Meanwhile, using only the review sequence shows slightly better performance than using only the category sequence, indicating that review sequences contribute more to user preference modeling. Additionally, removing the GRU and simply concatenating features leads to a relatively minor performance degradation, but still performs worse than Rec-SSP. This result confirms that the improved performance of Rec-SSP is not merely due to the addition of auxiliary information.

By contrast, Rec-SSP effectively utilizes review-level and category-level sequences and employs a GRU to refine short-term preference learning, ensuring that sufficient information is captured. Overall, this experiment suggests that while review sequences provide more critical information than category sequences, jointly considering diverse and meaningful sequences contributes to improved recommendation performance.

5.4. Validation of the Effectiveness of Long-Term and Short-Term Preference Fusion Mechanism (RQ4)

In this study, Rec-SSP proposed balanced long-term and short-term user preferences through a gated fusion mechanism to obtain the final user preferences. In this section, performance comparison experiments were conducted by applying concatenation, multiplication, and averaging operation methods to the part where the gated fusion mechanism was applied to analyze the effect of fusion methods on performance. The experimental results are listed in Table 6.

Table 6. Performance comparison based on fusion methods.

Method	Video Games		CDs and Vinyl		Musical Instruments		Yelp	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Concatenation	0.689	0.981	0.570	0.830	0.641	0.926	0.761	0.982
Multiplication	0.699	0.983	0.545	0.829	0.629	0.928	0.766	0.981
Average	0.683	0.981	0.546	0.832	0.631	0.925	0.761	0.980
Rec-SSP	0.679	0.979	0.537	0.824	0.621	0.924	0.753	0.977

The experimental results demonstrate that Rec-SSP, which applies the gated fusion mechanism, consistently outperforms other fusion methods, maintaining a stable and superior performance. This result indicates that Rec-SSP dynamically models the importance of long-term and short-term preference information based on context, allowing each preference type to significantly influence the final purchase decisions in relevant situations.

Concatenation connects two vectors to minimize information loss. However, it fails to adequately emphasize the key factors contributing to user preferences, leading to performance degradation [43]. Multiplication captures the interactions between two vectors and reflects their correlations, but does not effectively contribute to deriving the final preferences [44]. Additionally, averaging integrates both types of preference information, but cannot emphasize more influential preferences, resulting in suboptimal performance [44]. By contrast, Rec-SSP utilizes a gated fusion mechanism to dynamically integrate long-term and short-term preferences by considering their relative importance. This approach

enables the model to capture user preferences more accurately, thereby enhancing recommendation performance.

5.5. Sensitivity Analysis of Key Hyperparameters

To further investigate the impact of key hyperparameters on model performance, we conducted a sensitivity analysis focusing on the short-term sequence length K and the number of attention heads in the multi-head attention mechanism. Although K was initially set to 5 following previous studies, we also performed a sensitivity analysis by varying K to examine its effectiveness. The experimental results on the Video Games dataset are listed in Table 7.

Table 7. Sensitivity analysis of the short-term sequence length.

Sequence Length	MAE	MSE	RMSE
3	0.697	0.975	0.987
4	0.680	0.979	0.989
5	0.679	0.959	0.979
6	0.685	0.979	0.989
7	0.695	0.974	0.987

The experimental results demonstrate that as the sequence length increases, recommendation performance initially improves due to richer short-term contextual information, and then peaks at 5. However, when the sequence length exceeds 5, performance gradually degrades, suggesting that excessively long sequences may introduce noise from less relevant historical interactions. Therefore, setting the sequence length to 5 represents a reasonable choice for the proposed Rec-SSP model.

Table 8 presents the recommendation performance of Rec-SSP under different numbers of attention heads.

Table 8. Sensitivity analysis of the number of attention heads.

Head	MAE	MSE	RMSE
4	0.703	0.973	0.987
6	0.679	0.959	0.979
8	0.688	0.970	0.985
10	0.683	0.962	0.981
12	0.690	0.972	0.986

The experimental results show that increasing the number of attention heads from 4 to 6 improves performance, indicating that multi-head attention benefits from capturing diverse interaction patterns in short-term sequences. However, when the number of attention heads exceeds 6, the performance gradually degrades. This suggests that an excessive number of attention heads may introduce redundancy and increase model complexity, making it more difficult to effectively aggregate informative signals. Therefore, setting the number of attention heads to 6 represents a reasonable trade-off between modeling capacity and efficiency for the proposed Rec-SSP model.

5.6. Case Study on the Gated Fusion Mechanism

To further enhance the interpretability of the proposed Rec-SSP model, we present a case study to illustrate how the gated fusion mechanism dynamically balances long-term and short-term user preferences. Figure 2 illustrates the preference fusion process for an individual user.

