From Past To Path: Masked History Learning for Next-Item Prediction in Generative Recommendation

Anonymous ACL submission

Abstract

Generative recommendation, which directly generates item identifiers, has emerged as a promising paradigm for recommendation systems. However, this left-to-right paradigm inherently biases the model towards local contexts, failing to capture the bidirectional dependencies necessary for understanding complex user intents. To address this limitation, we propose **Masked History Learning (MHL)**, a novel training framework that shifts the objective from simple next-step prediction to deep comprehension of history. MHL augments the standard autoregressive objective with an auxiliary task of reconstructing masked historical items, compelling the model to understand "why" an item path is formed from the user's past behaviors, rather than just "what" item comes next. We introduce two key contributions to enhance this framework: (1) an **entropy-guided masking** policy that intelligently targets the most informative historical items for reconstruction, and (2) a **curriculum learning** scheduler that progressively transitions from history reconstruction to future prediction. Experiments on three public datasets show that our method significantly outperforms state-of-the-art generative models, highlighting that a comprehensive understanding of the past is crucial for accurately predicting a user's future path. The code will be released. Anonymous code is available at https://anonymous.4open.science/r/MHL_anonymous-57C7.

1 Introduction

Recommender systems have become essential tools for navigating the vast digital landscape, evolving from collaborative filtering (Wang et al., 2015; Li et al., 2024; Chen et al., 2018) to sequential models that capture user behavior dynamics (Purificato et al., 2024; Yuan et al., 2023; He et al., 2023). A new paradigm, *generative recommendation* (Rajput et al., 2023; Muennighoff et al., 2025), has recently emerged, offering powerful new ways to model user preferences. This approach adapts pre-trained language models like T5 (Rajput et al., 2023; Bao et al., 2023) and utilizes large language models (Hou et al., 2025b) to directly generate a sequence of semantic IDs representing the items to be recommended (Hua et al., 2023; Zhai et al., 2024), thus providing unprecedented flexibility.

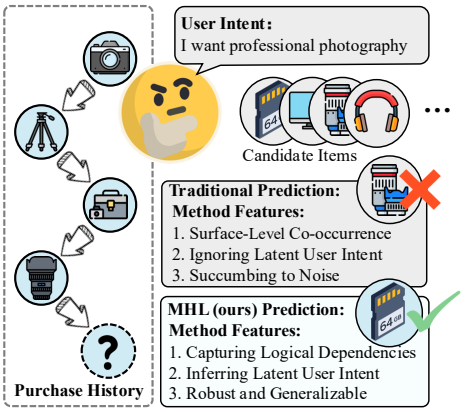

Figure 1: Comparison between the traditional generative recommendation system and our MHL framework.

However, despite their architectural diversity, these models share a fundamental limitation: they are trained almost exclusively to predict the next single item, rather than to understand the path that led there. This narrow focus on autoregressive next-item prediction, while intuitive, prioritizes local transitions over global understanding of user behavior. We contend that this paradigm yields models proficient in local forecasting yet deficient in global comprehension, rendering them susceptible to noise and short-term deviations (i.e., short-term myopia). Our preliminary study (see Figure 6 in Appendix E) reveals a startling phenomenon, i.e., existing SOTA models suffer a **performance collapse** when recent interaction history is truncated, while MHL maintains robustness and **outperforms the leading baseline by over 40%**. This exposes their over-reliance on recency bias rather than deep

intent understanding. In contrast, MHL maintains robustness by leveraging global context.

For example, as shown in Figure 1, consider the purchasing path of a photography enthusiast, who interacts with the following items in order: *camera body*, *tripod*, *camera bag*, and *camera lens*. Although the ground truth for the subsequent purchase is the *memory card*, existing models, fixated on recent item (*camera lens*), often incorrectly predict other lens-related accessories. The user's intention is a direct continuation of purchasing the initial *camera body*, but this intention is obscured by intermediate items. Due to the inability to fully internalize the underlying intention associations behind items along the purchasing path, existing autoregressive models are trained merely to predict "what comes next," but cannot effectively understand "why this path matters."

To address this limitation, we introduce **Masked History Learning** (MHL), a novel training framework for generative recommendation. Specifically, we augment next-item prediction with an auxiliary objective of reconstructing masked items within historical paths. This approach shifts the learning paradigm from predicting results to understanding the process, yielding three key advantages: **(a) Capturing Logical Dependencies.** MHL compels the model to understand the intrinsic associations between masked items and other items, thereby shifting the focus from statistical co-occurrence to the logical structure of a user's path. For example, by reconstructing the masked historical item "tripod", the model is forced to learn that "tripod" is a logical complement to a "camera body" purchase. **(b) Inferring Latent User Intent.** The deep understanding of paths enables models to look beyond a user's explicit behaviors and infer the latent intent driving them. Models can learn to comprehend a coherent and higher-level goal (like "pursuing professional photography") from seemingly disparate items (such as cameras, bags, and future accessories). **(c) Learning Robust and Generalizable Representations.** The history reconstruction objective inherently enhances the quality of the learned representations. To reconstruct history accurately, the model must prioritize strong, logically consistent signals while learning to discount irrelevant or noisy interactions that provide poor contextual clues. This focus on the core signal results in item representations that are more stable and less susceptible to incidental deviations in behavior.

We validate the proposed MHL at multiple granularities: item-, token-, and mixed-level, consistently observing performance gains. To refine this learning process, we introduce two key innovations. First, moving beyond random masking, we propose an adaptive strategy guided by information theory (MacKay, 2002). We selectively mask items sharing high **entropy** with others, creating challenging training signals that focus on significant behavioral patterns. Second, we employ **curriculum learning** (Bengio et al., 2009) to connect the history reconstruction training with autoregressive inference. The training process begins with a warmup phase (He et al., 2016) using random masking, followed by a transition to a high masking ratio guided by entropy to build deep contextual understanding. The ratio is gradually reduced to prepare the model for path generation.

We conduct extensive experiments on three categories of the Amazon Reviews 2014 dataset (McAuley et al., 2015). The results demonstrate that understanding the past significantly enhances the model's ability to predict future paths, outperforming state-of-the-art baselines on metrics like Recall@K and NDCG@K. The contributions of this paper can be summarized as follows:

- We identify a key limitation in generative recommenders: training dominated by next-step prediction overlooks deep understanding of user history. We address this by proposing *Masked History Learning*, which learns to reconstruct a user's *past* to better predict their future *path*.
- We design two strategies to enhance our framework: *entropy-guided masking* to focus on the most informative historical parts, and *curriculum learning* to bridge the gap between understanding history and generating future paths.
- Extensive experiments on three categories of the Amazon Reviews 2014 dataset validate our approach's effectiveness, achieving new state-of-the-art results for generative recommendation.

2 Related Work

Sequential Recommendation. Sequential recommendation models user behavior over time to predict future interactions. Early methods use Markov chains to capture item-to-item transitions (Rendle et al., 2010). Deep learning has since transformed this field. Modern approaches employ various neural architectures including recurrent neural networks (Hidasi et al., 2016; Li et al., 2017; Yue et al.,

2024), convolutional neural networks (Tang and Wang, 2018), Transformers (Kang and McAuley, 2018; Sun et al., 2019), and graph neural networks (Chang et al., 2021; Wu et al., 2019).

While most sequential models are trained autoregressively, some studies have explored alternative learning objectives that go beyond simple next-item prediction. BERT4Rec (Sun et al., 2019) and S^3-Rec (Zhou et al., 2020) use masked item prediction with bidirectional encoders to learn rich contextual representations for discriminative recommendation. These models randomly mask items in user sequences and learn to reconstruct them using full bidirectional context. This approach helps models capture richer dependencies compared to purely left-to-right training. Despite this success, the generative recommendation paradigm has remained largely autoregressive.

Generative Recommendation. Recent advances in generative models have has promoted the transformation of recommendation systems from discriminative to generative (Rajput et al., 2023). Inspired by generative retrieval (Tay et al., 2022; Wang et al., 2022), these methods tokenize items into discrete semantic identifiers. Sequence-to-sequence models can then directly generate these identifiers as recommendations.

Two main approaches have emerged in generative recommendation. The first leverages Large Language Models (LLMs) through zero-shot prompting (Gao et al., 2023; Harte et al., 2023) and instruction tuning (Muennighoff et al., 2025) to align LLMs with user behaviors. The second focuses on semantic ID-based generation, where items are first encoded as discrete token sequences (Rajput et al., 2023) derived from quantizing dense representations (Hua et al., 2023; Wang et al., 2024), then autoregressively decoded to produce recommendations (Zhai et al., 2024).

