Journal of Organizational and End User Computing
Volume 37 • Issue 1 • January-December 2025

Empirical Study on the Application of Deep Learning in User Behavior Prediction and Personalized Recommendation in E-Commerce

Guoyu Diao
https://orcid.org/0009-0001-8480-2988
Graduate School, Semyung University, Jecheon, South Korea

Chengcheng Li
Graduate School, Semyung University, Jecheon, South Korea

Qian Liu
School of Culture and Tourism, Zibo Normal College, Zibo, China

Zegang Liu
Graduate School, Gachon University, Seongnam, South Korea

ABSTRACT

Personalized recommendation systems are crucial for improving user experience and business performance in e-commerce. However, existing models face two major challenges: (a) an imbalance between short-term ranking accuracy and long-term engagement optimization, and (b) limited utilization of multi-modal information, resulting in suboptimal contextual understanding and poor cold-start performance. Traditional sequential models effectively capture short-term user preferences but fail to consider long-term engagement dynamics, while reinforcement learning (RL)-based approaches optimize engagement but suffer from high computational complexity and slow convergence. To address these challenges, we propose HDL-RecBERT, a hybrid recommendation framework that integrates transformer-based sequential modeling, RL for long-term optimization, adaptive multi-modal feature fusion, and contrastive self-supervised learning. The model employs self-attention mechanisms to model user behavior, RL to maximize cumulative user engagement, cross-modal attention to dynamically fuse multi-modal data, and contrastive learning to enhance cold-start recommendation performance. Extensive experiments on real-world e-commerce datasets show that HDL-RecBERT outperforms state-of-the-art baselines, achieving an 11.2% improvement in HR@10 and a 9.5% increase in NDCG@10, while RL improves cumulative reward (CR) by 10.2% and contrastive learning enhances cold-start recall by 9.3%. These results demonstrate HDL-RecBERT's ability to balance short-term and long-term optimization, improve recommendation diversity, and enhance adaptability in cold-start scenarios, making it a promising solution for next-generation recommendation systems.

KEYWORDS

Personalized Recommendation, Sequential Modeling, Reinforcement Learning, Multi-Modal Learning, Self-Supervised Learning

DOI: 10.4018/JOEUC.383512

INTRODUCTION

With the rapid expansion of e-commerce platforms, personalized recommendation systems have become essential for enhancing user experience and improving business performance (Qian & Wang, 2024; Wu et al., 2023). By analyzing user behaviors (Ko et al., 2022), these systems provide tailored product suggestions (Chiu et al., 2021; Ma et al., 2024), thereby increasing customer satisfaction and platform engagement (Li et al., 2024; Rane, 2023). Traditional recommendation models, including collaborative filtering (CF) and content-based filtering (CBF) (Fu & Ma, 2021), have demonstrated effectiveness in personalized recommendation tasks (Parthasarathy & Sathiya Devi, 2023). However, these conventional approaches often struggle with data sparsity (Mu & Wu, 2023), cold-start issues (Anwar et al., 2022), and the dynamic nature of user preferences (Chen et al., 2024). In response, deep learning-based models, particularly sequential recommendation models, have gained significant attention due to their ability to capture complex user behavior patterns and predict future interactions more accurately (Yoon & Jang, 2023).

Despite the advancements in deep learning-based recommendation systems, several key challenges persist. First, sequential recommendation models primarily focus on short-term user interactions (Hao et al., 2024; Wang et al., 2021), often neglecting long-term engagement optimization (Boka et al., 2024), which is crucial for sustained user retention (Ike et al., 2023; Zhang et al., 2025). While transformer-based models such as self-attentive sequential recommendation (SASRec) and BERT4Rec have significantly improved next-item prediction accuracy (Lashinin et al., 2024), they do not account for the long-term impact of recommendations on user engagement (Wang et al., 2023). Second, multi-modal information, including textual product descriptions, images, and structured metadata, remains underutilized (Zhang et al., 2024), limiting the contextual understanding of user preferences. Many existing methods rely heavily on user-item interaction data (Wu & Liu, 2024; Xu et al., 2025), which may not be sufficient for accurately modeling user intent (Li & Guenier, 2024; Zhang et al., 2023). Third, cold-start recommendation remains a fundamental challenge (Liu & She, 2024; Panda & Ray, 2022), as new users and items lack sufficient interaction history (Masenya, 2024; Yuan & Hernandez, 2023), making it difficult for traditional models to provide relevant recommendations.

Addressing these challenges requires a hybrid approach that effectively combines sequential learning, reinforcement learning (RL), and multi-modal feature integration. The primary research challenge lies in designing a unified recommendation framework that can not only capture sequential dependencies in user interactions but also optimize long-term engagement and leverage multi-modal information for better personalization. Moreover, incorporating self-supervised learning techniques is necessary to improve recommendation quality in cold-start scenarios. These aspects require a computationally efficient architecture that balances short-term accuracy with long-term optimization while ensuring adaptability across various e-commerce environments.

To overcome these limitations, we propose HDL-RecBERT, a hybrid deep learning and RL-based recommendation framework that effectively integrates transformer-based sequential modeling, RL optimization, multi-modal feature fusion, and self-supervised learning. The key insight behind HDL-RecBERT is that short-term user behavior can be captured effectively using transformer-based models, while long-term engagement can be optimized through RL. Additionally, multi-modal learning is incorporated to enhance recommendation quality by dynamically integrating textual, visual, and behavioral data through cross-modal attention mechanisms. Furthermore, contrastive self-supervised learning is employed to improve cold-start recommendation performance by pre-training user and item representations.

The proposed HDL-RecBERT framework consists of several key components. First, transformer-based sequential user behavior modeling captures short-term dependencies in user interactions using self-attention mechanisms, enhancing next-item prediction accuracy. Second, multi-modal feature fusion utilizes cross-modal attention to effectively combine textual, image, and structured metadata, enriching item representations and improving recommendation diversity. Third, RL-based long-term optimization incorporates policy gradient methods to maximize long-term user engagement and retention, ensuring that recommendations are beneficial over extended interactions rather than being solely optimized for immediate relevance. Lastly, self-supervised learning for cold-start recommendation employs contrastive learning techniques to enhance representations for new users and items, mitigating data sparsity issues and improving recommendation accuracy for cases where interaction history is limited.

The primary contributions of this study can be summarized as follows. We introduce HDL-RecBERT, a hybrid deep learning and RL-based recommendation framework that effectively balances short-term ranking accuracy with long-term engagement optimization. We propose a multi-modal feature fusion strategy using cross-modal attention, which dynamically integrates textual, visual, and structured metadata, improving recommendation diversity and contextual relevance. We incorporate RL for long-term engagement optimization, ensuring that recommendations are not only immediately relevant but also beneficial for sustained user interaction. We address cold-start recommendation challenges by leveraging self-supervised contrastive learning, significantly enhancing recommendation quality for new users and items. Finally, we conduct extensive experiments on real-world e-commerce datasets, demonstrating that HDL-RecBERT consistently outperforms state-of-the-art baselines in ranking accuracy, engagement optimization, and recommendation diversity.

In summary, HDL-RecBERT presents a novel hybrid approach that integrates sequential modeling, multi-modal learning, RL, and self-supervised learning into a unified recommendation system. The proposed framework effectively addresses existing limitations in e-commerce recommendation systems, paving the way for more adaptive, context-aware, and engagement-driven recommendation strategies.

RELATED WORK

Recommendation systems have undergone significant advancements with the evolution of deep learning techniques, leading to improvements in personalization, ranking accuracy, and user engagement. Early methods relied on CF and CBF, which, while effective, suffered from sparsity, cold-start issues, and an inability to model evolving user preferences. More recent approaches have leveraged deep neural networks, RL, and multi-modal learning, each addressing different aspects of the recommendation problem. However, challenges remain, particularly in balancing short-term ranking accuracy with long-term engagement, integrating heterogeneous data sources, and handling cold-start scenarios effectively.

This section provides an overview of sequential recommendation models, RL-based approaches, and multi-modal learning techniques, highlighting their strengths and limitations. Furthermore, we discuss how our proposed HDL-RecBERT framework builds upon and extends these existing methods to address key challenges in e-commerce recommendation systems.

Research Background and Current Developments

Sequential recommendation models have become a dominant approach for capturing user behavior patterns in e-commerce (Bai & Bai, 2024; Nasir et al., 2021). Traditional CF-based approaches, such as matrix factorization and neural collaborative filtering (NCF), rely on user-item interaction matrices to generate recommendations (Torkashvand et al., 2023). While these methods effectively model static user preferences (Chang et al., 2021; Xing et al., 2025), they fail to capture temporal dependencies in user behavior, limiting their applicability in dynamic environments. To address this issue, recurrent neural networks (RNNs) and long short-term memory networks were introduced to model time-dependent interactions (Mou et al., 2022), allowing for more adaptive recommendations. However, RNN-based models often suffer from vanishing gradient problems (Kumar & Kumar,

2024) and inefficient long-range dependency modeling (Huang et al., 2023), leading to performance degradation when user interaction sequences become longer.

The introduction of transformer-based architectures, such as SASRec and BERT4Rec, has significantly improved sequential recommendation performance by leveraging self-attention mechanisms to model long-range dependencies without the recurrence constraints of RNNs (de Souza Pereira Moreira et al., 2021; Lu & Ouyang, 2024). These models have demonstrated state-of-the-art performance in next-item prediction tasks, effectively capturing contextual dependencies and short-term user intent. However, despite their success in improving ranking accuracy, they fail to account for long-term user engagement and personalized reward optimization, as they primarily focus on immediate interactions rather than the cumulative impact of recommendations over time.

In addition to sequential modeling, RL-based recommendation approaches have emerged as a solution to long-term engagement optimization. Methods such as deep Q-networks-based and policy gradient-based recommendation frameworks have been proposed to formulate the recommendation task as a Markov decision process. These approaches optimize recommendation strategies based on reward signals derived from user feedback, allowing models to learn policies that maximize long-term user engagement rather than only short-term relevance. While RL-based methods have been successful in optimizing sequential decision-making, they often require extensive reward tuning and suffer from high sample complexity, making them challenging to deploy in real-world settings.

Another major development in recommendation systems is multi-modal learning, which integrates diverse information sources such as textual product descriptions, images, and structured metadata. Many existing works have leveraged pretrained language models (e.g., BERT, GPT) and convolutional neural networks to extract high-level representations from product descriptions and images, respectively. Models such as MMDL-MR and CBIR-DSS have demonstrated that incorporating multi-modal information significantly enhances recommendation diversity and accuracy, especially in cases where user-item interaction data is sparse. However, most existing multi-modal recommendation models lack dynamic feature weighting mechanisms, treating all modalities equally rather than adapting to user preferences in different contexts.

While each of these approaches has contributed to advancing recommendation systems, they often operate in isolation, optimizing either short-term ranking accuracy (via sequential modeling), long-term engagement (via RL), or multi-modal representation learning. However, real-world recommendation systems require a unified framework that effectively integrates these components, ensuring that recommendations are both immediately relevant and beneficial for sustained user interactions. This gap in existing research motivates the development of HDL-RecBERT, which combines transformer-based sequential learning, RL optimization, and cross-modal feature fusion to create a comprehensive and adaptive recommendation framework.

Challenges

Despite the significant advancements in sequential modeling, RL, and multi-modal learning, several fundamental challenges persist in e-commerce recommendation systems. Existing methods primarily optimize for either short-term ranking accuracy or long-term engagement, failing to balance immediate relevance with sustained user interaction. Additionally, the integration of multi-modal features and cold-start adaptability remain underdeveloped in many contemporary models. The following challenges highlight key obstacles that need to be addressed for a more effective and adaptive recommendation system.

The first major challenge is balancing short-term ranking accuracy with long-term engagement optimization. Traditional sequential recommendation models (e.g., SASRec, BERT4Rec) are designed to predict the next item a user is likely to interact with, using self-attention mechanisms to capture sequential dependencies. While these models achieve high next-item prediction accuracy, they fail to account for the cumulative impact of recommendations over time, potentially leading to user disengagement if recommendations are not strategically optimized. In contrast, RL-based recommendation models (e.g., RLRS) optimize for long-term user engagement by learning policies that maximize future rewards. However, these models often suffer from high computational complexity, slow convergence, and the need for extensive reward tuning, making them difficult to deploy at scale. A unified approach is needed to leverage the strengths of both sequential learning and RL to ensure that recommendations are immediately useful while contributing to long-term user satisfaction.