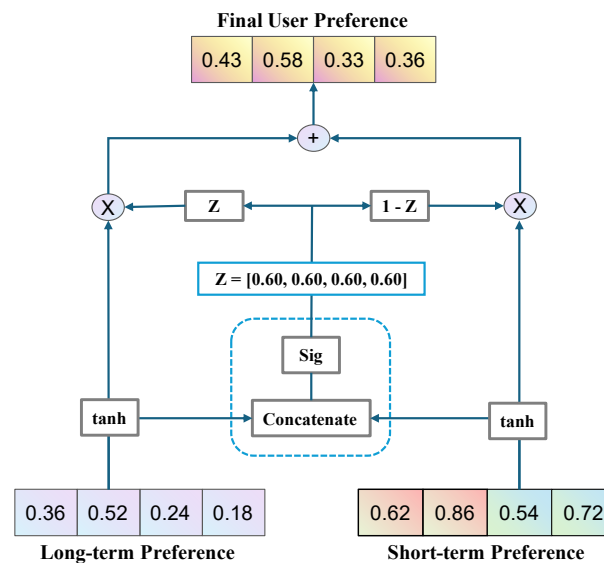


Figure 2. Example of preference fusion with the gated fusion mechanism.

The gated fusion mechanism adaptively controls the relative contributions of long-term and short-term preferences. In this example, the model assigns a higher overall weight to the long-term preference component, indicating that the user's historical interests play an important role in the current recommendation decision, while recent interactions still contribute complementary information. It demonstrates the flexibility of the gated fusion mechanism, which allows Rec-SSP to adjust its focus across different user contexts rather than relying on a fixed combination strategy. Therefore, if a user exhibits stable preference patterns, long-term preferences are emphasized, whereas short-term dynamics can be strengthened if recent behavioral signals become more important.

6. Conclusions

With the advancement of internet technologies and the growth of information, the importance of recommender systems that effectively capture user preferences and suggest suitable items has become increasingly significant. This study was conducted to address the limitations of existing recommender systems, which primarily focus on long-term preferences and fail to accurately reflect evolving short-term user preferences. Therefore, this study introduces a novel approach that captures short-term preferences by utilizing user review-level and category-level sequences, while integrating them in a balanced manner with long-term preferences extracted from users' aggregated reviews. The proposed Rec-SSP effectively captures both long-term and short-term preferences and achieves superior recommendation performance by applying a gated fusion mechanism to dynamically integrate these preferences, reflecting the complex interdependencies within user interactions. The experimental results validate that Rec-SSP outperforms existing baseline models across all datasets. Through ablation and comparative experiments, this study systematically analyzed the contribution of each component to performance improvement and demonstrated that Rec-SSP maintained robust effectiveness across datasets with varying characteristics. In conclusion, Rec-SSP demonstrates significant potential for enhancing recommendation accuracy by effectively leveraging user reviews and category information and integrating long-term and short-term preferences in a balanced manner. This study contributes to recommender systems by providing a foundational approach for accurately modeling user preferences and delivering precise recommendations.

However, this study had several limitations. First, the proposed methodology relies on a sequence pattern analysis, which often incurs a high computational cost. In particular,

when processing large-scale datasets, the computational burden increases significantly, posing challenges for real-time predictions and large-scale deployments. Future research should explore lightweight models that efficiently analyze sequence data while reducing computational complexity. Second, although this study utilizes review and item category sequences, these data alone may not fully capture user preferences. For example, users who repeatedly purchase items within the same category and explore diverse categories exhibit distinct behaviors that the current methodology may struggle to differentiate. Therefore, incorporating additional information, such as user profile data and metadata (e.g., item descriptions and brand information), could be a promising direction for future research. Finally, although this study focuses on offline preference modeling, Rec-SSP also provides a flexible foundation for extension to online and continuous learning scenarios. In real-world recommender systems, user feedback can be used to adapt the model's behavior over time. Inspired by recent advances in reinforcement learning with human feedback [45], future work could integrate such feedback to dynamically adjust the gated fusion mechanism and continuously rebalance long-term and short-term preferences.

Author Contributions: Conceptualization, L.J., X.L., and J.K.; Methodology, L.J., X.L., and S.K.; Software, L.J., and S.K.; Data curation, S.K.; Writing—original draft, L.J.; Writing—review & editing, L.J., X.L., and J.K.; Supervision, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research is (partially) funded by the BK21 FOUR program of Graduate School, Kyung Hee University (GS-1-JO-ON-20241900).

Data Availability Statement: The Amazon and Yelp datasets are publicly available at <https://amazon-reviews-2023.github.io/> and <https://business.yelp.com/data/resources/open-dataset/> (accessed on 2 February 2026).