Despite their flexibility and scalability, existing generative recommenders (Rajput et al., 2023; Wang et al., 2024; Hou et al., 2025b) share a common limitation: they rely almost exclusively on autoregressive training that predicts the next item token given previous tokens. This left-to-right approach focuses on local transitions but may miss internal dependencies and underlying user intent.

Our Contribution. Our study addresses this gap by introducing history reconstruction learning to generative recommendation. Unlike BERT4Rec and S^3-Rec, which employ masked prediction within bidirectional encoders to learn representations for discriminative scoring tasks, our proposed MHL augments standard *unidirectional, decoder-only* model with an auxiliary historical reconstruction objective. This design preserves the model's native autoregressive generation capability while enriching the training signal through deeper historical understanding. We further introduce entropy-guided masking to focus learning on the most informative historical patterns and curriculum learning to seamlessly transition from history understanding to future path generation. Together, these contributions establish a new training paradigm for generative recommenders that emphasizes understanding the past to better predict future paths.

3 Method

We propose **Masked History Learning (MHL)**, a training framework that augments generative recommendation with an auxiliary objective of reconstructing masked historical items. As illustrated in Figure 2, MHL jointly optimizes next-item prediction and masked history reconstruction, guided by entropy-based masking and curriculum scheduling.

3.1 Preliminaries

Semantic ID Representation. Recent generative recommendation systems (Rajput et al., 2023; Hou et al., 2025a) represent each item as a sequence of discrete semantic tokens. In this work, each item i from the item set $\mathcal{I}$ is encoded as a K-digit semantic ID $\phi(i) = \{w_i^1, \ldots, w_i^K\}$, where each digit w_i^k is drawn from a codebook $\mathcal{W}^k$. We construct a unified vocabulary $\mathcal{V} = \mathcal{V}_{\text{code}} \cup \mathcal{V}_{\text{mask}}$, where codebook tokens occupy non-overlapping ID ranges and position-aware mask tokens enable unambiguous reconstruction (details in Appendix A).

Base Architecture. Given a user history $S_T = (i_1, \ldots, i_T)$, item i_t obtains its item-level representations H_t by performing mean pooling on K-digit embeddings, and generates context hidden state $\mathbf{h}_t \in \mathbb{R}^d$ using a Transformer decoder. For prediction, K lightweight heads $\{g_k\}_{k=1}^K$ derive digit-wise states $\mathbf{h}_t^k = g_k(\mathbf{h}_t)$, which are classified using temperature-scaled cosine similarity over codebook embeddings. We denote the resulting probability distribution over codebook $\mathcal{W}^k$ as $P_\theta^k(\cdot \mid \mathbf{h}_t)$. Full architectural details are provided in Appendix B.

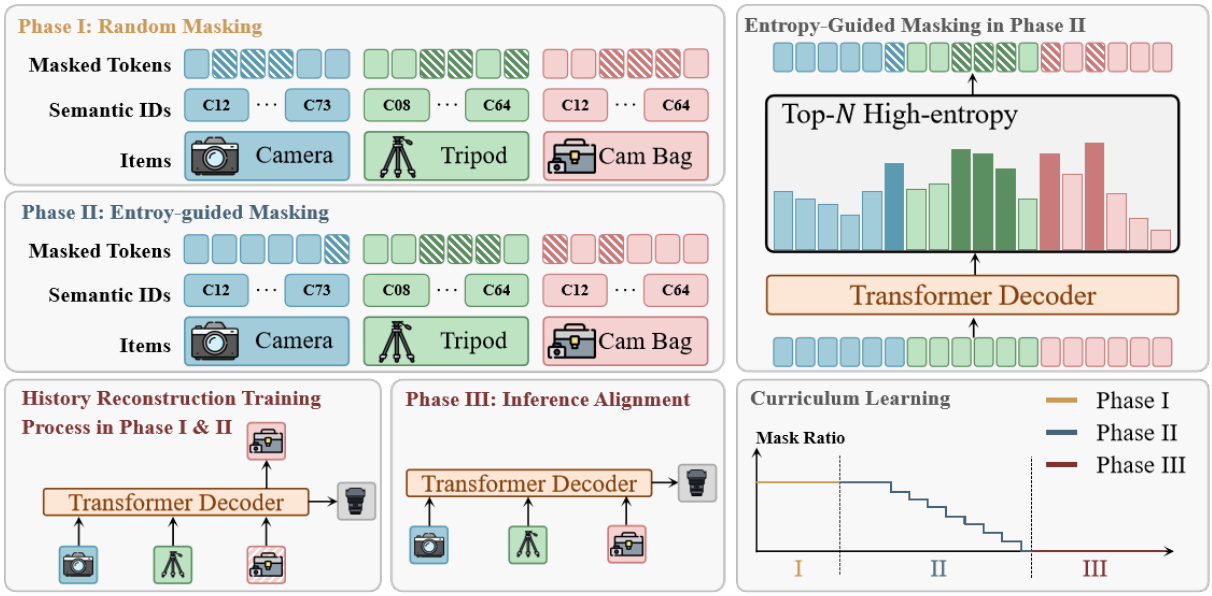

Figure 2: MHL framework overview. Online entropy computation selects challenging positions for masking. Curriculum training progresses from random to adaptive masking, then aligns with inference conditions.

3.2 Masked History Learning

The core idea of MHL is to train the model not only to predict *what comes next*, but also to understand *why the historical path is formed*. We achieve this by jointly optimizing two objectives: (i) next-item prediction and (ii) masked history reconstruction.

Next-Item Prediction. Given the final hidden state $\mathbf{h}_T$, we predict the semantic ID of the next item i_{T+1} using K parallel codebook classifiers:

$$\mathcal{L}_{\text{next}} = -\frac{1}{K}\sum_{k=1}^{K}\log P_\theta^k(w_{i_{T+1}}^k \mid \mathbf{h}_T). \tag{1}$$

Masked History Reconstruction. We mask selected positions in the user history and train the model to reconstruct the original tokens. Let $\mathcal{M}$ denote the set of masked positions. For each masked token $w_{i_t}^k \in \mathcal{M}$, we predict it from the contextualized state at position t:

$$\mathcal{L}_{\text{mask}} = -\frac{1}{|\mathcal{M}|}\sum_{w_{i_t}^k \in \mathcal{M}}\log P_\theta^k(w_{i_t}^k \mid \mathbf{h}_t). \tag{2}$$

This auxiliary objective forces the model to capture logical dependencies between items—not merely statistical co-occurrence, but the underlying intent that connects them.

Multi-Granularity Masking Strategies. We design three masking granularities to provide diverse learning signals:

- **Item-level masking** replaces all K digits of selected items. This treats each item as a complete semantic unit, compelling the model to learn inter-item dependencies that align with user-level behaviors.
- **Token-level masking** replaces individual digits within items. This enables fine-grained learning of intra-item token relationships, yielding more compact representations and better generalization by preventing over-reliance on complete item patterns.
- **Mixed-level masking** stochastically applies item- or token-level masking per item, balancing coarse and fine granularity for robustness.

Overall Objective. The final training objective combines both losses:

$$\mathcal{L}_{\text{MHL}} = \lambda_1\mathcal{L}_{\text{next}} + \lambda_2\mathcal{L}_{\text{mask}}, \tag{3}$$

where λ_1 and λ_2 are hyperparameters balancing next-item generation and history reconstruction.

3.3 Entropy-Guided Masking

Standard random masking treats all historical interactions equally, ignoring the fact that user behaviors have uneven information density. To address this, we propose **Entropy-Guided Masking**, an uncertainty-driven strategy that dynamically targets the most ambiguous and informative positions for reconstruction.

Online Entropy Computation. We employ an **online self-assessment** mechanism to estimate the uncertainty of each historical item. In each training step, we perform a gradient-free forward pass on the unmasked sequence S_T using the current model parameters θ. For each position t and codebook k, the predictive uncertainty is quantified by the Shannon entropy of the output distribution:

$$\mathcal{H}_t^k = -\sum_{v \in \mathcal{W}^k} P_\theta^k(v \mid \mathbf{h}_t) \log P_\theta^k(v \mid \mathbf{h}_t). \tag{4}$$

where $\mathbf{h}_t$ is the hidden state encoding the history up to t. High entropy $\mathcal{H}_t^k$ indicates that the model is uncertain about the item at t based on the context. These high-entropy positions often correspond to **critical decision points** or **complex semantic units**, such as the *tripod* in Figure 1, which is a logical complement to the camera body rather than a generic frequent item.