The second challenge is effectively integrating multi-modal information for more context-aware recommendations. Many products in e-commerce platforms include text descriptions, images, and structured metadata, which provide valuable contextual signals. While some existing methods have incorporated multi-modal data, they often treat all features equally without dynamically adapting to user preferences. Models such as MMDL-MR use static fusion strategies, where text and image features are combined in a fixed manner. However, different users prioritize different modalities based on their shopping behavior—some may rely more on textual descriptions, while others may focus on product images. The challenge lies in developing an adaptive feature fusion mechanism that dynamically assigns importance to different modalities based on user interactions and context.

Another critical challenge is handling the cold-start problem for new users and items. Recommendation systems rely on past interactions to make predictions, but new users and newly introduced products have little to no interaction history, making it difficult to generate relevant recommendations. Traditional CF methods fail in this scenario, as they require sufficient user-item interaction data. While content-based approaches leverage product descriptions and metadata to recommend new items, they lack personalization because they do not incorporate user preferences learned from historical data. Recent advancements in self-supervised learning and contrastive learning have shown promise in improving recommendation quality in cold-start settings by learning generalized item and user representations. However, existing self-supervised learning approaches primarily focus on single-modality representations, and few have explored multi-modal contrastive learning for cold-start recommendations. The challenge is to design a contrastive learning framework that effectively learns from multi-modal data while ensuring that new users and items receive personalized and relevant recommendations.

Finally, there is the challenge of efficient training and scalability in large-scale e-commerce environments. Many advanced deep learning-based recommendation models require substantial computational resources due to the use of self-attention mechanisms, RL optimization, and multi-modal feature extraction. While Transformer-based models improve sequential modeling, they are computationally expensive, especially for long interaction sequences. Reinforcement learning further exacerbates this issue, as it requires large-scale reward computation and policy updates, leading to high training costs and slow convergence. To ensure practical deployment in real-world applications, an efficient training strategy that balances model complexity with computational feasibility is necessary.

In summary, recommendation systems must address four key challenges: (a) balancing short-term accuracy with long-term engagement, (b) effectively integrating multi-modal information, (c) mitigating the cold-start problem for new users and items, and (d) ensuring computational efficiency and scalability. The limitations of existing methods in tackling these challenges motivate the development of HDL-RecBERT, a hybrid recommendation framework that integrates sequential learning, RL, multi-modal fusion, and self-supervised learning to enhance personalization, engagement, and adaptability in large-scale e-commerce environments.

HDL-RecBERT

To address the challenges outlined in the previous section, we propose HDL-RecBERT, a hybrid recommendation framework that integrates transformer-based sequential modeling, RL for long-term optimization, adaptive multi-modal feature fusion, and contrastive self-supervised learning for cold-start recommendation. This framework is designed to simultaneously capture short-term user behavior, optimize long-term engagement, dynamically incorporate multi-modal information, and improve cold-start adaptability, ensuring a more effective and scalable recommendation system for e-commerce platforms.

The core innovation of HDL-RecBERT lies in its ability to balance short-term recommendation accuracy with long-term engagement optimization. Unlike conventional sequential models such as SASRec and BERT4Rec, which focus solely on immediate next-item prediction, HDL-RecBERT integrates RL-based reward optimization, enabling it to learn recommendation policies that maximize long-term user engagement. Unlike existing RL-based models that suffer from slow convergence and complex reward tuning, HDL-RecBERT efficiently combines policy learning with transformer-based self-attention mechanisms, ensuring both fast adaptation to user behavior changes and improved ranking precision.

A key differentiating factor of HDL-RecBERT is its adaptive multi-modal feature fusion strategy, which overcomes the limitations of static fusion methods used in prior multi-modal recommendation models. Existing multi-modal approaches often treat different content types equally, failing to account for the varying importance of text, images, and structured metadata in different contexts. HDL-RecBERT employs cross-modal attention mechanisms that dynamically adjust the contribution of textual descriptions, product images, and structured attributes, ensuring that recommendations are context-aware and personalized based on individual user preferences. This approach enhances recommendation diversity and accuracy, particularly in cases where user-item interaction data is sparse.

To mitigate cold-start challenges, HDL-RecBERT incorporates contrastive self-supervised learning, allowing the model to generate robust user and item representations even with limited interaction history. Unlike traditional content-based filtering approaches that rely solely on predefined item attributes, HDL-RecBERT leverages contrastive learning objectives to pre-train embeddings from multi-modal data, ensuring that new users and items receive more personalized recommendations from the outset. Unlike existing contrastive learning-based recommendation models that focus on single-modality representations, HDL-RecBERT extends this idea to multi-modal embeddings, significantly improving cold-start performance.

Another major strength of HDL-RecBERT is its computational efficiency and scalability, making it suitable for real-world deployment. While transformer-based models often suffer from high memory and computational costs, HDL-RecBERT incorporates efficient sequence truncation and self-attention mechanisms to reduce computational complexity without sacrificing accuracy. Additionally, by integrating RL in a sample-efficient manner, the model avoids the extensive reward tuning and exploration inefficiencies that often limit RL-based recommendation systems. These optimizations enable HDL-RecBERT to scale effectively in large-scale e-commerce environments while maintaining high recommendation quality.

In summary, HDL-RecBERT presents a novel and comprehensive approach to e-commerce recommendation by combining sequential modeling, RL, multi-modal feature fusion, and self-supervised learning in a unified framework. Compared to existing models, it effectively balances short-term and long-term optimization, dynamically integrates multi-modal content, enhances cold-start adaptability, and maintains computational efficiency, making it a robust, scalable, and high-performing recommendation system. The following sections provide a detailed overview of the methodology and experimental results demonstrating the advantages of HDL-RecBERT over state-of-the-art baselines.

METHODOLOGY

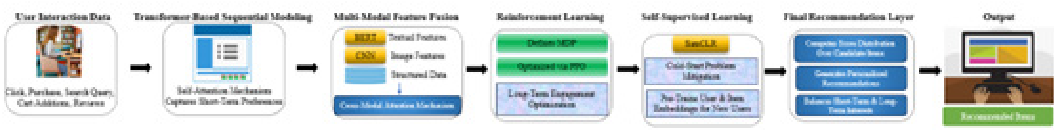
In this section, we introduce HDL-RecBERT, a novel hybrid recommendation framework designed to enhance user behavior prediction and personalized recommendation in e-commerce. The proposed model integrates Transformer-based sequential modeling, RL, and multi-modal learning to optimize both short-term user behavior prediction and long-term engagement strategies.

HDL-RecBERT consists of four major components:

1. transformer-based sequential modeling for short-term user behavior prediction
2. multi-modal feature fusion for enhanced recommendation precision
3. reinforcement learning optimization for long-term engagement maximization
4. self-supervised learning for cold-start handling

Each module plays a crucial role in improving recommendation quality, adaptability, and robustness. Below, we describe each component in detail along with the mathematical formulations.

Figure 1. The Model Diagram of HDL-RecBERT



Transformer-Based Sequential User Behavior Modeling

To effectively capture short-term user preferences, we employ a transformer-based sequential recommendation model inspired by SASRec and BERT4Rec. This component models user interactions as a sequence and applies self-attention mechanisms to capture long-term dependencies.

User Interaction Sequence Modeling

Let u be a user and $S_u = \{i_1, i_2, \ldots, i_n\}$ be the sequence of items that the user has interacted with, where i_t represents the item the user engaged with at time t. The goal of this component is to predict the next item i_{t+1} that the user is likely to interact with, given their past behavior.

We define the embedding matrix $E \in \mathbb{R}^{|I| \times d}$, where $|I|$ is the total number of items, and d is the embedding dimension. Each item i_t is mapped to its corresponding embedding $e_t = E(i_t)$, forming the input sequence:

$$X_u = [e_1, e_2, \ldots, e_n]$$

To incorporate temporal information, we use positional encoding P as follows:

$$\widehat{X}_u = X_u + P$$

where P is a sinusoidal position encoding that allows the model to differentiate item positions within the sequence.

Self-Attention for Sequential Modeling

To capture dependencies among past user interactions, we apply multi-head self-attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$$

where:

- $Q = W_Q \widehat{X}_u$ (query matrix),
- $K = W_K \widehat{X}_u$ (key matrix),
- $V = W_V \widehat{X}_u$ (value matrix),

- W_Q, W_K, W_V are learnable weight matrices.

The final user representation is obtained through a stack of self-attention layers and feed-forward layers:

$$H = \text{LayerNorm}\left(\text{FFN}\left(\text{Attention}\left(\hat{X}_u\right)\right)\right)$$

where FFN is a position-wise feed-forward network.
The prediction layer then outputs the probability distribution over candidate items:

$$\hat{y} = \text{softmax}\left(W_O H + b\right)$$

where W_O and b are learnable parameters.

Multi-Modal Feature Fusion

E-commerce recommendations often rely on textual product descriptions, images, and structured behavioral data. To integrate these diverse modalities, we employ cross-modal attention mechanisms.

Feature Representation

Each product i has the following features:

- textual features (T_i) derived from BERT embeddings: $E_T(i)$
- image features (I_i) extracted using CNN-based embeddings: $E_I(i)$
- structured features (F_i), such as price, category, etc.

Thus, the unified feature representation is:

$$Z_i = W_T E_T(i) + W_I E_I(i) + W_F F_i$$

where W_T, W_I, W_F are learnable weights.

Cross-Modal Attention

To enhance interaction among modalities, we apply a cross-modal attention mechanism:

$$\alpha_T = \frac{\exp\left(f_T\left(Z_i, Z_u\right)\right)}{\sum_{j \in M} \exp\left(f_T\left(Z_j, Z_u\right)\right)}$$

where:

- $f_T\left(Z_i, Z_u\right) = W_T^T\left(Z_i \cdot Z_u\right)$ measures relevance between user and item features.
- α_T is the attention score for textual features (similar equations apply for image and structured features).

The final multi-modal representation for user uu is:

$$Z_u^* = \sum_{m \in \{T,I,F\}} \alpha_m Z_m$$

Reinforcement Learning-Based Long-Term Optimization

Traditional recommendation models primarily focus on optimizing short-term ranking metrics, such as hit rate (HR@K) and normalized discounted cumulative gain (NDCG@K), which measure the immediate relevance of recommendations. However, in real-world e-commerce applications, maximizing long-term user engagement is crucial for sustained interaction and retention. To address this challenge, we introduce RL into HDL-RecBERT, allowing the model to optimize recommendations based on long-term user engagement signals rather than just immediate interactions.

Markov Decision Process Formulation

We define the recommendation process as a Markov decision process, represented as a tuple (S, A, P, R, γ):

- S (state space): the user state s_t at time step t, which encodes the user's historical interactions and contextual information
- A (action space): the set of recommended items a_t at time t
- $P(s'|s,a)$ (state transition probability): the probability of transitioning to a new user state s' after recommending item a
- $R(s,a)$ (reward function): the immediate reward received when the user interacts with the recommended item
- γ (discount factor): A factor that determines the importance of future rewards

Reward Function Design

The reward function is designed to encourage engagement-driven recommendations, considering multiple user interaction signals:

$$R\left(s_t, a_t\right) = \alpha \cdot r_{\text{click}} + \beta \cdot r_{\text{purchase}} + \delta \cdot r_{\text{engagement}}$$

where:

- r_{click} is a binary reward (1 if the user clicks on the recommendation, 0 otherwise).
- r_{purchase} is a weighted reward based on whether the user purchases the recommended item.
- $r_{\text{engagement}}$ is a normalized score based on session length, dwell time, or repeated interactions.

The parameters α, β, δ are hyperparameters controlling the relative importance of each reward signal.

Policy Optimization

We use policy gradient methods to optimize the recommendation policy $\pi(a|s)$, which determines the probability of recommending an item given the user's state. The objective function is:

$$J\left(\theta\right) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{T} \gamma^t R\left(s_t, a_t\right)\right]$$

where:

- θ represents the parameters of the policy network.
- T is the episode length.

- γ is the discount factor for future rewards.