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

References

1. Yang, S.; Li, Q.; Lim, H.; Kim, J. An attentive aspect-based recommendation model with deep neural network. *IEEE Access* **2024**, *12*, 5781–5791.
2. Guan, X.; Cheng, Z.; He, X.; Zhang, Y.; Zhu, Z.; Peng, Q.; Chua, T.-S. Attentive aspect modeling for review-aware recommendation. *ACM Trans. Inf. Syst. (TOIS)* **2019**, *37*, 1–27.
3. Sun, Y.; Zhang, Y. Conversational recommender system. In Proceedings of the 41st International Acm Sigir Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, 8–12 July 2018; pp. 235–244.
4. Liu, H.; Wang, Y.; Peng, Q.; Wu, F.; Gan, L.; Pan, L.; Jiao, P. Hybrid neural recommendation with joint deep representation learning of ratings and reviews. *Neurocomputing* **2020**, *374*, 77–85.
5. Hyun, D.; Park, C.; Cho, J.; Yu, H. Learning to utilize auxiliary reviews for recommendation. *Inf. Sci.* **2021**, *545*, 595–607.
6. Wang, R.; Jiang, Y.; Lou, J. Attention-based dynamic user preference modeling and nonlinear feature interaction learning for collaborative filtering recommendation. *Appl. Soft Comput.* **2021**, *110*, 107652.
7. Gao, J.; Lin, Y.; Wang, Y.; Wang, X.; Yang, Z.; He, Y.; Chu, X. Set-sequence-graph: A multi-view approach towards exploiting reviews for recommendation. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Galway, Ireland, 19–23 October 2020; pp. 395–404.
8. Guo, W.; Tian, J.; Li, M. Price-aware enhanced dynamic recommendation based on deep learning. *J. Retail. Consum. Serv.* **2023**, *75*, 103500.
9. Cho, J.; Hyun, D.; Kang, S.; Yu, H. Learning heterogeneous temporal patterns of user preference for timely recommendation. In Proceedings of the Web Conference 2021, Ljubljana, Slovenia, 19–23 April 2021; pp. 1274–1283.
10. Zhou, R.; Qian, H.; Zhang, J.; Liu, C.; Wan, J.; Ren, Y.; Xiong, N.; Zhao, N.; Zhang, S. Self-attention mechanism enhanced user interests modeling for personalized recommendation services in cyber-physical-social systems. *IEEE Trans. Netw. Sci. Eng.* **2021**, *9*, 3835–3846.