Adaptive Mask Selection. Based on the computed uncertainty, we adaptively construct the mask set $\mathcal{M}$ using a stochastic budgeting strategy. First, we aggregate token-level entropies into an item-level importance score $\bar{\mathcal{H}}_t = \frac{1}{K}\sum_{k=1}^{K} \mathcal{H}_t^k$. Then, we rank all historical positions in descending order of their importance scores. To prevent the model from overfitting to a fixed masking density, we sample the mask budget N from a uniform distribution $\mathcal{U}(1, \lfloor \gamma \cdot T \rfloor)$, where γ is the current curriculum ratio. Finally, we select the top-N positions with the highest entropy to form $\mathcal{M}$. This adaptive strategy ensures that the reconstruction objective ($\mathcal{L}_{\text{mask}}$) is always applied to the positions where the model lacks understanding, thereby maximizing the gradient information gain and forcing the model to capture deeper bidirectional dependencies.

3.4 Curriculum Training Scheduler

Directly applying entropy-guided masking from the start leads to unstable training, as the model's early-stage entropy estimates are unreliable. Moreover, training with masking creates a discrepancy with inference, where no masking occurs. We address both issues through a three-phase curriculum:

Phase I: Random Masking Warm-up. Training begins with random masking at a low ratio. This phase serves two purposes: (1) establishing stable optimization before entropy estimates become reliable, and (2) building baseline reconstruction ability through simple, unbiased masking.

Phase II: Entropy-Guided Masking with Adaptive Decay. Once the model stabilizes, we switch to entropy-guided masking to focus on the most informative positions. The masking ratio γ starts high to encourage deep historical understanding. When validation performance plateaus, we decay γ exponentially:

$$\gamma \leftarrow \max(\gamma_{\min},\ \gamma \cdot \eta), \tag{5}$$

where $\gamma_{\min}$ is the minimum ratio threshold and $\eta \in (0, 1)$ is the decay factor. This gradually shifts the model's focus from history reconstruction toward future prediction.

Phase III: Inference Alignment. Finally, we set $\gamma = 0$ and train exclusively with $\mathcal{L}_{\text{next}}$. This phase eliminates the train-test discrepancy: since inference involves no masking, inference-aligned training ensures the model is fully aligned with its deployment condition. This smooth transition—from reconstruction to generation—bridges the gap between understanding "why this path matters" and predicting "what comes next."

The complete training procedure is summarized in Algorithm 1 (Appendix C).

4 Experiment

4.1 Experimental Settings

Dataset. We evaluate models on three Amazon product categories: *Sports and Outdoors*, *Beauty*, and *Toys and Games* from the Amazon Reviews 2014 dataset (McAuley et al., 2015). We preprocess each category with core-5 filtering (He and McAuley, 2016). This retains only users and items with at least five interactions to ensure sufficient density for sequential modeling. For item metadata, we concatenate title, price, brand, feature, categories, and description into natural language sentences. This facilitates semantic representation learning following recent practice in generative recommendation (Wang et al., 2024). Table 11 from Appendix G shows detailed dataset statistics.

Baselines. We evaluate against comprehensive baselines in two categories: item ID-based methods and semantic ID-based approaches. Item ID-based methods operate directly on item IDs: GRU4Rec (Hidasi et al., 2016), HGN (Ma et al., 2019), SASRec (Kang and McAuley, 2018), FDSA (Hao et al., 2023), BERT4Rec (Sun et al., 2019), Caser (Tang and Wang, 2018), S^3-Rec (Zhou et al., 2020). Semantic ID-based

Model	Beauty				Toys and Games				Sports and Outdoors			
	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10
Item ID-based												
Caser	.0205	.0131	.0347	.0176	.0166	.0107	.0270	.0141	.0116	.0072	.0194	.0097
GRU4Rec	.0164	.0099	.0283	.0137	.0097	.0059	.0176	.0084	.0129	.0086	.0204	.0110
HGN	.0325	.0206	.0512	.0266	.0321	.0221	.0497	.0277	.0189	.0120	.0313	.0159
BERT4Rec	.0203	.0124	.0347	.0170	.0116	.0071	.0203	.0099	.0115	.0075	.0191	.0099
SASRec	.0387	.0249	.0605	.0318	.0463	.0306	.0675	.0374	.0233	.0154	.0350	.0192
FDSA	.0267	.0163	.0407	.0208	.0228	.0140	.0381	.0189	.0182	.0122	.0288	.0156
S^3-Rec	.0387	.0244	.0647	.0327	.0443	.0294	.0700	.0376	.0251	.0161	.0385	.0204
Semantic ID-based												
RecJPQ	.0311	.0167	.0482	.0222	.0331	.0182	.0484	.0231	.0141	.0076	.0220	.0102
VQ-Rec	.0457	.0317	.0664	.0383	.0497	.0346	.0737	.0423	.0208	.0144	.0300	.0173
TIGER	.0454	.0321	.0648	.0384	.0521	.0371	.0712	.0432	.0264	.0181	.0400	.0225
HSTU	.0469	.0314	.0704	.0389	.0433	.0281	.0669	.0357	.0258	.0165	.0414	.0215
RPG*	.0500	.0358	.0745	.0436	.0550	.0386	.0778	.0460	.0284	.0197	.0436	.0246
MHL (ours)	**.0574**	**.0424**	**.0795**	**.0495**	**.0672**	**.0489**	**.0903**	**.0564**	**.0359**	**.0249**	**.0511**	**.0298**

Table 1: Performance comparison of Item ID-based and Semantic ID-based models across three datasets. * denotes results reproduced using the authors' code and parameters. **Best** and second-best results are bolded and underlined.

approaches tokenize items into discrete semantic identifiers for generative recommendation: VQRec (Hou et al., 2023), RecJPQ (Petrov and Macdonald, 2024), TIGER (Rajput et al., 2023), HSTU (Zhai et al., 2024), RPG (Hou et al., 2025a).

Evaluation Metrics. We evaluate recommendation performance using Recall@K and NDCG@K with K=5 and 10. Following prior works (Kang and McAuley, 2018; Rajput et al., 2023; Sun et al., 2019; Hou et al., 2025a), we adopt standard leave-one-out strategy. For each user sequence, the last item is reserved for testing, the second-to-last for validation, and the remaining items for training.

Implementation Details. We encode item metadata (e.g., title, brand, price) with Sentence-T5-base (Ni et al., 2022). The resulting 768-dimensional embeddings are reduced to 128 dimensions via PCA and then discretized into sequences of 32 semantic tokens using FAISS-based optimized product quantization (OPQ) (Ge et al., 2014). Our backbone is a Transformer decoder, identical to the one used in RPG (Hou et al., 2025a), featuring a hidden size of 448, two layers, and four attention heads. The model is trained to jointly optimize next-item prediction and masked token reconstruction with equal weights. We apply an entropy-guided curriculum masking strategy, and early stopping is used when the mask ratio decays to zero. During optimization, we use AdamW with a learning rate of 5e-4, a batch size of 64, and cosine scheduling with 10k warmup steps. Inference is performed with graph-constrained beam search (Hou et al., 2025a) (beam size 50, 3 propagation steps). The models are trained for up to 300 epochs on NVIDIA RTX A6000 GPUs. More details can be found in Appendix H.

4.2 Experiment Results

Overall Performance. In all experiments, we performed three runs with different random seeds and report the average results to account for variability. Table 1 presents the results across three Amazon product categories. We can find that MHL consistently achieves state-of-the-art performance. In addition, the results confirm that semantic ID-based models outperform traditional item ID-based approaches, with MHL leading all baselines, including strong competitors like TIGER and HSTU. The performance improvements are substantial. For example, MHL achieves a 37.6% improvement over TIGER in the NDCG@5 score for *Sports and Outdoors*. This validates our claim: understanding why a user path is formed is crucial for predicting what comes next. MHL's superior performance demonstrates three key benefits. First, by reconstructing masked historical items, the model learns logical dependencies between items rather than co-occurrence patterns. Second, the entropy-guided masking forces the model to focus on the most informative and challenging positions in user history, precisely where latent intent is obscured. Third, the curriculum learning bridges the gap between history understanding and future prediction, ensuring