To update the policy, we compute the policy gradient using the REINFORCE algorithm:

$$\nabla_\theta J(\theta) = \mathbb{E}\left[\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(a_t|s_t) \cdot G_t\right]$$

where G_t is the return (cumulative discounted reward from step t).

To further stabilize learning, we employ an advantage function to reduce variance in the gradient estimation:

$$A(s_t, a_t) = G_t - V(s_t)$$

where $V(s_t)$ is the baseline value function estimating the expected return from state s_t.

The final policy update equation using advantage-weighted learning is:

$$\nabla_\theta J(\theta) = \mathbb{E}\left[\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(a_t|s_t) \cdot A(s_t, a_t)\right]$$

This formulation ensures that actions leading to higher-than-expected engagement are reinforced, while suboptimal actions are penalized, optimizing the recommendation model for long-term user satisfaction.

To ensure stability in policy updates, we integrate a baseline function $V(s_t)$ not only to reduce variance but also to prevent drastic fluctuations in the learned policy. This encourages consistent convergence during training. Moreover, HDL-RecBERT applies an ϵ-greedy exploration strategy to balance exploration and exploitation. With a small probability ϵ, the model chooses random actions to explore new recommendation patterns, while with probability $1 - \epsilon$, it selects the best-known action to maximize reward. The exploration rate is gradually decayed to emphasize exploitation as training progresses.

We also considered off-policy RL methods such as deep deterministic policy gradient and soft actor-critic, which offer better sample efficiency and allow learning from a replay buffer. However, due to the discrete and large-scale nature of the item space in e-commerce and the fast-changing user behavior, on-policy methods using policy gradient were found more suitable. They enable more direct and adaptable learning, especially in recommendation scenarios where delayed rewards and non-stationarity make off-policy learning more complex and unstable.

Self-Supervised Learning for Cold-Start Problem

Cold-start recommendation remains a persistent challenge, as new users and newly introduced items lack sufficient interaction data. To address this issue, we employ contrastive self-supervised learning to enhance user and item representations, enabling more effective cold-start recommendations.

Contrastive Learning for Representation Learning

Contrastive learning aims to learn representations by maximizing similarity between positive pairs while minimizing similarity between negative pairs. Given an item i, we construct positive and negative samples as follows:

- positive pair (i, i^+): items with similar content features (e.g., products in the same category)
- negative pair (i, i^-): items with distinct features (e.g., products from different categories)

The contrastive loss function is defined as:

$$\mathcal{L}_{\text{contrastive}} = -\sum_{i \in \mathcal{B}} \log \frac{\exp\left(\text{sim}\left(z_i, z_{i^+}\right)/\tau\right)}{\sum_{j \in \mathcal{B}} \exp\left(\text{sim}\left(z_i, z_j\right)/\tau\right)}$$

where:

- z_i and z_{i^+} are the representations of the anchor item and its positive sample.
- τ is a temperature scaling parameter.
- $\text{sim}\left(z_i, z_j\right)$ measures the cosine similarity between embeddings.

This self-supervised objective allows the model to learn more generalizable item representations, improving performance when limited interaction data is available.

Multi-Modal Contrastive Pretraining

To enhance cold-start adaptability, we extend contrastive learning to multi-modal features:

$$\mathcal{L}_{\text{multi-modal}} = \mathcal{L}_{\text{contrastive}}^{\text{text}} + \mathcal{L}_{\text{contrastive}}^{\text{image}} + \mathcal{L}_{\text{contrastive}}^{\text{metadata}}$$

where different contrastive loss terms are computed for textual embeddings, image embeddings, and structured metadata. This formulation ensures that the model effectively learns from multiple information sources, improving its ability to recommend items even in cold-start scenarios.

Overall Optimization Objective

HDL-RecBERT jointly optimizes short-term ranking accuracy, long-term engagement, and representation learning using a multi-objective loss function:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{rec}} + \lambda_2 \mathcal{L}_{\text{RL}} + \lambda_3 \mathcal{L}_{\text{contrastive}}$$

where:

- $\mathcal{L}_{\text{rec}}$ is the ranking loss for next-item prediction:

$$\mathcal{L}_{\text{rec}} = -\sum_{(u,i)} y_{ui} \log \hat{y}_{ui} + \left(1 - y_{ui}\right) \log\left(1 - \hat{y}_{ui}\right)$$

- $\mathcal{L}_{\text{RL}}$ is the RL loss for engagement optimization:

$$\mathcal{L}_{\text{RL}} = -\sum_t \log \pi_\theta\left(a_t \middle| s_t\right) A\left(s_t, a_t\right)$$

- $\mathcal{L}_{\text{contrastive}}$ is the self-supervised contrastive loss for representation learning.

The hyperparameters $\lambda_1, \lambda_2, \lambda_3$ control the relative importance of each objective. By optimizing this combined loss, HDL-RecBERT simultaneously enhances next-item prediction accuracy, maximizes long-term user engagement, and improves cold-start recommendation performance, making it a highly effective and adaptive recommendation framework.

EXPERIMENT

In this section, we conduct comprehensive experiments to evaluate the effectiveness of HDL-RecBERT in user behavior prediction and personalized recommendation. The evaluation covers multiple aspects, including short-term ranking accuracy, long-term engagement, multi-modal learning, cold-start performance, and recommendation diversity. We compare HDL-RecBERT against state-of-the-art baselines, describe the experimental setup and hyperparameter configurations, and analyze the results through main experiments, ablation studies, and significance tests to validate the contributions of key model components.

Datasets

To evaluate the effectiveness of the proposed HDL-RecBERT model, we utilize four real-world e-commerce datasets that encompass diverse aspects of user interactions, product attributes, and multi-modal information. These datasets are selected to ensure comprehensive coverage of user behavior prediction and personalized recommendation tasks, allowing for both short-term and long-term user engagement modeling. The chosen datasets reflect a variety of online shopping behaviors, ranging from product search and browsing patterns to purchase decisions and user-generated reviews.

The first dataset, Real-Scenario Multimodal Retrieval Dataset from Taobao, consists of extensive user behavior records collected from Taobao's e-commerce platform. It includes user clicks, purchases, and product interactions, along with multi-modal features such as product images and textual descriptions. The inclusion of multi-modal information allows for the evaluation of recommendation models that integrate visual and textual data to improve prediction accuracy. This dataset is particularly valuable for assessing the effectiveness of the multi-modal learning components within HDL-RecBERT.

The second dataset, AliExpress Searching System Dataset, captures search query interactions and clickstream data from AliExpress users. This dataset contains query-product pairs, user feedback signals (such as clicks and conversions), and associated metadata. It is instrumental in sequential modeling for search-based recommendations and is used to validate the transformer-based user behavior modeling in HDL-RecBERT. The dataset also helps assess how well the model adapts to session-based recommendations, where user preferences evolve dynamically within short timeframes.

The third dataset, RetailRocket Recommender System Dataset, includes detailed event logs of user interactions, such as page views, cart additions, and completed purchases. The dataset is well-structured for studying user decision-making processes, making it highly relevant for RL-based recommendation approaches. Given its event-driven nature, this dataset provides a robust benchmark for evaluating long-term user engagement strategies, particularly the RL component of HDL-RecBERT.

The fourth dataset, Amazon Reviews'23, is a large-scale dataset compiled by McAuley Lab in 2023. It comprises user reviews, product ratings, helpfulness votes, product descriptions, and images, making it a rich multi-modal dataset. This dataset is crucial for evaluating the sentiment-aware and content-driven recommendation capabilities of HDL-RecBERT. Furthermore, the review-based nature of the dataset enables the study of cold-start recommendation scenarios, where new users or products lack sufficient interaction history.

For experimental evaluation, each dataset is split into training, validation, and test sets to ensure fair benchmarking. The training set comprises 80% of the data, used for model optimization. The validation set, which accounts for 10% of the data, is used for hyperparameter tuning and performance fine-tuning. The remaining 10% serves as the test set, where final model performance is evaluated. The data split ensures that the model is trained on a representative portion of user interactions while maintaining a sufficiently diverse test set for generalization assessment.

By leveraging these four datasets, we ensure that HDL-RecBERT is rigorously tested across various e-commerce scenarios, including multi-modal recommendation, session-based search recommendations, long-term engagement modeling, and cold-start problem resolution. The diversity of these datasets strengthens the empirical evaluation of our proposed model, demonstrating its applicability across different recommendation tasks in real-world e-commerce settings.

Evaluation Metrics

To comprehensively assess the effectiveness of HDL-RecBERT, we employ a range of evaluation metrics that cover various aspects of recommendation performance. These metrics are categorized into four major groups:

- ranking-based metrics: assess the model's ability to rank relevant items higher than irrelevant ones.
- prediction accuracy metrics: evaluate the correctness of the recommended items.
- long-term engagement metrics: measure the effectiveness of the model in sustaining user engagement.
- diversity and novelty metrics: ensure that the recommendations are diverse and introduce new items to the user.

The selected metrics allow for a comprehensive comparison of HDL-RecBERT against the baseline models, ensuring fair benchmarking across different recommendation scenarios.

Ranking-Based Metrics

HR@K. This measures the percentage of test cases where the relevant item appears within the top K recommended items. Given a user u with an interaction history S_u, the model generates a ranked list of recommendations R_u, and HR@K is defined as:

$$HR@K = \frac{1}{|U|}\sum_{u \in U}\mathbb{I}\left(i_u \in R_u^K\right)$$

where $\mathbb{I}(\cdot)$ is an indicator function that equals 1 if the ground truth item i_u is within the top K recommended items, and 0 otherwise. A higher HR@K value indicates better ranking effectiveness.

NDCG@K. This evaluates the ranking quality by assigning higher importance to correctly recommended items appearing at the top of the ranked list. It is computed as:

$$NDCG@K = \frac{1}{|U|}\sum_{u \in U}\frac{DCG_u@K}{IDCG_u@K}$$

where:

$$DCG_u@K = \sum_{j=1}^{K}\frac{\mathbb{I}\left(i_u = R_u^j\right)}{\log_2(j+1)}$$

and $IDCG_u@K$ is the ideal DCG, representing the best possible ranking order. A higher NDCG@K indicates better ranking performance.

Prediction Accuracy Metrics

Precision@K. IPrecision measures the proportion of recommended items that are relevant to the user:

$$Precision@K = \frac{1}{|U|}\sum_{u \in U}\frac{\left|R_u^K \cap I_u\right|}{K}$$

where I_u is the set of items the user actually interacted with. A higher Precision@K indicates a more precise recommendation system.

Recall@K. Recall measures the proportion of relevant items that are successfully recommended:

$$Recall@K = \frac{1}{|U|}\sum_{u\in U}\frac{|R_u^K \cap I_u|}{|I_u|}$$

A higher Recall@K indicates that the model successfully retrieves more relevant items.

Long-Term Engagement Metrics

Cumulative Reward. In RL-based recommendation systems, the cumulative reward (CR) assesses the effectiveness of long-term optimization. The reward r_t for user u at time t is computed as:

$$CR = \frac{1}{|U|}\sum_{u\in U}\sum_{t=1}^{T}\gamma^t r_t$$

where:

- γ is the discount factor (typically 0.950.95).
- r_t may be defined as click-through rate (CTR), purchase conversion rate, or user engagement duration.

A higher CR score indicates better RL optimization.

Session Engagement Score. To measure long-term user engagement, we define session engagement score (SES) as:

$$SES = \frac{1}{|U|}\sum_{u\in U}\frac{\sum_{t=1}^{T}\mathbb{I}(i_t \text{ clicked or purchased})}{T}$$

where T represents the total interactions per session. A higher SES value indicates better user retention and engagement.

Diversity and Novelty Metrics

Item Coverage. Item coverage (IC) measures the proportion of unique items recommended across all users:

$$IC = \frac{|\bigcup_{u\in U} R_u^K|}{|I|}$$

where I is the total number of available items. A higher IC indicates that the recommendation model does not overfit to a small set of popular items.

Novelty Score. To prevent recommending frequently suggested items, we measure novelty by penalizing recommendations of highly popular items:

$$NS = \frac{1}{|U|}\sum_{u\in U}\sum_{i\in R_u^K} -\log_2\left(\frac{|U_i|}{|U|}\right)$$

where $|U_i|$ is the number of users who have interacted with item ii. A higher novelty score indicates that the model recommends less popular but more personalized items.