11. Zang, T.; Zhu, Y.; Zhu, J.; Xu, Y.; Liu, H. MPAN: Multi-parallel attention network for session-based recommendation. *Neuro-computing* **2022**, *471*, 230–241.
12. Hao, Y.; Zhang, T.; Zhao, P.; Liu, Y.; Sheng, V.S.; Xu, J.; Liu, G.; Zhou, X. Feature-level deeper self-attention network with contrastive learning for sequential recommendation. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 10112–10124.
13. Duan, Y.; Liu, P.; Lu, Y. MhSa-GRU: Combining user's dynamic preferences and items' correlation to augment sequence recommendation. *J. Intell. Inf. Syst.* **2023**, *61*, 225–248.
14. Li, W.; Xu, B. Aspect-based fashion recommendation with attention mechanism. *IEEE Access* **2020**, *8*, 141814–141823.
15. Liu, X.; Wu, S.; Zhang, Z.; Shen, C. Unify local and global information for top-N recommendation. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, 11–15 July 2022; pp. 1262–1272.
16. Zhu, Z.; Yan, M.; Deng, X.; Gao, M. Rating prediction of recommended item based on review deep learning and rating probability matrix factorization. *Electron. Commer. Res. Appl.* **2022**, *54*, 101160.
17. Chen, L.; Chen, G.; Wang, F. Recommender systems based on user reviews: The state of the art. *User Model. User-Adapt. Interact.* **2015**, *25*, 99–154.
18. Kanwal, S.; Nawaz, S.; Malik, M.K.; Nawaz, Z. A review of text-based recommendation systems. *IEEE Access* **2021**, *9*, 31638–31661.
19. Da'ud, A.; Salim, N.; Idris, R. An adaptive deep learning method for item recommendation system. *Knowl. -Based Syst.* **2021**, *213*, 106681.
20. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent trends in deep learning based natural language processing. *IEEE Comput. Intell. Mag.* **2018**, *13*, 55–75.
21. Zheng, L.; Noroozi, V.; Yu, P.S. Joint deep modeling of users and items using reviews for recommendation. In Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, Cambridge, UK, 6–10 February 2017; pp. 425–434.
22. Chen, C.; Zhang, M.; Liu, Y.; Ma, S. Neural attentional rating regression with review-level explanations. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 1583–1592.
23. Pugoy, R.A.; Kao, H.-Y. BERT-based neural collaborative filtering and fixed-length contiguous tokens explanation. In Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing, Suzhou, China, 4–7 December 2020; pp. 143–153.
24. Zhang, J.; Mu, X.; Zhao, P.; Kang, K.; Ma, C. Improving current interest with item and review sequential patterns for sequential recommendation. *Eng. Appl. Artif. Intell.* **2021**, *104*, 104348.
25. Alabduljabbar, R.; Alshareef, M.; Alshareef, N. Time-aware recommender systems: A comprehensive survey and quantitative assessment of literature. *IEEE Access* **2023**, *11*, 45586–45604.
26. Kang, W.-C.; McAuley, J. Self-attentive sequential recommendation. In Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM), Singapore, 17–20 November 2018; pp. 197–206.
27. Rendle, S.; Freudenthaler, C.; Schmidt-Thieme, L. Factorizing personalized markov chains for next-basket recommendation. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 811–820.
28. Hou, Y.; Mu, S.; Zhao, W.X.; Li, Y.; Ding, B.; Wen, J.-R. Towards universal sequence representation learning for recommender systems. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022; pp. 585–593.
29. Li, C.; Niu, X.; Luo, X.; Chen, Z.; Quan, C. A review-driven neural model for sequential recommendation. *arXiv* **2019**, arXiv:1907.00590.
30. Choe, B.; Kang, T.; Jung, K. Recommendation system with hierarchical recurrent neural network for long-term time series. *IEEE Access* **2021**, *9*, 72033–72039.
31. Ghasemi, N.; Momtazi, S. Neural text similarity of user reviews for improving collaborative filtering recommender systems. *Electron. Commer. Res. Appl.* **2021**, *45*, 101019.
32. Brännvall, R.; Forsgren, H.; Sandin, F.; Liwicki, M. ReLU and Addition-based Gated RNN. *arXiv* **2023**, arXiv:2308.05629.
33. Lai, C.-H.; Liu, D.-R.; Lien, K.-S. A hybrid of XGBoost and aspect-based review mining with attention neural network for user preference prediction. *Int. J. Mach. Learn. Cybern.* **2021**, *12*, 1203–1217.
34. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
35. Xie, Q.; Zhang, X.; Ding, Y.; Song, M. Monolingual and multilingual topic analysis using LDA and BERT embeddings. *J. Informetr.* **2020**, *14*, 101055.

36. Lei, J.; Li, Y.; Yang, S.; Shi, W.; Wu, Y. Two-stage sequential recommendation for side information fusion and long-term and short-term preferences modeling. *J. Intell. Inf. Syst.* **2022**, *59*, 657–677.
37. Arevalo, J.; Solorio, T.; Montes-y-Gómez, M.; González, F.A. Gated multimodal units for information fusion. *arXiv* **2017**, arXiv:1702.01992.
38. Seo, S.; Huang, J.; Yang, H.; Liu, Y. Interpretable convolutional neural networks with dual local and global attention for review rating prediction. In Proceedings of the Eleventh ACM Conference on Recommender Systems, Como, Italy, 27–31 August 2017; pp. 297–305.
39. Isinkaye, F.O.; Folajimi, Y.O.; Ojokoh, B.A. Recommendation systems: Principles, methods and evaluation. *Egypt. Inform. J.* **2015**, *16*, 261–273.
40. He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; Chua, T.-S. Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web, Perth, Australia, 3–7 April 2017; pp. 173–182.
41. Huang, B.; Luo, J.; Du, W.; Pan, W.; Ming, Z. Cascaded Cross Attention for Review-based Sequential Recommendation. In Proceedings of the 2023 IEEE International Conference on Data Mining (ICDM), Shanghai, China, 1–4 December 2023; pp. 170–179.
42. Wu, S.; Tang, Y.; Zhu, Y.; Wang, L.; Xie, X.; Tan, T. Session-based recommendation with graph neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 346–353.
43. Beutel, A.; Covington, P.; Jain, S.; Xu, C.; Li, J.; Gatto, V.; Chi, E.H. Latent cross: Making use of context in recurrent recommender systems. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina del Rey, CA, USA, 5–9 February 2018; pp. 46–54.
44. Sabetzadeh, F.; Moaiyeri, M.H.; Ahmadijeh, M. A majority-based imprecise multiplier for ultra-efficient approximate image multiplication. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2019**, *66*, 4200–4208.
45. Wong, M.F.; Tan, C.W. Aligning crowd-sourced human feedback for reinforcement learning on code generation by large language models. *arXiv* **2024**, arXiv:2503.15129.

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.