Mask Strategy	Curriculum Strategy	Beauty				Toys and Games				Sports and Outdoors			
		R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10
No Mask	Direct Inference	.0500	.0358	.0745	.0436	.0550	.0386	.0778	.0460	.0284	.0197	.0436	.0246
Token-level	Random	.0554	.0379	**.0797**	.0457	**.0647**	.0402	**.0914**	.0488	.0321	.0179	.0482	.0231
	Entropy-guided	.0491	.0338	.0715	.0410	.0232	.0151	.0359	.0191	.0110	.0070	.0171	.0090
	R→Inf	.0554	.0404	.0791	.0480	.0623	.0452	.0876	**.0533**	**.0337**	.0233	**.0486**	.0281
	E→Inf	.0518	.0374	.0736	.0445	.0471	.0322	.0702	.0396	.0264	.0180	.0402	.0224
	R→E→Inf	**.0574**	**.0424**	.0795	**.0495**	.0644	**.0456**	.0883	**.0533**	.0334	**.0239**	.0474	**.0284**
Item-level	Random	.0500	.0355	.0694	.0418	.0564	.0392	.0808	.0470	**.0287**	.0198	**.0420**	.0240
	Entropy-guided	.0441	.0315	.0660	.0386	.0522	.0359	.0750	.0432	.0225	.0151	.0336	.0187
	R→Inf	.0494	.0361	.0695	.0426	**.0575**	**.0410**	**.0816**	**.0487**	.0270	.0188	.0415	.0234
	E→Inf	.0467	.0335	.0672	.0401	.0545	.0376	.0768	.0448	.0286	**.0199**	.0414	.0240
	R→E→Inf	**.0501**	**.0363**	**.0704**	**.0429**	.0553	.0390	.0802	.0470	.0282	.0197	.0419	**.0241**
Mixed-level	Random	.0521	.0359	**.0789**	.0444	.0589	.0393	**.0859**	.0480	.0318	.0197	**.0493**	.0254
	Entropy-guided	.0491	.0344	.0728	.0420	.0566	.0377	.0833	.0463	.0285	.0186	.0428	.0232
	R→Inf	**.0542**	**.0391**	.0752	**.0459**	**.0593**	.0420	.0835	.0498	.0295	.0203	.0435	.0248
	E→Inf	.0520	.0371	.0743	.0442	.0535	.0376	.0761	.0449	.0289	.0202	.0418	.0243
	R→E→Inf	.0537	.0384	.0757	.0455	.0592	**.0421**	.0845	**.0503**	**.0325**	**.0228**	.0478	**.0278**

Table 2: Ablation study comparing masking strategies and curriculum learning approaches with codebook size 16 and mask ratio 0.15. **Best** and second-best results are bolded and underlined.

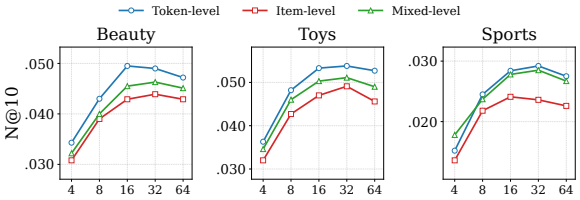

Figure 3: Impact of varying codebook sizes on model performance across three mask strategies (Token-level, Item-level, Mixed-level) on 3 datasets.

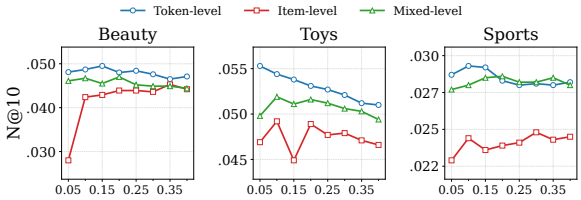

Figure 4: Effect of different mask ratios on model performance under three mask strategies (Token-level, Item-level, Mixed-level) across 3 datasets.

a smooth transition from learning "why this path matters" to predicting "what comes next". These targeted learning mechanisms enable MHL to consistently outperform baselines. The framework's effectiveness is particularly evident on the complex dataset like *Sports and Outdoors*, where logical item relationships are more nuanced and user intent is harder to infer.

Ablation Study. We conduct systematic ablation studies to understand each component's contribution within MHL. We evaluate six model variants across three masking strategies: Direct Inference (Inf) without masking, Random masking (R), and Entropy-guided masking (E). We also test three curriculum learning strategies: R→Inf, E→Inf, and the complete R→E→Inf framework. Table 2 validates our framework design through three key findings. First, all masking variants significantly outperform direct inference, demonstrating that reconstructing user history provides a richer learning signal. Second, entropy-guided masking consistently surpasses random masking, indicating that targeting high-entropy predictions is more effective for guiding the model to understand user intent. Finally, the complete R→E→Inf curriculum learning framework achieves optimal performance.

This validates our curriculum design: starting with basic pattern learning through random masking, progressing to targeted understanding via entropy guidance, and finally inference-only optimization. This progression mirrors the learning objective of transitioning from "why this path matters" to "what comes next".

4.3 Further Analysis

Impact of Semantic ID Length. We study the effect of codebook size (4–64) on model performance. As shown in Figure 3, performance improves as codebook size increases from 4 to 32, with sizes 16 and 32 yielding optimal dataset-wide results. Size 4 degrades due to insufficient semantic granularity, while size 64 declines slightly from sparsity issues. Based on these findings, we adopt size 16 for Beauty and 32 for Toys and Sports. Detailed results are in Appendix F.1.

Sensitivity to Masking Ratio. We vary the masking ratio from 0.05 to 0.40 across three masking strategies. As Figure 4 shown, MHL demonstrates

Mask Strategy	Training Method	Beauty					Toys and Games					Sports and Outdoors				
		R@5	N@5	R@10	N@10	Δ%	R@5	N@5	R@10	N@10	Δ%	R@5	N@5	R@10	N@10	Δ%
Text w/o Mask	RPG	.0297	.0212	.0439	.0258	–	.0323	.0234	.0446	.0273	–	.0134	.0094	.0203	.0117	–
MHL (ours)	R→E→Inf	.0338	.0238	.0483	.0285	**+10.5**	.0347	.0249	.0498	.0297	**+8.8**	.0150	.0106	.0237	.0134	**+14.5**

Table 3: Generalization study comparing MHL with RPG baseline on text token sequences using mixed masking. MHL uses mask ratio 0.15 and reconstruction loss weight 0.5. Δ% denotes the relative improvement of N@10.

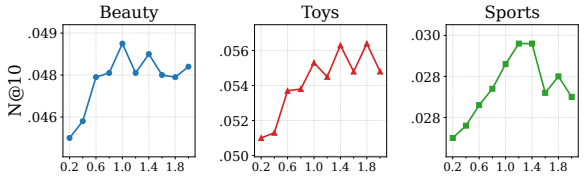

Figure 5: Evaluation of model performance under different reconstruction loss values across multiple datasets, with a focus on N@10 performance.

Historical Purchase Sequence	
Footwear Adhesive → Running Waist Pack → Cardio Trampoline → Heavyweight T-Shirt → BMX Pads → ?	
MHL (Top-5)	**RPG (Top-5)**
Youth Multi-Sport Helmet ✓	Crew Sock
NBA Street Basketball	Eco Open Bottom Pant
Mini Basketball Hoop	Training T-shirt
Indoor/Outdoor Basketball	Jersey Pants
NBA Game Ball Mini	Long Sleeve Cotton T-Shirt

Table 4: Case study comparison between MHL and the RPG baseline.Top-5 recommendations are listed in descending order of predicted relevance.

strong robustness to ratio variations. Token-level masking with lower ratios (0.05–0.15) achieves the best overall performance, while mixed-level masking provides a balanced solution within the 0.10–0.20 range. Detailed results are in Appendix F.2.

Effect of Reconstruction Loss Weight. We vary the reconstruction loss weight from 0.2 to 2.0 while keeping the prediction loss fixed at 1.0. As shown in Figure 5, optimal weights are dataset-dependent but consistently fall within 1.0–1.8. Performance degrades when the weight drops below 0.8, indicating that reconstruction serves as an effective auxiliary signal requiring careful balancing. Detailed results are in Appendix F.3.

Generalizability to Text Sequences. To demonstrate MHL's broad applicability, we apply our framework to unstructured token sequences derived from item titles. Table 3 compares our complete R→E→Inf strategy against the RPG baseline on raw text tokens. MHL consistently outperforms RPG across all metrics on the Beauty dataset. This result is significant as it shows that MHL's core principle—reconstructing the past to predict the future—generalizes beyond semantic IDs to complex and noisy text sequences. The success on text sequences validates that MHL captures fundamental learning dynamics rather than exploiting specific properties of semantic ID representations. This generalizability highlights MHL's potential for broader sequential modeling applications where historical context understanding is key for future prediction.