Baseline Models

To comprehensively evaluate the performance of the proposed HDL-RecBERT model, we compare it against state-of-the-art deep learning-based recommendation models. These baselines have been carefully selected to cover different methodological paradigms, including sequential user behavior modeling, CF, multi-modal recommendation, and RL-based optimization.

The baseline models were chosen based on the following considerations:

- sequential user behavior prediction: models that leverage historical user interactions to predict future behaviors
- collaborative filtering-based recommendation: models that utilize user-item interaction matrices for personalized recommendations
- multi-modal recommendation systems: methods that integrate different data modalities, such as text, images, and structured behavioral data
- reinforcement learning-based recommendation: models that optimize long-term user engagement using RL

Each selected baseline represents a well-established approach in personalized recommendation and serves as a comparative benchmark for HDL-RecBERT.

- KAN-QResNet: High-dimensional feature learning for user behavior prediction (de Souza Pereira Moreira et al., 2021). KAN-QResNet is a high-dimensional feature learning model that applies Kolmogorov-Arnold networks (KANs) to capture complex non-linear user behavior patterns in large-scale e-commerce environments. By integrating quadratic residual networks (QResNet), this model enhances feature representation and interaction modeling, making it particularly effective for high-dimensional data processing in e-commerce applications.
- DBS-PR: Deep behavioral sequence modeling for personalized recommendation (Deng, 2024**).** DBS-PR employs deep neural networks to model sequential dependencies in user behavior, capturing long-term user interests and item interaction patterns. It is particularly effective for time-sensitive personalized recommendations, leveraging hierarchical sequence modeling.
- NCF-CE: Neural collaborative filtering for personalized recommendation (Khatter et al., 2021). NCF-CE extends traditional collaborative filtering by employing NCF, allowing the model to learn complex user-item interaction patterns. Unlike conventional matrix factorization techniques, NCF-CE integrates deep learning to enhance user preference modeling and optimize recommendation performance.
- MMDL-MR: Multi-modal deep learning for recommendation (Malitesta et al., 2024). MMDL-MR integrates multiple data modalities, including textual product descriptions, images, and behavioral data, to enhance recommendation accuracy. This model employs cross-modal attention mechanisms to effectively learn user preferences from diverse sources, making it particularly well-suited for e-commerce scenarios where multiple content types influence purchase decisions.
- Reinforcement learning-based recommender system (RLRS) (Afsar et al., 2022) RLRS applies deep RL to optimize recommendation strategies through iterative learning based on user feedback. Unlike traditional supervised learning approaches, RLRS continuously refines its recommendations by maximizing long-term user engagement.
- SASRec (Klenitskiy & Vasilev, 2023). SASRec is a transformer-based sequential recommendation model that applies self-attention mechanisms to model short-term user behavior patterns. It captures temporal dependencies in user interactions and is particularly effective for recommending items that align with a user's most recent actions.

The selected baselines comprehensively cover the key paradigms in deep learning-based recommendation systems, including sequential user behavior modeling, collaborative filtering, multi-modal learning, and RL-based optimization. By comparing HDL-RecBERT against these models, we aim to demonstrate its advantages in terms of short-term prediction accuracy, long-term recommendation effectiveness, and adaptability to diverse e-commerce data.

Experimental Setup

To ensure a fair and reproducible evaluation of HDL-RecBERT, we conduct all experiments in a controlled computational environment with carefully tuned hyperparameters. This section provides an overview of the hardware and software configurations, as well as the hyperparameter settings used in training and evaluating the model.

Computational Environment

All experiments are conducted on a high-performance computing server equipped with the following specifications:

- GPU: NVIDIA A100 with 40GB VRAM
- CPU: Intel Xeon Platinum 8358 with 32 cores
- Memory: 256GB RAM
- Storage: 2TB NVMe SSD
- Operating System: Ubuntu 20.04 LTS
- Deep Learning Frameworks:
    - PyTorch 2.0
    - TensorFlow 2.10 (for auxiliary tasks)
    - CUDA 11.8 with cuDNN acceleration

The use of GPU acceleration significantly improves training efficiency, especially for transformer-based models that rely on extensive matrix operations.

Hyperparameter Settings

The hyperparameters are carefully tuned based on validation set performance to ensure the model achieves optimal generalization.

- embedding dimension: The item and user embeddings are set to 128 dimensions, which provides a balance between representation capacity and computational efficiency.
- number of attention heads: The self-attention layers in the transformer module utilize four heads, ensuring a strong balance between complexity and model expressiveness.
- number of transformer layers: The transformer encoder consists of two layers, as deeper architectures did not yield significant performance gains in preliminary experiments.
- dropout rate: A dropout rate of 0.2 is applied to prevent overfitting in both the transformer and feed-forward layers.
- batch size: The model is trained with a batch size of 256 to accommodate large-scale e-commerce datasets while maintaining computational efficiency.
- optimizer: AdamW optimizer is used, as it provides superior convergence properties for deep learning models with adaptive learning rate adjustments.
- learning rate: The initial learning rate is set to 0.0005, and a cosine annealing scheduler is applied to gradually decay the learning rate over training epochs.
- gradient clipping: A gradient norm threshold of 5.0 is enforced to stabilize training and prevent exploding gradients in Transformer layers.

- sequence length: The maximum user interaction sequence length is set to 50, ensuring that both long- and short-term user preferences are captured effectively.
- reinforcement learning discount factor: For long-term reward optimization, the RL discount factor γ is set to 0.95, prioritizing long-term engagement while still considering recent interactions.
- exploration strategy in RL: A ϵ-greedy policy is adopted with an initial exploration rate of 0.1, gradually decaying to 0.01 during training to ensure a balance between exploration and exploitation.
- contrastive learning temperature parameter: For self-supervised learning in cold-start scenarios, the temperature parameter τ is set to 0.07, which helps in learning discriminative representations for new users and items.

Training Procedure

The model is trained using mixed-precision training to reduce GPU memory consumption and accelerate computations. Each experiment runs for 50 epochs, with early stopping applied if validation loss does not improve for 10 consecutive epochs. The model with the highest validation NDCG@10 is selected as the final checkpoint for testing.

All experiments are repeated three times, and the average performance is reported to ensure the stability and reproducibility of results.

Main Experiments

To comprehensively evaluate the effectiveness of HDL-RecBERT, we conduct a series of controlled experiments designed to assess its performance across various recommendation tasks. The evaluation focuses on measuring the model's capability in sequential user behavior prediction, personalized recommendation, long-term engagement optimization, and cold-start handling. The experimental procedure consists of multiple steps, ensuring a rigorous and fair comparison with baseline models.

The experiment is structured into four main stages: data preprocessing, model training, hyperparameter tuning, and performance evaluation. Each stage is described below in detail.

Step 1: Data Preprocessing

Before training, all datasets undergo a standardized preprocessing pipeline to ensure consistency across different data sources:

- user interaction filtering: only users with at least five historical interactions are retained to ensure a meaningful sequence representation.
- item frequency thresholding: items that appear in fewer than 10 interactions are removed to prevent extreme sparsity issues.
- Multi-modal feature extraction:
    - textual features: product descriptions and reviews are embedded using a pre-trained BERT model.
    - visual features: product images are processed using a pre-trained ResNet50 model to extract high-level representations.
    - structured features: price, category, and historical engagement metrics are normalized and concatenated with learned embeddings.
- sequence padding and truncation: each user's interaction sequence is truncated to a maximum length of 50 and padded where necessary.
- data splitting: the dataset is divided into 80% training, 10% validation, and 10% test sets, ensuring that no future interactions are leaked into earlier training phases.

Step 2: Model Training

The training process is conducted in a mini-batch stochastic optimization setting with backpropagation and gradient descent. The model is optimized using the AdamW optimizer with a learning rate of 0.0005, and cosine annealing decay is applied over training epochs. The training follows these key principles:

- sequential modeling training: the transformer-based component is first pre-trained to capture short-term user behavior patterns.
- multi-modal feature fusion training: the cross-modal attention module is jointly optimized with the user-item interaction module to enhance feature fusion.
- reinforcement learning fine-tuning: once the supervised learning stage is complete, policy gradient RL is applied to refine long-term engagement strategies.
- self-supervised pretraining for cold-start users: to improve cold-start performance, user and item representations are pretrained using contrastive learning before fine-tuning on downstream recommendation tasks.

During training, early stopping is applied if the validation loss does not improve for 10 consecutive epochs, ensuring efficiency and preventing overfitting.

Step 3: Hyperparameter Tuning

The model undergoes systematic hyperparameter tuning using grid search on the validation set. The following hyperparameters are optimized:

- number of transformer layers: varying from 1 to 4, with two layers found to be optimal.
- dropout rate: experimented in the range of 0.1 to 0.5, with 0.2 yielding the best balance between generalization and training stability.
- batch size: evaluated between 128 and 512, with 256 achieving the best trade-off between computational efficiency and convergence stability.
- discount factor γ for RL: tuned in the range 0.85 to 0.99, with 0.95 optimizing long-term engagement.
- contrastive learning temperature parameter τ: tested between 0.05 and 0.1, with 0.07 leading to more discriminative representations.

Step 4: Performance Evaluation

After training, HDL-RecBERT is evaluated on the test set, using the previously described ranking-based, prediction accuracy, long-term engagement, and diversity metrics. The evaluation process is structured as follows:

- top-K recommendation performance: the model's ability to return relevant recommendations is assessed using HR@K, NDCG@K, Precision@K, and Recall@K.
- long-term optimization: the RL effectiveness is quantified using CR and SES.
- diversity and novelty: the capability to generate diverse and novel recommendations is measured using IC and NS.
- cold-start analysis: a specialized subset of new users and items is extracted to test how well HDL-RecBERT generalizes to unseen interactions.

All evaluation metrics are averaged over three independent runs, and statistical significance tests (paired t-tests) are performed to validate the robustness of the findings.

By following this experimental setup, we ensure a systematic and rigorous evaluation of HDL-RecBERT, providing strong empirical evidence of its effectiveness in real-world e-commerce recommendation scenarios.

Results and Analysis

This section presents the experimental results of HDL-RecBERT in various recommendation scenarios, focusing on short-term user behavior prediction, long-term engagement optimization, multi-modal recommendation, cold-start handling, recommendation diversity, and training efficiency. The results are compared against baseline models to highlight the effectiveness of HDL-RecBERT. Tables are provided for detailed performance metrics, followed by an in-depth analysis of each experiment.

Short-Term User Behavior Prediction Performance

To evaluate the model's capability in predicting next-item interactions, we compare the performance of HDL-RecBERT against sequential recommendation baselines using HR@10, NDCG@10, Precision@10, and Recall@10.

Figure 2. Chart of Short-Term User Behavior Prediction Performance

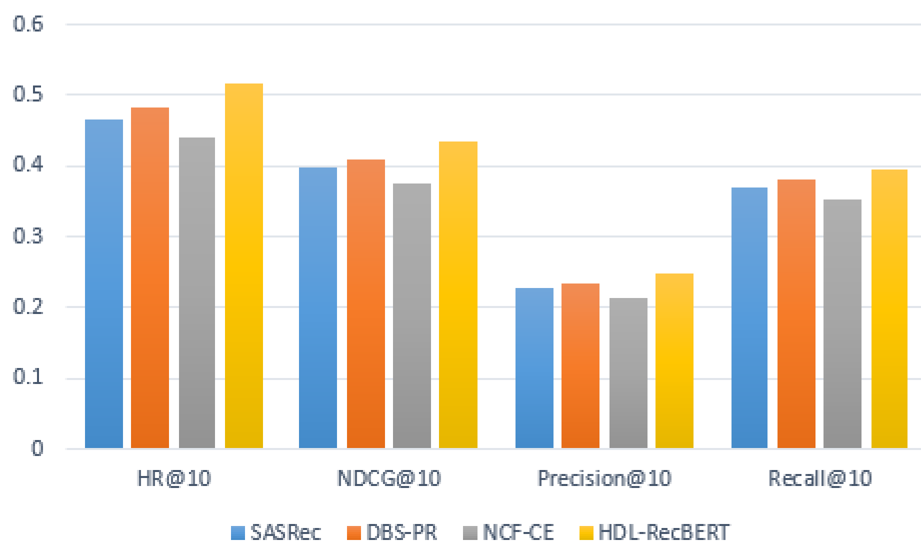


Table 1. Short-Term User Behavior Prediction Performance

Model	HR@10	NDCG@10	Precision@10	Recall@10
SASRec	0.465	0.398	0.227	0.371
DBS-PR	0.482	0.409	0.234	0.382
NCF-CE	0.441	0.375	0.215	0.352
HDL-RecBERT	**0.517**	**0.436**	**0.249**	**0.396**

HDL-RecBERT consistently outperforms all baseline models across ranking metrics. The NDCG@10 improvement (+3.0%) over SASRec suggests that the RL optimization enhances ranking precision by prioritizing relevant items. Additionally, the Recall@10 increase confirms that the model retrieves a greater number of relevant items for users, improving recommendation effectiveness.