Case Study. As illustrated in Table 4, the baseline RPG model, trained solely on autoregressive next-item prediction, often misinterprets a user's intent by overemphasizing transient, noisy signals, such as the mid-sequence clothing items. For example, its predictions for items like "Crew Sock" deviate from the user's primary and recurring interest in athletic gear and accessories. In contrast, MHL framework requires the model to reconstruct a user's historical trajectory, and encourages it to identify and prioritize the core underlying intent. MHL can look beyond short-term deviations and accurately predict the next item "Youth Multi-Sport Helmet", which aligns logically with the user's sustained interest in firearm-related products.

5 Conclusion

Existing generative recommenders focus on predicting "what comes next" but fail to understand "why this path matters". We introduce MHL, a simple and effective framework that learns from masked history reconstruction alongside next-item prediction. MHL incorporates entropy-guided masking to target informative historical positions and curriculum learning to transition from history understanding to future prediction. Experiments on three datasets show state-of-the-art performance and successful generalization to text-based Item IDs. Our findings confirm that understanding the past is crucial for predicting the future. MHL represents a significant step toward recommendation systems that comprehend user behavior patterns rather than merely statistical co-occurrence.

Limitations

MHL constructs semantic IDs solely from textual item metadata and does not explicitly incorporate visual or other multi-modal signals. Nevertheless, the proposed masking and reconstruction mechanism operates on discrete tokens and is modality-agnostic, making it applicable to multi-modal features once encoded into codebook-based representations; here we focus on text-only inputs to isolate the effect of masked history learning.

Entropy-guided masking introduces additional training-time cost due to an extra forward pass for uncertainty estimation. While this increases training time, it provides more informative supervision and leads to consistent performance gains, and the overhead is confined to training without affecting inference efficiency.

Ethics Statement

This paper uses publicly available pretrained models and datasets. The datasets employed in this work are widely adopted in the research community and contain no private user data or personally identifiable information. All models are evaluated as-is, without any additional training or fine-tuning that could amplify harmful behaviors. We therefore believe that this study complies with the ACL Ethics Policy.

References

Keqin Bao, Jizhi Zhang, Yang Zhang, Wenjie Wang, Fuli Feng, and Xiangnan He. 2023. Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems (RecSys)*, pages 1007–1014.

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning (ICML)*, pages 41–48.

Jianxin Chang, Chen Gao, Yu Zheng, Yiqun Hui, Yanan Niu, Yang Song, Depeng Jin, and Yong Li. 2021. Sequential recommendation with graph neural networks. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 378–387.

Rui Chen, Qingyi Hua, Yan-Shuo Chang, Bo Wang, Lei Zhang, and Xiangjie Kong. 2018. A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks. *IEEE Access*, 6:64301–64320.

Yunfan Gao, Tao Sheng, Youlin Xiang, Yun Xiong, Haofen Wang, and Jiawei Zhang. 2023. Chat-rec: Towards interactive and explainable llms-augmented recommender system. *arXiv preprint arXiv:2303.14524*.

Tiezheng Ge, Kaiming He, Qifa Ke, and Jian Sun. 2014. Optimized product quantization. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 36(4):744–755.

Yongjing Hao, Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Guanfeng Liu, and Xiaofang Zhou. 2023. Feature-level deeper self-attention network with contrastive learning for sequential recommendation. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 35(10):10112–10124.

Jesse Harte, Wouter Zorgdrager, Panos Louridas, Asterios Katsifodimos, Dietmar Jannach, and Marios Fragkoulis. 2023. Leveraging large language models for sequential recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems (RecSys)*, pages 1096–1102.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pages 770–778.

Ruining He and Julian McAuley. 2016. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on World Wide Web (WWW)*, pages 507–517.

Zhicheng He, Weiwen Liu, Wei Guo, Jiarui Qin, Yingxue Zhang, Yaochen Hu, and Ruiming Tang. 2023. A survey on user behavior modeling in recommender systems. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI*, pages 6656–6664.

Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based recommendations with recurrent neural networks. In *4th International Conference on Learning Representations (ICLR)*.

Yupeng Hou, Zhankui He, Julian McAuley, and Wayne Xin Zhao. 2023. Learning vector-quantized item representation for transferable sequential recommenders. In *Proceedings of the ACM Web Conference 2023 (WWW)*, pages 1162–1171.

Yupeng Hou, Jiacheng Li, Ashley Shin, Jinsung Jeon, Abhishek Santhanam, Wei Shao, Kaveh Hassani, Ning Yao, and Julian McAuley. 2025a. Generating long semantic ids in parallel for recommendation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pages 956–966.

Yupeng Hou, An Zhang, Leheng Sheng, Zhengyi Yang, Xiang Wang, Tat-Seng Chua, and Julian McAuley. 2025b. Generative recommendation models: Progress and directions. In *Companion Proceedings of the ACM on Web Conference 2025 (WWW)*, pages 13–16.

Wenyue Hua, Shuyuan Xu, Yingqiang Ge, and Yongfeng Zhang. 2023. How to index item ids for recommendation foundation models. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region (SIGIR-AP)*, pages 195–204.

Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, pages 197–206.

Jing Li, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tao Lian, and Jun Ma. 2017. Neural attentive session-based recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM)*, pages 1419–1428.

Pang Li, Shahrul Azman Mohd Noah, and Hafiz Mohd Sarim. 2024. A survey on deep neural networks in collaborative filtering recommendation systems. *arXiv preprint arXiv:2412.01378*.

Chen Ma, Peng Kang, and Xue Liu. 2019. Hierarchical gating networks for sequential recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, pages 825–833.

David J. C. MacKay. 2002. *Information Theory, Inference & Learning Algorithms*. Cambridge University Press, USA.

Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton van den Hengel. 2015. Image-based recommendations on styles and substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 43–52.

Niklas Muennighoff, Hongjin SU, Liang Wang, Nan Yang, Furu Wei, Tao Yu, Amanpreet Singh, and Douwe Kiela. 2025. Generative representational instruction tuning. In *The Thirteenth International Conference on Learning Representations (ICLR)*.

Jianmo Ni, Gustavo Hernandez Abrego, Noah Constant, Ji Ma, Keith Hall, Daniel Cer, and Yinfei Yang. 2022. Sentence-t5: Scalable sentence encoders from pretrained text-to-text models. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 1864–1874.

Aleksandr V. Petrov and Craig Macdonald. 2024. Recjpq: Training large-catalogue sequential recommenders. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM)*, pages 538–547.

Erasmo Purificato, Ludovico Boratto, and Ernesto William De Luca. 2024. User modeling and user profiling: A comprehensive survey. *arXiv preprint arXiv:2402.09660*.

Shashank Rajput, Nikhil Mehta, Anima Singh, Raghunandan Hulikal Keshavan, Trung Vu, Lukasz Heldt, Lichan Hong, Yi Tay, Vinh Q. Tran, Jonah Samost, Maciej Kula, Ed H. Chi, and Maheswaran Sathiamoorthy. 2023. Recommender systems with generative retrieval. In *Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*.

Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web (WWW)*, pages 811–820.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 1441–1450.

Jiaxi Tang and Ke Wang. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM)*, pages 565–573.

Yi Tay, Vinh Q. Tran, Mostafa Dehghani, Jianmo Ni, Dara Bahri, Harsh Mehta, Zhen Qin, Kai Hui, Zhe Zhao, Jai Gupta, Tal Schuster, William W. Cohen, and Donald Metzler. 2022. Transformer memory as a differentiable search index. In *Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS)*.

Hao Wang, Naiyan Wang, and Dit-Yan Yeung. 2015. Collaborative deep learning for recommender systems. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 1235–1244.

Wenjie Wang, Honghui Bao, Xinyu Lin, Jizhi Zhang, Yongqi Li, Fuli Feng, See-Kiong Ng, and Tat-Seng Chua. 2024. Learnable item tokenization for generative recommendation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM)*, pages 2400–2409.

Yujing Wang, Yingyan Hou, Haonan Wang, Ziming Miao, Shibin Wu, Hao Sun, Qi Chen, Yuqing Xia, Chengmin Chi, Guoshuai Zhao, Zheng Liu, Xing Xie, Hao Allen Sun, Weiwei Deng, Qi Zhang, and Mao Yang. 2022. A neural corpus indexer for document retrieval. In *Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS)*.

Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-based recommendation with graph neural networks. In *Proceedings of*

the AAAI Conference on Artificial Intelligence, pages 346–353.

Zheng Yuan, Fajie Yuan, Yu Song, Youhua Li, Junchen Fu, Fei Yang, Yunzhu Pan, and Yongxin Ni. 2023. Where to go next for recommender systems? id- vs. modality-based recommender models revisited. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 2639–2649.

Zhenrui Yue, Yueqi Wang, Zhankui He, Huimin Zeng, Julian Mcauley, and Dong Wang. 2024. Linear recurrent units for sequential recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM)*, pages 930–938.