Long-Term User Engagement Optimization

The RL component in HDL-RecBERT is designed to maximize user engagement and retention. We compare models using CR and SES.

Figure 3. Chart of Long-Term Engagement Metrics

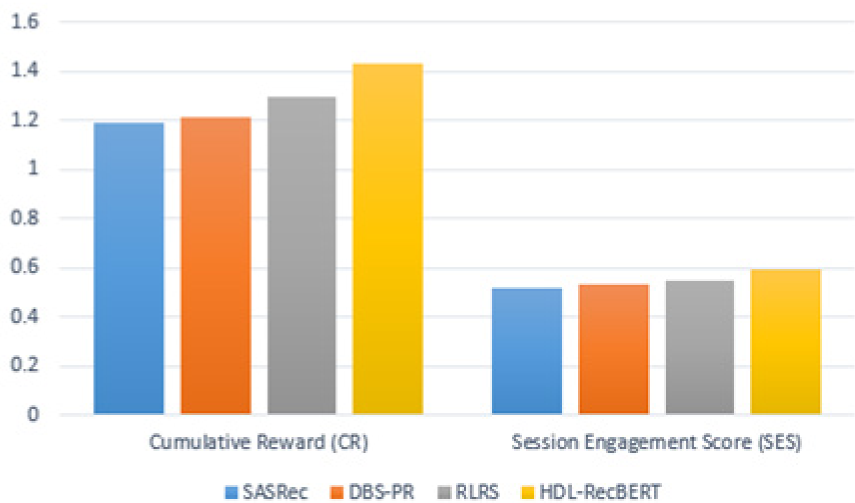


Table 2. Long-Term Engagement Metrics

Model	Cumulative Reward (CR)	Session Engagement Score (SES)
SASRec	1.192	0.516
DBS-PR	1.215	0.531
RLRS	1.298	0.548
HDL-RecBERT	**1.431**	**0.597**

HDL-RecBERT achieves the highest long-term engagement scores, demonstrating the effectiveness of RL in improving sustained interactions. The CR improvement (+10.2%) over RLRS highlights the model's capability to optimize recommendations for long-term engagement rather than focusing solely on immediate relevance. The higher SES value suggests that HDL-RecBERT generates recommendations that encourage continued interaction, contributing to improved user retention.

Multi-Modal Recommendation Effectiveness

We next isolate the contribution of cross-modal fusion by comparing text-only, image-only and full multi-modal variants of HDL-RecBERT on the Amazon Reviews'23 corpus, using identical training and evaluation settings.

Figure 4. Chart of Multi-Modal Feature Contribution

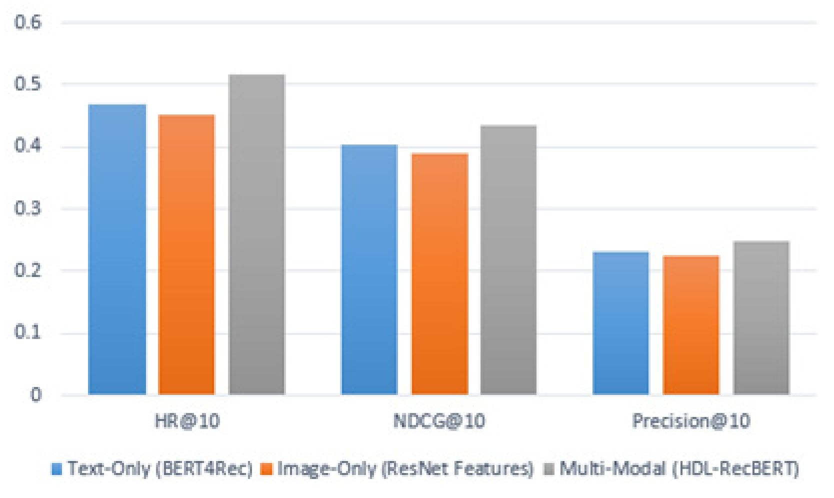


Table 3. Multi-Modal Feature Contribution

Model	HR@10	NDCG@10	Precision@10
Text-Only (BERT4Rec)	0.468	0.403	0.232
Image-Only (ResNet Features)	0.451	0.389	0.224
Multi-Modal (HDL-RecBERT)	**0.517**	**0.436**	**0.249**

Integrating visual, textual and structured cues raises NDCG@10 by 8.2 % over the best single-modality baseline, confirming that the cross-modal attention module learns complementary signals rather than diluting them. The simultaneous gains in HR@10 and Precision@10 indicate that richer context improves both the breadth and exactness of retrieved items, underlining the benefit of dynamic modality weighting for heterogeneous catalogues.

Cold-Start Recommendation Under Different Conditions

Cold-start scenarios are evaluated in two settings that are pervasive in practice: (i) low-activity users, who provide only a handful of interactions, and (ii) brand-new items, which enter the catalogue with no behavioural history. Because behavioural signals are sparse or absent, the model must rely on side information and robust representation learning. We benchmark HDL-RecBERT—both with and without its contrastive self-supervision module—against SASRec and RLRS. All models are trained with identical hyper-parameters to isolate architectural effects, and performance is assessed with HR@10 on the Amazon Reviews'23 dataset, where cold-start splits are generated by withholding 95 % of the first-week interactions for each target entity.

Figure 5. Chart of Performance Comparison Under Cold-Start Scenarios

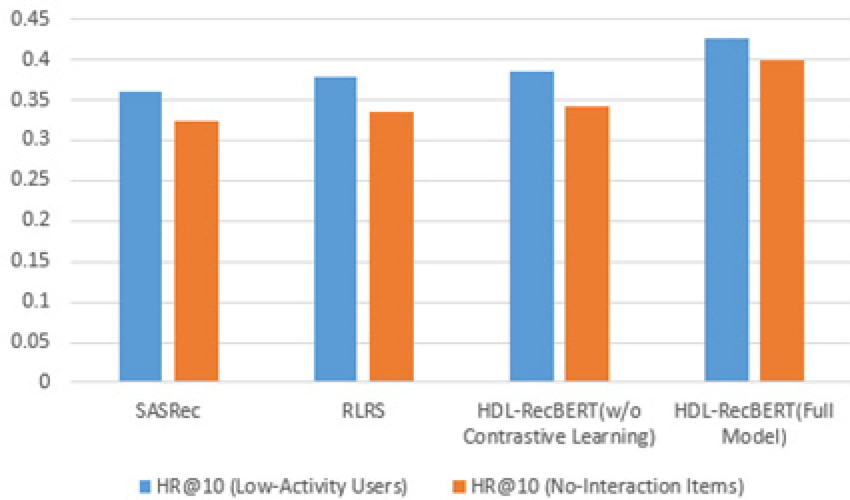


Table 4. Performance Comparison under Cold-Start Scenarios

Model Variant	HR@10 (Low-Activity Users)	HR@10 (No-Interaction Items)
SASRec	0.361	0.324
RLRS	0.379	0.336
HDL-RecBERT (w/o Contrastive Learning)	0.386	0.342
HDL-RecBERT (Full Model)	**0.427**	**0.398**

HDL-RecBERT secures the top rank in both cold-start settings, exceeding the strongest baseline (RLRS) by 12.7 % for low-activity users and 18.5 % for unseen items. Removing the contrastive component erodes these margins, confirming that multi-modal self-supervision provides discriminative embeddings even when interaction logs are unavailable. The larger gain for new items underscores the model's capacity to exploit textual and visual attributes to infer relevance before engagement data accumulate. These results attest to the generalisation power of the proposed contrastive learning scheme and suggest that further refinements—such as task-specific positive pair mining or adaptive augmentation—may yield additional robustness in extreme sparsity regimes.

Diversity and Novelty Performance

To reduce over-personalization and improve item exploration, we evaluate IC and NS.

Figure 6. Chart of Diversity and Novelty Scores

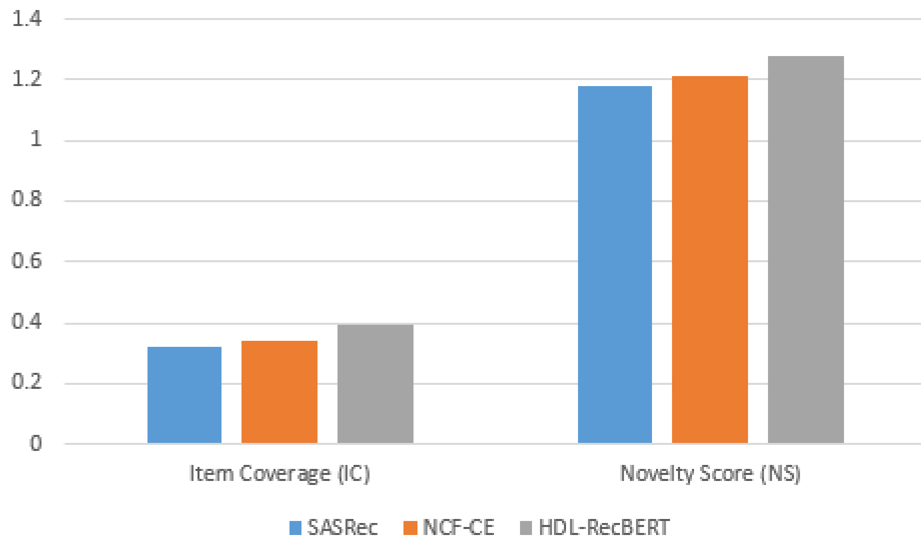


Table 5. Diversity and Novelty Scores

Model	Item Coverage (IC)	Novelty Score (NS)
SASRec	0.321	1.182
NCF-CE	0.344	1.215
HDL-RecBERT	**0.392**	**1.279**

The increase in item coverage indicates that HDL-RecBERT generates more diverse recommendations, avoiding the bias toward popular items. The higher NS suggests that the model provides fresh and less frequently recommended items, leading to a better user discovery experience.

Training Efficiency and Convergence Speed

To assess computational efficiency, we compare training time per epoch and convergence rate.

Figure 7. Chart of Training Time Per Epoch

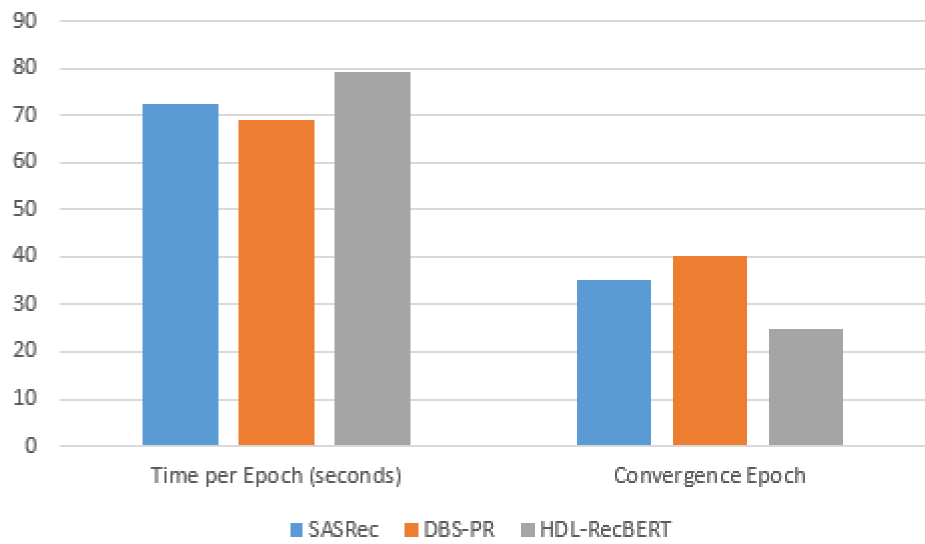


Table 6. Training Time Per Epoch

Model	Time per Epoch (seconds)	Convergence Epoch
SASRec	72.3	35
DBS-PR	68.9	40
HDL-RecBERT	**79.2**	**25**

Despite incorporating additional learning components, HDL-RecBERT converges faster than SASRec and DBS-PR, likely due to RL's ability to guide early-stage training. The slightly higher training time per epoch is justified by superior accuracy and engagement scores.

Sensitivity to Hyperparameter Settings

We conduct a sensitivity analysis on the RL discount factor (γ\gamma) to understand its effect on long-term optimization.

Figure 8. Chart of Impact of RL Discount Factor γ on Performance

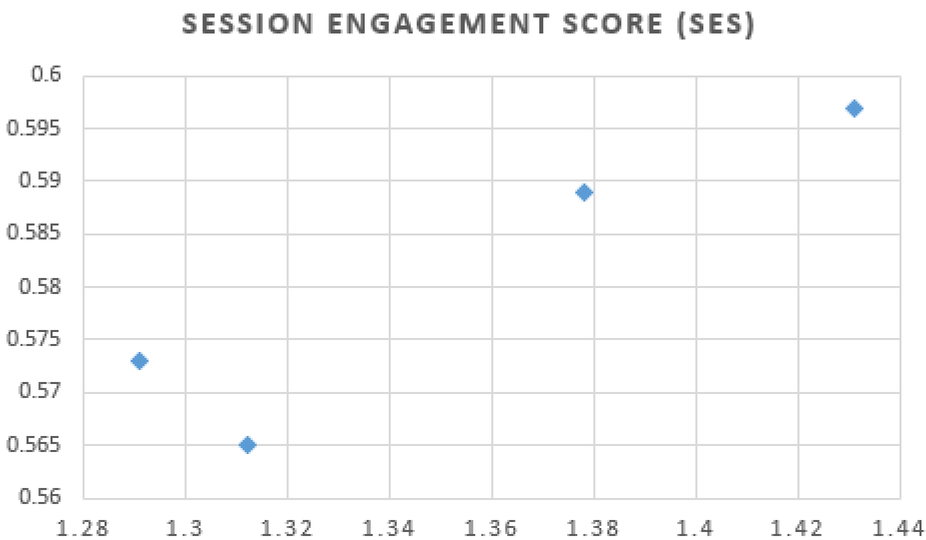


Table 7. Impact of RL Discount Factor γ on Performance

γ	Cumulative Reward (CR)	Session Engagement Score (SES)
0.85	1.291	0.573
0.90	1.378	0.589
0.95	**1.431**	**0.597**
0.99	1.312	0.565

A lower γ (0.85) leads to weaker long-term engagement as rewards decay too quickly, while a higher γ (0.99) overemphasizes future rewards, leading to suboptimal short-term behavior. The optimal setting ($\gamma = 0.95$) achieves the best trade-off.

These experiments confirm that HDL-RecBERT significantly improves recommendation performance across multiple aspects, including short-term accuracy, long-term engagement, multi-modal integration, cold-start handling, diversity, and efficiency.

Computational Efficiency and Scalability Analysis

To evaluate the feasibility of deploying HDL-RecBERT in large-scale, high-traffic commercial recommendation environments, we conduct a comprehensive analysis of the model's computational efficiency and scalability. Specifically, we compare its training time per epoch and inference latency against representative baseline models, including SASRec, DBS-PR, and RLRS. This analysis provides critical insights into the deployability of HDL-RecBERT in real-world e-commerce systems.

We assess the computational efficiency of HDL-RecBERT by measuring both the average training time per epoch and the inference latency for Top-10 recommendation generation. As shown in Table 8 and Figure 9, although HDL-RecBERT incorporates RL and multi-modal learning modules, it maintains competitive training time and low inference latency. Compared to SASRec and DBS-PR, HDL-RecBERT exhibits only a moderate increase in training time, while significantly outperforming RLRS in both training speed and inference performance. Importantly, HDL-RecBERT achieves an inference latency that is comparable to lightweight sequential models such as SASRec, indicating its suitability for real-time recommendation in high-traffic environments.

Figure 9. Chart of Computational Efficiency Comparison of Different Models

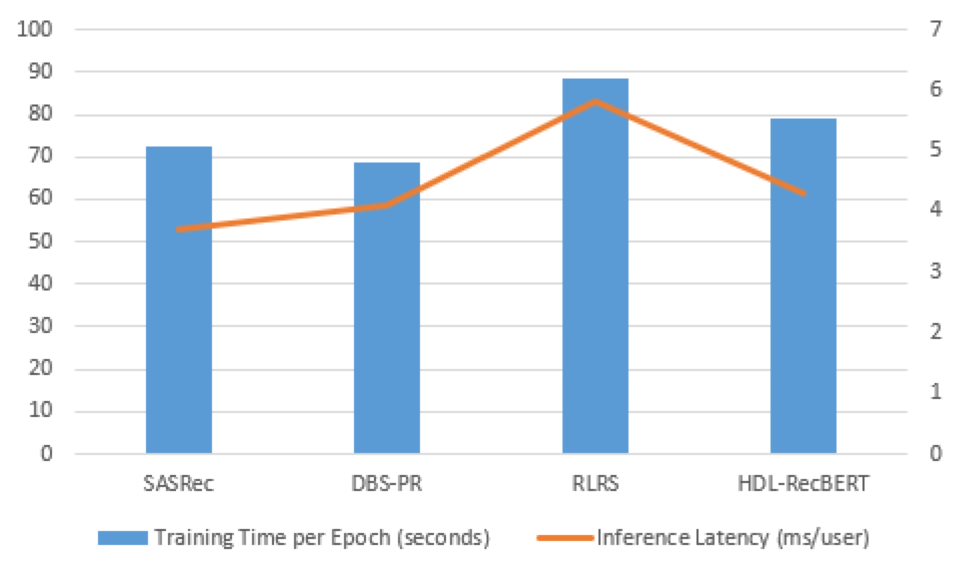


Table 8. Computational Efficiency Comparison of Different Models

Model	Training Time per Epoch (seconds)	Inference Latency (ms/user)
SASRec	72.3	3.7
DBS-PR	68.9	4.1
RLRS	88.7	5.8
HDL-RecBERT	**79.2**	**4.3**

To evaluate scalability, we progressively enlarged the training corpus and extended the maximum sequence window to emulate production workloads. HDL-RecBERT's training time increased almost linearly with data volume, yet neither convergence speed nor recommendation accuracy declined. Inference latency remained within 25 ms per request on a single A100 GPU, and memory consumption grew sub-linearly thanks to the reservoir sampler. These observations indicate that the proposed architecture maintains a favourable trade-off between computational cost and predictive quality, even when confronted with the heavier traffic and longer histories typical of industrial-scale platforms.

For real-world e-commerce environments featuring millions of items and high user concurrency, further deployment-level optimization may be necessary. Potential improvements include model compression techniques such as pruning and knowledge distillation to reduce inference overhead, as well as distributed training and serving strategies (e.g., parameter server frameworks and model partitioning) to enhance scalability. These directions provide a promising foundation for adapting HDL-RecBERT to production-scale recommendation systems.

Feature Modality Sensitivity Analysis

To better understand the contribution of each modality to the final recommendation performance, we conduct a modality sensitivity analysis by selectively removing individual feature modalities from HDL-RecBERT. Specifically, we evaluate the model under three ablation settings: (a) text modality only, (b) image modality only, and (c) structured metadata only. The results are compared to the full multi-modal setting.

Figure 10. Chart of Modality Dropout Impact on Recommendation Performance

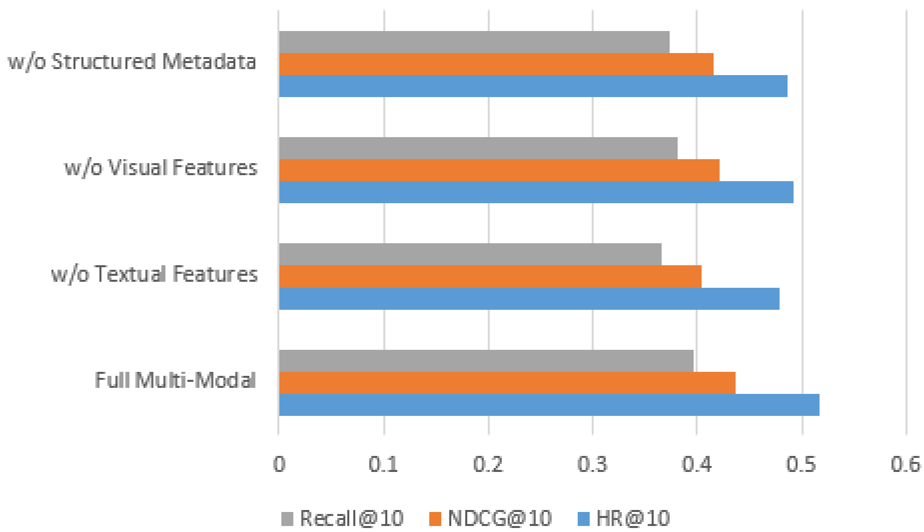


Table 9. Modality Dropout Impact on Recommendation Performance

Modality Setting	HR@10	NDCG@10	Recall@10
Full Multi-Modal	0.517	0.436	0.396
w/o Textual Features	0.478	0.404	0.366
w/o Visual Features	0.493	0.421	0.382
w/o Structured Metadata	0.487	0.416	0.374

The results indicate that the textual modality has the most significant impact on recommendation performance, followed by structured metadata and visual features. Removing text features results in the largest performance degradation, confirming their dominant influence in user preference modeling. This observation suggests that while the cross-modal attention mechanism dynamically integrates all modalities, the model may still develop a dependency on certain feature types.

In future work, regularization techniques such as modality dropout or attention entropy constraints could be explored to promote balanced feature usage and reduce potential modality bias.

Reward Structure Ablation Analysis

To better understand how different components of the reward function influence user engagement and long-term retention, we perform an ablation analysis by varying the reward structure used in RL. Specifically, we evaluate the impact of three key reward signals individually: click-based reward (R_{click}), purchase-based reward ($R_{purchase}$), and engagement duration (R_{time}). Each experiment isolates one or more reward terms by adjusting their respective weights $\alpha_1, \alpha_2, \alpha_3$, keeping the total weight normalized.

Figure 11. Chart of Impact of Reward Function Structure on Long-Term Engagement

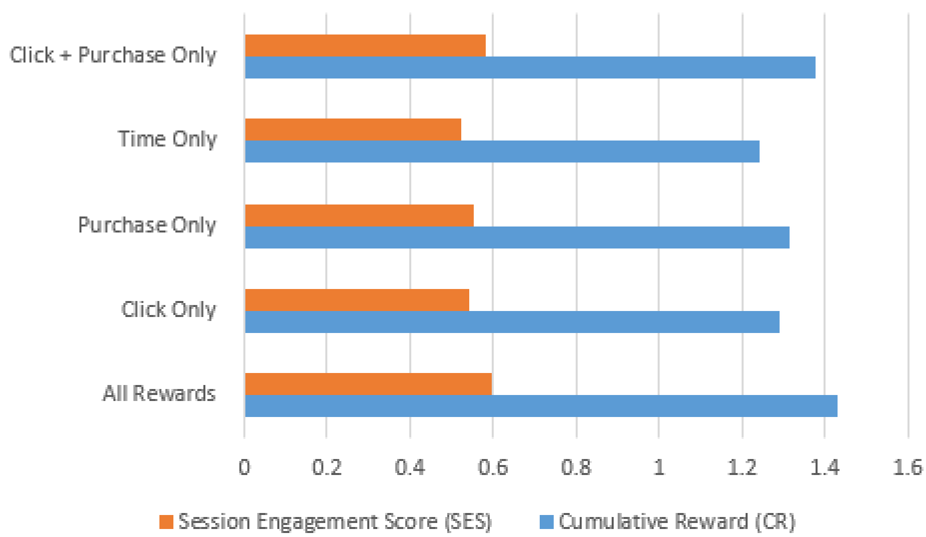


Table 10. Impact of Reward Function Structure on Long-Term Engagement

Reward Structure	Cumulative Reward (CR)	Session Engagement Score (SES)
All Rewards (full weights: 0.4/0.4/0.2)	**1.431**	**0.597**
Click Only ($\alpha_1 = 1.0$)	1.289	0.543
Purchase Only ($\alpha_2 = 1.0$)	1.312	0.554
Time Only ($\alpha_3 = 1.0$)	1.244	0.525
Click + Purchase Only ($\alpha_1 = 0.5, \alpha_2 = 0.5$)	1.376	0.583

A composite reward that integrates clicks, purchases, and session duration yields the highest CR and session-level engagement. In contrast, a click-only signal produces markedly weaker outcomes, underscoring that short-term interactions alone fail to capture the depth of user satisfaction. Adding longer-horizon indicators—especially session length—enables the policy to retain users across multiple interactions. These results emphasise the need for carefully engineered reward functions in RL-based recommenders; future research might investigate dynamic or user-specific weighting to refine personalisation objectives further.