Jiaqi Zhai, Lucy Liao, Xing Liu, Yueming Wang, Rui Li, Xuan Cao, Leon Gao, Zhaojie Gong, Fangda Gu, Jiayuan He, Yinghai Lu, and Yu Shi. 2024. Actions speak louder than words: Trillion-parameter sequential transducers for generative recommendations. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, pages 58484–58509.

Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM)*, pages 1893–1902.

A Vocabulary Construction Details

Let $\mathcal{I}$ denote the item set. Each item $i \in \mathcal{I}$ is represented by a K-digit semantic ID using K codebooks $\{\mathcal{W}^1, \ldots, \mathcal{W}^K\}$, where $|\mathcal{W}^k| = V_k$ denotes the vocabulary size of the k-th codebook:

$$\phi(i) = \{w_i^1, \ldots, w_i^K\}, \quad w_i^k \in \mathcal{W}^k. \tag{6}$$

We construct a unified vocabulary that contains both semantic tokens and mask tokens:

$$\mathcal{V} = \mathcal{V}_{\text{code}} \cup \mathcal{V}_{\text{mask}}, \tag{7}$$

where $\mathcal{V}_{\text{code}} = \bigcup_{k=1}^{K} \mathcal{V}_{\text{code}}^k$ and each codebook block $\mathcal{V}_{\text{code}}^k$ occupies a non-overlapping ID range:

$$\mathcal{V}_{\text{code}}^k = \Big[1 + \sum_{j<k} V_j,\ \sum_{j\le k} V_j\Big]. \tag{8}$$

We allocate position-aware mask tokens $\mathcal{V}_{\text{mask}}$ for each position-digit pair (t, k) up to the maximum sequence length $T_{\max}$:

$$\mathcal{V}_{\text{mask}} = \Big[\sum_{k=1}^{K} V_k + 1,\ \sum_{k=1}^{K} V_k + K \cdot T_{\max}\Big], \tag{9}$$

ensuring that each masked position can be uniquely identified, making the masking process lossless and unambiguous.

B Model Architecture Details

Token Embedding. Each token ID in $\mathcal{V}$ is mapped to a d-dimensional embedding by a shared lookup table (word token embedding, WTE). For a user history $S_T = (i_1, \ldots, i_T)$, we expand each item into its K semantic tokens $\{w_{i_t}^k\}_{k=1}^K$ and retrieve their embeddings:

$$e_t^k = \text{WTE}(w_{i_t}^k) \in \mathbb{R}^d. \tag{10}$$

Item-level Representation via Mean Pooling. We aggregate each item's K digit embeddings into a single item-level representation:

$$H_t = \frac{1}{K} \sum_{k=1}^{K} e_t^k \in \mathbb{R}^d. \tag{11}$$

Sequence Modeling. The item representations $\{H_1, \ldots, H_T\}$ are fed into a Transformer decoder to obtain contextualized hidden states:

$$\mathbf{h}_{1:T} = \text{Dec}(H_{1:T}), \quad \mathbf{h}_t \in \mathbb{R}^d. \tag{12}$$

Digit-wise Prediction Heads. To predict each semantic digit independently, we use K lightweight prediction heads $\{g_k\}_{k=1}^K$ that derive digit-specific states from the shared item state:

$$\mathbf{h}_t^k = g_k(\mathbf{h}_t), \quad k = 1, \ldots, K, \quad \mathbf{h}_t^k \in \mathbb{R}^d. \tag{13}$$

Cosine Classifier with Temperature Scaling. For codebook k, let $E^k \in \mathbb{R}^{V_k \times d}$ denote the embedding matrix where each row corresponds to a token embedding. We compute prediction logits using temperature-scaled cosine similarity. Both the digit state and token embeddings are L2-normalized:

$$\hat{\mathbf{h}}_t^k = \frac{\mathbf{h}_t^k}{\|\mathbf{h}_t^k\|_2}, \quad \hat{E}_v^k = \frac{E_v^k}{\|E_v^k\|_2}, \tag{14}$$

$$\boldsymbol{\ell}_t^k = \frac{\hat{\mathbf{h}}_t^k (\hat{E}^k)^\top}{\tau} \in \mathbb{R}^{V_k}, \tag{15}$$

where τ is the temperature hyperparameter. The probability distribution over codebook $\mathcal{W}^k$ is then:

$$P_\theta^k(w = v \mid \mathbf{h}_t) = \text{softmax}(\boldsymbol{\ell}_t^k)_v. \tag{16}$$

C Training Algorithm

Algorithm 1 summarizes the complete MHL training procedure with curriculum scheduling.

D Notation Summary

Table 5 summarizes the key notations used throughout this paper.

E Pilot Experiment

To illustrate the motivation of this paper, we conduct a pilot experiment to examine whether models trained with the standard next-item prediction paradigm overly rely on recent interactions, potentially neglecting the user's past behaviors. Specifically, on the *Toys and Games* dataset, for sequences longer than 20 in the test set, we truncate each sequence by removing the last 15 items. For example, given an original sequence $[i_1, i_2, \ldots, i_{25}] \rightarrow i_{26}$, the truncated version becomes $[i_1, i_2, \ldots, i_{10}] \rightarrow i_{11}$. This setup evaluates how well models capture long-range dependencies when recent interactions are unavailable, rather than relying on short-term patterns near the target item.

The results of full-sequence and truncated-sequence evaluation are shown in Table 9. Under the full sequence setting, MHL outperforms RPG across all metrics, achieving an 18.23% improvement on N@10. In the truncated setting, which emphasizes longer-range dependencies, the improvement is even larger, reaching 43.95%, indicating that MHL not only captures both short-term and long-term user preferences, but also better understands the overall sequence context. This comparison demonstrates that MHL more effectively models user behavior, whereas RPG tends to rely more heavily on recent interactions.

To further validate MHL's ability to capture long-term user intent, we conduct a length-stratified analysis. We bucket the test sequences by length and compute N@10 for both RPG and MHL. The performance trends are visualized in Figure 6, with detailed numerical results reported in Table 10. RPG performs reasonably well on medium-length sequences (30–50 items) but struggles on very short and very long sequences. For instance, sequences longer than 50 items see RPG's N@10 drop to 0.0375, while MHL boosts it to 0.0577, corresponding to a 53.33% relative improvement. Overall, MHL consistently outperforms RPG across all length buckets, and the relative improvement is most pronounced for the extremely long sequences. These results confirm that MHL effectively captures long-term user preferences rather than relying primarily on recent interactions, further supporting the motivation for our proposed approach.

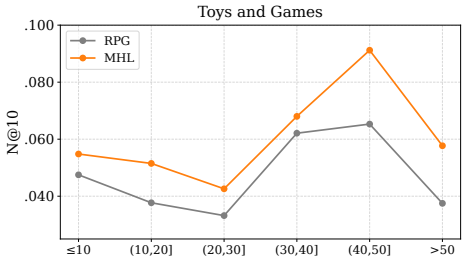

Figure 6: Length-stratified N@10 comparison on Toys and Games. MHL adopts a 16-bit codebook consistent with RPG, and employs a mask ratio of 0.10. MHL outperforms RPG across all sequence length bins.

F Hyperparameter Sensitivity Analysis

F.1 Codebook Size

We investigate the influence of codebook size on model efficacy by varying the size within the range of 4 to 64. Performance exhibits a consistent upward trend as the codebook size expands from 4 to 32 across all three masking strategies and datasets. Notably, codebook sizes of 16 and 32 yield the optimal results; specifically, size 16 proves marginally superior on the Beauty dataset, whereas 32 dominates on Toys and Sports. Conversely, a size of 4 leads to significant performance deterioration due to inadequate semantic granularity, while size 64 incurs a slight decline, likely stemming from sparsity issues and increased optimization difficulty caused by excessive capacity.

These results suggest that codebook size should balance expressiveness and learnability. Based on these findings, we adopt codebook sizes of 16 for Beauty and 32 for Toys and Sports in our main experiments. Table 6 provides a detailed performance comparison across different codebook sizes with a mask ratio of 0.15, including results for three mask strategies (Token-level, Item-level, and Mixed-level) on all three datasets.

F.2 Mask Ratio

We investigate the impact of the masking ratio on model robustness by varying it from 0.05 to 0.40. MHL maintains relatively stable performance across three masking strategies (Token-level, Item-level, and Mixed-level) and all datasets (Beauty, Toys, and Sports), utilizing dataset-specific codebook sizes. Notably, while the model exhibits slight performance fluctuations depending on the specific strategy and dataset, it generally demonstrates strong resilience to ratio variations.

For token-level masking, lower ratios between

0.05 and 0.15 yield consistently strong performance, with 0.05 achieving optimal results on Toys and Games and 0.10–0.15 performing best on Beauty and Sports and Outdoors. Item-level masking shows greater variability across datasets, with optimal ratios ranging from 0.10 to 0.35 depending on the specific dataset. Mixed-level masking demonstrates stable performance within the 0.10–0.20 range, often achieving competitive results across all datasets. While mixed-level masking provides a balanced and robust solution, token-level masking with lower ratios tends to achieve the highest overall performance, particularly on datasets with richer textual information.

The observed stability across mask ratios indicates that MHL is robust to variations in masking, making it suitable for real-world deployment where flexibility in masking strategies is important. Table 7 presents detailed performance across different mask ratios for all masking strategies and datasets.

F.3 Reconstruction Loss

To investigate the contribution of the reconstruction objective in MHL, we vary the reconstruction loss weight from 0.2 to 2.0 while keeping the prediction loss fixed at 1.0. We observe that the optimal reconstruction weight is dataset-dependent: Beauty achieves best performance at 1.0, Toys and Games peaks at 1.8, and Sports and Outdoors performs optimally around 1.2–1.4. Despite this variation, the best results consistently fall within the range of 1.0–1.8, suggesting that a moderate-to-high reconstruction weight is generally beneficial.

Performance degrades noticeably when the weight drops below 0.8, indicating that insufficient reconstruction supervision weakens the learned token representations. Conversely, excessively high weights (e.g., 2.0) do not yield further improvements and may even hurt performance, likely due to the reconstruction objective overshadowing the prediction task. These findings highlight that reconstruction loss serves as an effective auxiliary signal that complements the primary prediction objective, but requires careful balancing to maximize performance. Table 8 reports the specific performance metrics across all reconstruction loss values.

G Statistics of the Dataset

The detailed statistics of Amazon Reviews 2014 datasets is shown in Table 11.

H Implementation Details

We encode item metadata (title, brand, price, features, categories, description) using Sentence-T5 and reduce 768-dimensional embeddings to 128 dimensions with PCA. Following RPG (Hou et al., 2025a), we discretize continuous representations into generative semantic IDs using FAISS-based OPQ. Each item is represented as a sequence of 32 tokens (32 codebooks with 256 codewords each). Our backbone is a Transformer decoder with the same parameter size as RPG (Hou et al., 2025a): hidden size 448, 2 layers, 4 attention heads, feed-forward dimension 1024, and GELU activation. The maximum sequence length is 50 with dropout 0.3 for embeddings and attention modules.

For training, we jointly optimize next-item prediction and masked token reconstruction with equal weights. We use entropy-guided curriculum masking: training starts with random masking, then switches to entropy-based masking; if validation does not improve for 5 consecutive evaluations, mask ratio decays linearly by $0.1 \times r_0$ (with $r_0 = 0.15$) until reaching 0. After this, the model trains purely on prediction with early stopping patience of 20. Entropy forward propagation stabilizes masking decisions using window size 3, decay factor 2.0, and residual mixing coefficient 0.2 across item-level and token-level entropies.

During inference, we follow RPG (Hou et al., 2025a) and apply graph-constrained beam search with beam size 50, each node keeping 50 edges, and 3 propagation steps. Optimization uses AdamW with learning rate 5e-4, batch size 64, weight decay 0.0, gradient clipping 1.0, 10k warmup steps, and cosine learning rate scheduling. We train for up to 300 epochs with early stopping patience of 20. All experiments use NVIDIA RTX A6000 GPUs with distributed training and mixed precision.

Algorithm 1 MHL Curriculum Training

Require: Training set $\mathcal{D}$, initial mask ratio γ_0, decay factor η, minimum ratio $\gamma_{\min}$, patience p

Initialize model parameters θ
$\gamma \leftarrow \gamma_0$, stage $\leftarrow$ `RANDOM`
// Phase I: Random Masking Warm-up
while stage $=$ `RANDOM` **do**
  for each batch $\mathcal{B} \in \mathcal{D}$ **do**
    Apply random masking with ratio γ
    Compute $\mathcal{L}_{\text{MHL}} = \lambda_1 \mathcal{L}_{\text{next}} + \lambda_2 \mathcal{L}_{\text{mask}}$
    Update θ via gradient descent
  end for
  if validation loss converges for p epochs **then**
    stage $\leftarrow$ `ENTROPY`
  end if
end while
// Phase II: Entropy-Guided Masking with Decay
while stage $=$ `ENTROPY` **and** $\gamma > 0$ **do**
  for each batch $\mathcal{B} \in \mathcal{D}$ **do**
    Compute entropy $\mathcal{H}_t^k$ via gradient-free forward pass
    Mask top-N highest-entropy positions
    Compute $\mathcal{L}_{\text{MHL}}$ and update θ
  end for
  if validation loss plateaus for p epochs **then**
    $\gamma \leftarrow \max(\gamma_{\min}, \gamma \cdot \eta)$
  end if
  if $\gamma \leq \gamma_{\min}$ **then**
    $\gamma \leftarrow 0$, stage $\leftarrow$ `PREDICT`
  end if
end while
// Phase III: Inference Alignment
while not converged **do**
  for each batch $\mathcal{B} \in \mathcal{D}$ **do**
    Compute $\mathcal{L}_{\text{next}}$ only (no masking)
    Update θ via gradient descent
  end for
end while
return Trained model parameters θ

Symbol	Description
Item Representation	
$\mathcal{I}$	Item set
i	An item in $\mathcal{I}$
K	Number of codebooks (digits per item)
$\mathcal{W}^k$	The k-th codebook vocabulary
V_k	Size of the k-th codebook, $\lvert\mathcal{W}^k\rvert$
$\phi(i)$	Semantic ID of item i
w_i^k	The k-th digit of item i's semantic ID
Vocabulary	
$\mathcal{V}$	Unified token vocabulary
$\mathcal{V}_{\text{code}}, \mathcal{V}_{\text{mask}}$	Codebook tokens and mask tokens
$T_{\max}$	Maximum sequence length
Sequence and Embeddings	
S_T	User history sequence of length T
T	Sequence length
t	Position index in sequence
d	Embedding dimension
e_t^k	Token embedding for position t, digit k
H_t	Item-level representation (mean-pooled)
Model Components	
θ	Model parameters
$\mathbf{h}_t$	Contextualized hidden state from decoder
$\mathbf{h}_t^k$	Digit-wise hidden state from head g_k
g_k	The k-th prediction head
E^k	Embedding matrix for codebook k
τ	Temperature for cosine classifier
$\boldsymbol{\ell}_t^k$	Prediction logits for position t, digit k
P_θ^k	Prediction distribution over $\mathcal{W}^k$
Training Objectives	
$\mathcal{M}$	Set of masked positions
$\mathcal{L}_{\text{next}}$	Next-item prediction loss
$\mathcal{L}_{\text{mask}}$	Masked reconstruction loss
$\mathcal{L}_{\text{MHL}}$	Combined MHL loss
λ_1, λ_2	Loss balancing weights
Entropy-Guided Masking	
$\mathcal{H}_t^k$	Entropy at position t, digit k
$\bar{\mathcal{H}}_t$	Item-level entropy (averaged)
Curriculum Training	
γ	Current masking ratio
γ_0	Initial masking ratio
$\gamma_{\min}$	Minimum masking ratio threshold
η	Decay factor for masking ratio
N	Number of positions to mask (budget)
p	Patience for early stopping

Table 5: Summary of notations.