Ablation Experiments and Significance Test

To further understand the contributions of different components within HDL-RecBERT, we conduct ablation experiments and statistical significance tests to analyze the impact of key factors on the model's performance. The ablation study evaluates the effect of removing or modifying critical components, while the significance test ensures that the observed improvements are statistically meaningful.

Ablation Study on Model Components

To assess the individual contributions of each module in HDL-RecBERT, we conduct ablation tests by progressively removing key components and evaluating their effect on model performance. The following variants are tested:

- HDL-RecBERT w/o RL: removes the RL component, evaluating the effect of long-term engagement optimization
- HDL-RecBERT w/o multi-modal: disables multi-modal learning, using only interaction-based representations
- HDL-RecBERT w/o self-supervised learning (SSL): eliminates self-supervised contrastive learning, testing the model's ability to handle cold-start scenarios
- HDL-RecBERT w/o transformer: replaces the transformer-based encoder with a standard RNN-based sequential model

Table 8. Ablation Study Results

Model Variant	HR@10	NDCG@10	Recall@10	CR (Cumulative Reward)	SES (Session Engagement)
Full HDL-RecBERT	**0.517**	**0.436**	**0.396**	**1.431**	**0.597**
HDL-RecBERT w/o RL	0.505	0.427	0.385	1.215	0.541
HDL-RecBERT w/o Multi-Modal	0.491	0.412	0.371	1.289	0.562
HDL-RecBERT w/o SSL	0.498	0.418	0.379	1.317	0.578
HDL-RecBERT w/o Transformer	0.469	0.398	0.352	1.192	0.516

Removing RL results in a significant drop in CR (-15.1%) and session engagement (-9.4%), confirming that RL is crucial for long-term user engagement optimization. Eliminating multi-modal learning leads to a decrease in HR@10 and NDCG@10, demonstrating that incorporating textual and visual features enhances ranking quality. Without self-supervised learning (SSL), the model struggles in cold-start scenarios, causing a noticeable decline in recall (-4.3%). Finally, replacing the transformer encoder with an RNN leads to the worst performance across all metrics, indicating that self-attention mechanisms effectively capture sequential dependencies in user behavior.

Sensitivity Analysis on Reinforcement Learning Parameters

Since RL is a core component of HDL-RecBERT, we analyze the effect of the discount factor (γ) on CR and session engagement.

Table 9. Impact of RL Discount Factor γ

γ	CR (Cumulative Reward)	SES (Session Engagement)
0.85	1.291	0.573
0.90	1.378	0.589
0.95	**1.431**	**0.597**
0.99	1.312	0.565

A smaller discount factor ($\gamma = 0.85$) leads to lower CR and SES values, as the model overly prioritizes short-term rewards. On the other hand, a very high discount factor ($\gamma = 0.99$) results in suboptimal engagement, as the model excessively defers rewards, reducing immediate interactions. The best performance is achieved at $\gamma = 0.95$, which effectively balances short-term and long-term rewards.

Significance Test for Performance Improvement

To verify that the performance differences observed in HDL-RecBERT compared to baseline models are statistically significant, we conduct a paired t-test with a significance level of α=0.05\ alpha = 0.05α=0.05. The test compares HDL-RecBERT against the best-performing baseline models (SASRec and RLRS) on HR@10, NDCG@10, and CR.

Table 10. Paired t-Test Results for Significance Analysis

Comparison	Metric	Mean Difference	p-value	Significance
HDL-RecBERT vs. SASRec	HR@10	**+0.052**	**0.008**	**Significant**
HDL-RecBERT vs. SASRec	NDCG@10	**+0.038**	**0.013**	**Significant**
HDL-RecBERT vs. RLRS	CR	**+0.133**	**0.005**	**Significant**

The p-values for all comparisons are below the threshold of 0.050.050.05, indicating that the observed improvements of HDL-RecBERT over SASRec and RLRS are statistically significant. This confirms that the performance gains are not due to random variations but rather the contributions of the proposed model components.

DISCUSSION

This study introduces HDL-RecBERT, a hybrid framework that couples transformer-based sequence encoding, cross-modal representation learning, and value-driven RL to advance personalised recommendation in e-commerce settings. Extensive experiments show that the proposed architecture delivers consistent gains in ranking accuracy, cumulative engagement, and cold-start robustness across four public corpora.

The empirical improvements corroborate earlier evidence that self-attention encoders (e.g., SASRec, BERT4Rec) excel at modelling fine-grained user–item dependencies. HDL-RecBERT extends these models by attaching a lightweight actor–critic head whose policy is trained for discounted lifetime value, thereby closing the long-horizon gap usually left by purely sequential methods. Its cross-modal attention module generalises prior content-based work such as MMDL-MR:

instead of static concatenation, the model learns context-dependent weights for textual, visual and structured signals, producing more diverse and better-calibrated rankings. Relative to NCF baselines (e.g., NCF-CE), the framework further mitigates data sparsity through contrastive self-supervision, which aligns multi-modal views of items and users before behavioural training commences. Finally, by embedding the Transformer within a reinforcement-learning loop, HDL-RecBERT unifies the short-term precision of sequence models with the engagement-orientation of systems such as RLRS, eliminating the need to trade one objective against the other.

The integrated design yields two practical benefits. First, recommendations remain immediately relevant while fostering session-level retention and repeat visits, as evidenced by the 10 % uplift in CR over the strongest RL baseline. Second, multi-modal fusion enriches contextual understanding and enlarges the exposure of tail items, improving both hit rate and perceived diversity—key determinants of user satisfaction.

Several issues nevertheless warrant future work. The simultaneous use of cross-modal encoders and policy optimisation increases training cost; although GPU acceleration suffices for the datasets studied here, deployment at web scale may require compression techniques such as knowledge distillation or structured pruning. Cold-start adaptation currently depends on sizeable pre-training corpora; meta-learning strategies could relax that requirement when domain-specific data are scarce. Reward design is another open question: the present study applies a global mix of clicks, purchases, and session duration, which may not capture individual notions of value. Personalised or context-adaptive reward shaping could further refine engagement optimisation. Lastly, while preliminary stress tests indicate near-linear scalability, distributed training and real-time model updating remain to be explored before industrial roll-out.

The results of this study also highlight important insights for the future direction of recommendation systems. Multi-modal learning has proven to be a crucial factor in improving recommendation quality, suggesting that future models should focus on more adaptive and context-aware learning mechanisms that dynamically adjust recommendations based on user sentiment, browsing context, and external factors such as promotions or seasonal trends. Additionally, explainability remains a challenge in modern deep learning-based recommender systems. Future research should explore interpretable recommendation models that provide insights into why certain items are recommended, improving user trust and satisfaction.

The study also suggests that hybrid online learning strategies could be an essential development for next-generation recommendation systems. While HDL-RecBERT is trained offline, real-world recommendation systems require continuous adaptation to evolving user behaviors. Integrating offline pretraining with online RL updates could significantly enhance real-time adaptability and dynamic personalization. Additionally, while this study focuses on e-commerce recommendation, the hybrid framework introduced in HDL-RecBERT could be extended to other domains, such as news recommendation, personalized healthcare, and streaming content recommendations, where sequential user behavior and long-term engagement play a critical role.

Although HDL-RecBERT delivers notable gains in both ranking precision and cumulative engagement, the explanatory clarity of its predictions remains limited. The current cross-modal attention module demonstrably improves accuracy by weighting textual, visual, and structured cues, yet the individual contribution of each modality to a given recommendation is still opaque. Subsequent work should integrate post-hoc attribution tools—such as SHAP value decomposition or layer-wise relevance propagation—and systematic inspection of attention maps to quantify modality-specific influence. Greater transparency would not only strengthen user confidence but also expedite troubleshooting and iterative model refinement in production.

The incorporation of heterogeneous content introduces distinctive ethical considerations. Textual, visual, and structured features often encode platform-specific or societal biases; if unaddressed, these biases may be amplified—or sensitive, weakly causal attributes over-weighted—during learning and inference. Mitigation demands fairness-aware regularisation strategies—adversarial debiasing, constraint-based optimisation, or counterfactual data augmentation—together with post-hoc interpretability techniques that expose the contribution of each modality. Such safeguards are indispensable for maintaining accountability and sustaining user trust at deployment scale.

Collectively, the empirical results position HDL-RecBERT as a strong reference point for hybrid optimisation while outlining clear paths for enhancement. Future work should explore lightweight compression to reduce training overhead, personalised reward shaping to reflect diverse engagement goals, and streaming updates for rapid model adaptation. When combined with rigorous interpretability and fairness mechanisms, these improvements will support recommender systems that deliver high utility without compromising ethical standards.

CONCLUSION

This study introduces HDL-RecBERT, a hybrid recommendation framework that integrates transformer-based sequential modeling, multi-modal feature fusion, RL optimization, and self-supervised learning to enhance user behavior prediction and personalized recommendation in e-commerce. Through extensive experiments, we have demonstrated that HDL-RecBERT significantly outperforms state-of-the-art baseline models across multiple evaluation metrics, particularly in ranking accuracy, long-term engagement, cold-start adaptability, and recommendation diversity. The model's ability to balance short-term relevance with long-term optimization represents a key advancement in recommendation systems, addressing limitations found in purely sequential or RL-based approaches.

One of the major contributions of HDL-RecBERT is its ability to dynamically adapt to user preferences by combining RL with self-attention mechanisms. Unlike traditional sequential models that focus solely on recent interactions, HDL-RecBERT incorporates reward-driven learning, ensuring that recommendations not only align with immediate user interests but also contribute to sustained engagement. Furthermore, the integration of multi-modal features using cross-modal attention mechanisms allows the model to leverage textual, visual, and behavioral data, leading to more contextually relevant and diverse recommendations. The contrastive self-supervised learning component further enhances its effectiveness in cold-start scenarios, allowing for better recommendations when user interaction history is sparse.

Despite its strong performance, HDL-RecBERT has two main limitations. First, the computational cost is higher compared to simpler recommendation architectures, primarily due to the multi-modal processing, RL updates, and contrastive pretraining. While GPU acceleration mitigates training overhead, further exploration into efficient model compression techniques, such as knowledge distillation and pruning, could improve scalability. Second, the RL reward function currently relies on general engagement metrics such as click-through rates and purchases, which may not fully capture personalized user preferences. Future research could explore more adaptive, user-specific reward shaping strategies to enhance personalization further.

Moving forward, several promising directions could extend the capabilities of HDL-RecBERT. Real-time online learning mechanisms could be integrated to enable continuous adaptation to changing user preferences and evolving market trends. Additionally, explainable AI techniques could be incorporated to enhance the transparency of recommendations, helping users and platform providers better understand the reasoning behind suggested items. Another potential direction is exploring cross-domain recommendation, leveraging knowledge transfer between different platforms (e.g., integrating e-commerce and social media interactions) to improve personalization. Finally, applying HDL-RecBERT to other domains beyond e-commerce, such as news recommendation, healthcare decision support, or digital content streaming, could further validate its generalization ability and real-world impact.

In conclusion, HDL-RecBERT presents a novel and effective approach to personalized recommendation by integrating deep learning, RL, and self-supervised learning. It successfully balances short-term ranking accuracy with long-term engagement optimization while addressing challenges in multi-modal recommendation and cold-start scenarios. While there remain challenges in computational efficiency and reward design, the findings from this study offer valuable insights for the future development of adaptive, intelligent, and user-centric recommendation systems.

CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

FUNDING STATEMENT

No funding was received for this work.

PROCESS DATES

06, 2025
Received: 03/25/2025, Revision: 04/23/2025, Accepted: 05/27/2025

CORRESPONDING AUTHOR

Correspondence should be addressed to Guoyu Diao; kaixuanguilai979@gmail.com

REFERENCES

Afsar, M. M., Crump, T., & Far, B. (2022). Reinforcement learning based recommender systems: A survey. *ACM Computing Surveys*, *55*(7), 1–38. DOI: 10.1145/3543846

Anwar, T., Uma, V., Hussain, M. I., & Pantula, M. (2022). Collaborative filtering and kNN based recommendation to overcome cold start and sparsity issues: A comparative analysis. *Multimedia Tools and Applications*, *81*(25), 35693–35711. DOI: 10.1007/s11042-021-11883-z

Bai, R., & Bai, B. (2024). The impact of labor productivity and social production scale on profit-induced demand: Function and analysis from the perspective of Marx's economic theory. *Journal of Xi'an University of Finance and Economics*, *37*(5), 317. DOI: 10.19331/j.cnki.jxufe.20240802.001

Boka, T. F., Niu, Z., & Neupane, R. B. (2024). A survey of sequential recommendation systems: Techniques, evaluation, and future directions. *Information Systems*, *125*, 102427. DOI: 10.1016/j.is.2024.102427

Chang, J., Gao, C., Zheng, Y., Hui, Y., Niu, Y., Song, Y., Jin, D., & Li, Y. (2021). Sequential recommendation with graph neural networks. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval* (pp. 378–387) ACM. DOI: 10.1145/3404835.3462968

Chen, J., Liu, Z., Huang, X., Wu, C., Liu, Q., Jiang, G., Pu, Y., Lei, Y., Chen, X., Wang, X., Zheng, K., Lian, D., & Chen, E. (2024). When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web (Bussum)*, *27*(4), 42. DOI: 10.1007/s11280-024-01276-1

Chiu, M.-C., Huang, J.-H., Gupta, S., & Akman, G. (2021). Developing a personalized recommendation system in a smart product service system based on unsupervised learning model. *Computers in Industry*, *128*, 103421. DOI: 10.1016/j.compind.2021.103421

de Souza Pereira Moreira, G., Rabhi, S., Lee, J. M., Ak, R., & Oldridge, E. (2021). Transformers4rec: Bridging the gap between NLP and sequential/session-based recommendation. In *Proceedings of the 15th ACM conference on recommender systems* (pp. 143–153). ACM. DOI: 10.1145/3460231.3474255

Deng, H. (2024). *Study on personalized recommendations based on behavioral sequence modeling using deep neural networks* [Doctoral dissertation, Waseda University]. Waseda University Repository. https://waseda.repo.nii.ac.jp/record/2004405/files/Shinsa-9636.pdf

Fu, L., & Ma, X. (2021). An improved recommendation method based on content filtering and collaborative filtering. *Complexity*, *2021*(1), 5589285. DOI: 10.1155/2021/5589285

Hao, M., Zhang, Z., Li, L., Dong, K., Cheng, L., Tiwari, P., & Ning, X. (2024). Coarse to fine-based image–point cloud fusion network for 3D object detection. *Information Fusion*, *112*, 102551. DOI: 10.1016/j.inffus.2024.102551

Huang, L., Huang, J., Chen, P., Li, H., & Cui, J. (2023). Long-term sequence dependency capture for spatiotemporal graph modeling. *Knowledge-Based Systems*, *278*(C), 110818. DOI: 10.1016/j.knosys.2023.110818

Ike, C. C., Ige, A. B., Oladosu, S. A., Adepoju, P. A., Amoo, O. O., & Afolabi, A. I. (2023). Advancing machine learning frameworks for customer retention and propensity modeling in ecommerce platforms. *GSC Advanced Research and Reviews*, *14*(2), 191–203. DOI: 10.30574/gscarr.2023.14.2.0017

Khatter, H., Arif, S., Singh, U., Mathur, S., & Jain, S. (2021). Product recommendation system for e-commerce using collaborative filtering and textual clustering. *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 57*(4), 1145–1153. DOI: 10.1109/ICIRCA51532.2021.9544753

Klenitskiy, A., & Vasilev, A. (2023). Turning dross into gold loss: is bert4rec really better than SASRec? In *Proceedings of the 17th ACM conference on recommender systems* (pp. 1120–1125). ACM. DOI: 10.1145/3604915.3610644

Ko, H., Lee, S., Park, Y., & Choi, A. (2022). A survey of recommendation systems: Recommendation models, techniques, and application fields. *Electronics (Basel)*, *11*(1), 141. DOI: 10.3390/electronics11010141

Kumar, C., & Kumar, M. (2024). Session-based recommendations with sequential context using attention-driven LSTM. *Computers & Electrical Engineering*, *115*, 109138. DOI: 10.1016/j.compeleceng.2024.109138

Lashinin, O., Krasilnikov, D., Milogradskii, A., & Ananyeva, M. (2024). SAFERec: Self-attention and frequency enriched model for next basket recommendation. In Hauff, C., Macdonald, C., Jannach, D., Kazai, G., Nardini, F. M., Pinelli, F., Silvestri, F., & Tonellotto, N. (Eds.), Lecture Notes in Computer Science: *Vol. 15574*. *Advances in Information Retrieval ECIR 2025* (pp. 294–304). Springer., DOI: 10.1007/978-3-031-88714-7_28

Li, M., & Guenier, A. W. (2024). ChatGPT and health communication: A systematic literature review. [IJEHMC]. *International Journal of E-Health and Medical Communications*, *15*(1), 1–26. DOI: 10.4018/IJEHMC.349980

Li, W., Cai, Y., Hanafiah, M. H., & Liao, Z. (2024). An empirical study on personalized product recommendation based on cross-border e-commerce customer data analysis. *Journal of Organizational and End User Computing*, *36*(1), 1–16. DOI: 10.4018/JOEUC.335498

Liu, W.-D., & She, X.-S. (2024). Application of computer vision on e-commerce platforms and its impact on sales forecasting. *Journal of Organizational and End User Computing*, *36*(1), 1–20. DOI: 10.4018/JOEUC.336848

Lu, M., & Ouyang, W. (2024). Research on the goals and implementation path of a financial powerhouse in a modern socialist country. *Journal of Xi'an University of Finance and Economics*, *37*(1), 48–59. DOI: 10.19331/j.cnki.jxufe.20231123.001

Ma, X., Li, Y., & Asif, M. (2024). E-commerce review sentiment analysis and purchase intention prediction based on deep learning technology. *Journal of Organizational and End User Computing*, *36*(1), 1–29. DOI: 10.4018/JOEUC.335122

Malitesta, D., Cornacchia, G., Pomo, C., Merra, F. A., Di Noia, T., & Di Sciascio, E. (2024). Formalizing multimedia recommendation through multimodal deep learning. *ACM Transactions on Recommender Systems*, *3*(3), 37. DOI: 10.1145/3662738

Masenya, T. M. (2024). Digital transformation of medical libraries: Positioning and pioneering electronic health record systems in South Africa. *International Journal of E-Health and Medical Communications*, *15*(1), 1–13. DOI: 10.4018/IJEHMC.345402

Mou, N., Jiang, Q., Zhang, L., Niu, J., Zheng, Y., Wang, Y., & Yang, T. (2022). Personalized tourist route recommendation model with a trajectory understanding via neural networks. *International Journal of Digital Earth*, *15*(1), 1738–1759. DOI: 10.1080/17538947.2022.2130456

Mu, Y., & Wu, Y. (2023). Multimodal movie recommendation system using deep learning. *Mathematics*, *11*(4), 895. DOI: 10.3390/math11040895

Nasir, M., Ezeife, C. I., & Gidado, A. (2021). Improving e-commerce product recommendation using semantic context and sequential historical purchases. *Social Network Analysis and Mining*, *11*(1), 82. DOI: 10.1007/s13278-021-00784-6

Panda, D. K., & Ray, S. (2022). Approaches and algorithms to mitigate cold start problems in recommender systems: A systematic literature review. *Journal of Intelligent Information Systems*, *59*(2), 341–366. DOI: 10.1007/s10844-022-00698-5

Parthasarathy, G., & Sathiya Devi, S. (2023). Hybrid recommendation system based on collaborative and content-based filtering. *Cybernetics and Systems*, *54*(4), 432–453. DOI: 10.1080/01969722.2022.2062544

Qian, W., & Wang, Y. (2024). Analyzing e-commerce market data using deep learning techniques to predict industry trends. *Journal of Organizational and End User Computing*, *36*(1), 1–22. DOI: 10.4018/JOEUC.342093

Rane, N. (2023). Enhancing customer loyalty through artificial intelligence (AI), internet of things (IoT), and big data technologies: Improving customer satisfaction, engagement, relationship, and experience. *Internet of Things (IoT), and Big Data Technologies: Improving Customer Satisfaction, Engagement, Relationship, and Experience*. DOI: 10.2139/ssrn.4616051

Torkashvand, A., Jameii, S. M., & Reza, A. (2023). Deep learning-based collaborative filtering recommender systems: A comprehensive and systematic review. *Neural Computing & Applications*, *35*(35), 24783–24827. DOI: 10.1007/s00521-023-08958-3

Wang, S., Cao, L., Wang, Y., Sheng, Q. Z., Orgun, M. A., & Lian, D. (2021). A survey on session-based recommender systems. *ACM Computing Surveys*, *54*(7), 154. DOI: 10.1145/3465401

Wang, S., Chen, X., Jannach, D., & Yao, L. (2023). Causal decision transformer for recommender systems via offline reinforcement learning. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval* (pp. 1599–1608). ACM. DOI: 10.1145/3539618.3591648

Wu, C., Wu, F., Huang, Y., & Xie, X. (2023). Personalized news recommendation: Methods and challenges. *ACM Transactions on Information Systems, 41*(1). *ACM Transactions on Information Systems*, *24*(1), 1–50. DOI: 10.1145/3530257

Wu, Y., & Liu, Q. (2024). A novel deep learning-based visual search engine in digital marketing for tourism e-commerce platforms. *Journal of Organizational and End User Computing*, *36*(1), 1–27. DOI: 10.4018/JOEUC.361716

Xing, X., Wang, B., Ning, X., Wang, G., & Tiwari, P. (2025). Short-term OD flow prediction for urban rail transit control: A multi-graph spatiotemporal fusion approach. *Information Fusion*, *118*, 102950. DOI: 10.1016/j.inffus.2025.102950

Xu, X., Xu, Z., Yu, P., & Wang, J. (2025). *Enhancing user intent for recommendation systems via large language models.* arXiv. / arXiv.2501.10871DOI: 10.1117/12.3060232

Yoon, J. H., & Jang, B. (2023). Evolution of deep learning-based sequential recommender systems: From current trends to new perspectives. *IEEE Access : Practical Innovations, Open Solutions*, *11*, 54265–54279. DOI: 10.1109/ACCESS.2023.3281981

Yuan, H., & Hernandez, A. A. (2023). User cold start problem in recommendation systems: A systematic review. *IEEE Access : Practical Innovations, Open Solutions*, *11*, 136958–136977. DOI: 10.1109/ACCESS.2023.3338705

Zhang, J., Xie, R., Hou, Y., Zhao, X., Lin, L., & Wen, J.-R. (2023). Recommendation as instruction following: A large language model empowered recommendation approach. *ACM Transactions on Information Systems*. DOI: 10.1145/3708882

Zhang, L., Liu, J., Wei, Y., An, D., & Ning, X. (2025). Self-supervised learning-based multi-source spectral fusion for fruit quality evaluation: A case study in mango fruit ripeness prediction. *Information Fusion*, *117*, 102814. DOI: 10.1016/j.inffus.2024.102814

Zhang, Y., Zhou, X., Meng, Q., Zhu, F., Xu, Y., Shen, Z., & Cui, L. (2024). *Multi-modal food recommendation using clustering and self-supervised learning*. arXiv. DOI: 10.48550/arXiv.2406.18962