Mask Strategy	Codebook Size	Beauty				Toys and Games				Sports and Outdoors			
		R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10
Token-level	4	.0406	.0291	.0568	.0343	.0434	.0298	.0638	.0363	.0177	.0120	.0277	.0152
	8	.0502	.0369	.0691	.0430	.0577	.0409	.0805	.0482	.0280	.0209	.0395	.0245
	16	**.0574**	**.0424**	.0795	**.0495**	**.0644**	.0456	**.0883**	.0533	.0334	.0239	.0474	.0284
	32	.0568	.0411	**.0814**	.0490	.0634	**.0458**	.0880	**.0538**	**.0344**	**.0244**	**.0495**	**.0292**
	64	.0552	.0390	.0805	.0472	.0618	.0448	.0861	.0527	.0325	.0226	.0475	.0275
Item-level	4	.0344	.0253	.0516	.0308	.0385	.0259	.0576	.0320	.0151	.0107	.0242	.0136
	8	.0458	.0329	.0646	.0390	.0503	.0359	.0716	.0427	.0249	.0180	.0366	.0218
	16	.0501	.0363	.0704	.0429	.0553	.0390	.0802	.0470	.0282	**.0197**	.0419	**.0241**
	32	**.0523**	**.0368**	**.0743**	**.0439**	**.0579**	**.0416**	**.0811**	**.0491**	**.0286**	.0193	**.0421**	.0236
	64	.0509	.0358	.0731	.0429	.0546	.0384	.0768	.0456	.0273	.0179	.0417	.0226
Mixed-level	4	.0376	.0272	.0531	.0322	.0422	.0287	.0603	.0346	.0203	.0146	.0302	.0178
	8	.0468	.0343	.0645	.0400	.0553	.0385	.0785	.0460	.0279	.0201	.0389	.0237
	16	**.0537**	**.0384**	.0757	.0455	.0592	.0421	.0845	.0503	.0325	.0228	.0478	.0278
	32	**.0537**	.0383	**.0784**	**.0463**	**.0613**	**.0434**	**.0855**	**.0511**	**.0334**	**.0235**	**.0489**	**.0285**
	64	.0533	.0377	.0760	.0451	.0579	.0411	.0824	.0490	.0318	.0218	.0471	.0267

Table 6: Performance comparison across different codebook sizes with mask ratio 0.15.

Mask Strategy	Mask Ratio	Beauty				Toys and Games				Sports and Outdoors			
		R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10
Token-level	0.05	.0562	.0410	.0782	.0481	**.0657**	**.0476**	**.0898**	**.0553**	.0341	.0236	**.0499**	.0287
	0.10	.0572	.0412	**.0805**	.0487	.0632	.0463	.0885	.0544	**.0346**	**.0246**	.0493	**.0293**
	0.15	**.0574**	**.0424**	.0795	**.0495**	.0634	.0458	.0880	.0538	.0344	.0244	.0495	.0292
	0.20	.0549	.0403	.0788	.0480	.0628	.0457	.0858	.0531	.0339	.0240	.0473	.0283
	0.25	.0563	.0409	.0795	.0484	.0624	.0447	.0873	.0527	.0339	.0239	.0467	.0280
	0.30	.0558	.0407	.0774	.0476	.0613	.0443	.0854	.0521	.0333	.0238	.0466	.0281
	0.35	.0532	.0389	.0767	.0465	.0615	.0439	.0841	.0512	.0333	.0236	.0472	.0280
	0.40	.0551	.0399	.0772	.0471	.0593	.0429	.0844	.0510	.0336	.0237	.0474	.0282
Item-level	0.05	.0327	.0225	.0499	.0280	.0565	.0400	.0778	.0469	.0277	.0183	.0418	.0229
	0.10	.0495	.0359	.0700	.0424	**.0583**	**.0420**	.0810	**.0492**	**.0301**	**.0201**	.0433	.0244
	0.15	.0501	.0363	.0704	.0429	.0553	.0384	.0755	.0449	.0286	.0193	.0421	.0236
	0.20	.0507	.0372	.0716	.0439	.0568	.0408	**.0819**	.0489	.0290	.0195	.0426	.0239
	0.25	.0511	.0372	.0717	.0439	.0568	.0403	.0800	.0477	.0292	.0193	.0439	.0241
	0.30	.0510	.0370	.0715	.0436	.0580	.0404	.0815	.0479	.0295	.0200	**.0447**	**.0248**
	0.35	**.0521**	**.0379**	**.0749**	**.0453**	.0571	.0395	.0807	.0471	.0285	.0192	.0446	.0243
	0.40	.0516	.0370	.0740	.0442	.0555	.0384	.0806	.0466	.0296	.0199	.0439	.0245
Mixed-level	0.05	.0544	.0392	.0758	.0461	.0596	.0426	.0821	.0498	.0323	.0225	.0485	.0277
	0.10	.0548	.0397	.0767	.0467	.0610	**.0440**	.0855	**.0519**	**.0342**	.0236	.0483	.0280
	0.15	.0537	.0384	.0757	.0455	.0613	.0434	.0855	.0511	.0334	.0235	.0489	.0285
	0.20	**.0560**	**.0401**	**.0775**	**.0470**	**.0617**	**.0440**	.0853	.0516	**.0342**	**.0238**	**.0490**	**.0286**
	0.25	.0529	.0382	.0746	.0452	**.0617**	.0433	**.0860**	.0512	.0335	.0235	.0480	.0282
	0.30	.0519	.0377	.0743	.0449	.0612	.0428	.0854	.0506	.0335	.0232	.0489	.0282
	0.35	.0517	.0372	.0757	.0449	.0598	.0422	.0848	.0503	.0340	.0237	**.0490**	.0285
	0.40	.0516	.0368	.0753	.0444	.0590	.0418	.0825	.0494	.0329	.0228	**.0490**	.0280

Table 7: Performance Sensitivity Analysis across Different Mask Ratios with Dataset-Specific Codebook Sizes (16 for Beauty, 32 for Toys and Games / Sports and Outdoors).

Mask Strategy	Reconstruction Loss	Beauty				Toys and Games				Sports and Outdoors			
		R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10	R@5	N@5	R@10	N@10
Token-level	2.0	.0562	.0414	.0779	.0484	.0638	.0475	.0868	.0548	.0337	.0241	.0472	.0285
	1.8	.0560	.0411	.0769	.0479	**.0672**	**.0489**	**.0903**	**.0564**	.0347	.0246	.0484	.0290
	1.6	.0564	.0415	.0767	.0480	.0653	.0476	.0877	.0548	.0344	.0241	.0484	.0286
	1.4	.0559	.0413	**.0800**	.0490	.0664	**.0489**	.0896	.0563	**.0361**	**.0253**	.0503	**.0298**
	1.2	.0559	.0412	.0774	.0481	.0646	.0474	.0867	.0545	.0359	.0249	**.0511**	**.0298**
	1.0	**.0574**	**.0424**	.0795	**.0495**	.0657	.0476	.0898	.0553	.0346	.0246	.0493	.0293
	0.8	.0554	.0404	.0793	.0481	.0642	.0458	.0890	.0538	.0348	.0244	.0482	.0287
	0.6	.0549	.0401	.0790	.0479	.0640	.0460	.0881	.0537	.0331	.0235	.0478	.0283
	0.4	.0533	.0382	.0768	.0458	.0611	.0435	.0856	.0513	.0330	.0235	.0465	.0278
	0.2	.0526	.0376	.0758	.0450	.0602	.0429	.0854	.0510	.0330	.0224	.0487	.0275

Table 8: Token-level performance under varying reconstruction loss values (predict loss fixed at 1.0). Mask ratios: 0.15 for Beauty, 0.05 for Toys, 0.10 for Sports; codebook sizes: 16 for Beauty and 32 for Toys/Sports.

Setting	Model	Toys and Games				Δ%
		R@5	N@5	R@10	N@10	
Full	RPG	.0550	.0386	.0778	.0460	–
Full	MHL (ours)	.0656	.0471	.0885	.0544	**+18.3**
Truncated	RPG	.6355	.4823	.7454	.5176	–
Truncated	MHL (ours)	.8460	.7208	.9199	.7451	**+44.0**

Table 9: Comparison of RPG and MHL on the *Toys* dataset. Both RPG and MHL use a 16-bit codebook, with MHL utilizing a mask ratio of 0.10. "Full" denotes evaluation on complete sequences, while "Truncated" denotes evaluation on the prefixes of long sequences. $\Delta\%$ denotes the relative improvement of N@10.

Test Set	Toys and Games		Δ%
	RPG N@10	MHL N@10	
Full	.0460	.0544	**+18.3**
≤ 10	.0475	.0548	**+15.4**
(10,20]	.0377	.0515	**+36.6**
(20,30]	.0332	.0426	**+28.3**
(30,40]	.0621	.0680	**+9.5**
(40,50]	.0653	.0912	**+39.7**
> 50	.0375	.0577	**+53.9**

Table 10: Length-stratified N@10 performance of RPG and MHL on the *Toys and Games* test set. Both RPG and MHL use a 16-bit codebook, and MHL employs a mask ratio of 0.10. $\Delta\%$ denotes the relative improvement of MHL over RPG.

Datasets	#Users	#Items	#Interactions	Avg. t
Beauty	22,363	12,101	176,139	8.87
Toys and Games	19,412	11,924	148,185	8.63
Sports and Outdoors	18,357	35,598	260,739	8.32

Table 11: Statistics of the Amazon Reviews 2014 datasets. "Avg. t" denotes the average number of interactions per input sequence.