Bridge the Unseen Gap: Enhancing Non-overlapping Cross-domain CTR Prediction via Profile Retrieval

Jingyang Bin[*,†]
Shenzhen Technology University
Shenzhen, China

Xing Tang[*]
Shenzhen Technology University
Shenzhen, China

Wei Zeng
Ant International, Ant Group
Chengdu, China

Jianan Su
Ant International, Ant Group
Chengdu, China

Kaixin Shen
Ant International, Ant Group
Chengdu, China

Jingtong Wu
Ant International, Ant Group
Hangzhou, China

Chaohua Yang[†]
Shenzhen University
Shenzhen, China

KaiLiang Hao[†]
Shenzhen University
Shenzhen, China

Dugang Liu
Shenzhen University
Shenzhen, China

Xiuqiang He[‡]
Shenzhen Technology University
Shenzhen, China

Abstract

Click-through rate (CTR) prediction is a fundamental task in industrial recommender systems. Cross-domain CTR prediction, which leverages data from a source domain to improve performance in a target domain, has emerged as a key strategy. However, most existing methods rely on overlapping users or items across domains to enable knowledge transfer, which fails in prevalent real-world scenarios where domains are functionally or geographically isolated (e.g., cross-country services). In this paper, we introduce a novel paradigm shift, from implicit representation alignment to explicit retrieval-based instance transfer. We propose **LLM-PRIT**, a framework for Large Language Model-generated Profile Retrieval & Instance Transfer. Our framework operates in three cohesive stages. First, it utilizes an LLM as a universal semantic interpreter to generate domain-agnostic, transferable profiles for users and items, encapsulating open-world knowledge. Second, instead of directly using these textual profiles, it employs them as semantic anchors to retrieve the most relevant historical instances from the source domain. This step explicitly establishes cross-domain correlations while avoiding the modality gap. Finally, it transfers knowledge by efficiently fine-tuning the target CTR model on the retrieved instances, preserving the model's inherent feature-interaction capabilities. We conduct extensive experiments on a public and a real-world industrial dataset. Both online and offline results demonstrate the effectiveness of our LLM-PRIT, bridging the unseen gap with open-world semantic information.

[*]Both authors contributed equally to this research
[†]Work done when they were research interns at Ant International, Ant Group.
[‡]Corresponding authors

RecSys '26, Minneapolis, MN, USA

ACM ISBN 979-8-4007-2284-4/2026/09
https://doi.org/10.1145/3773078.3831834

CCS Concepts

• **Information systems → Recommender systems**.

Keywords

Click-through Rate Prediction, Non-overlapping Cross-domain , User Profile

ACM Reference Format:
Jingyang Bin, Xing Tang, Wei Zeng, Jianan Su, Kaixin Shen, Jingtong Wu, Chaohua Yang, KaiLiang Hao, Dugang Liu, and Xiuqiang He. 2026. Bridge the Unseen Gap: Enhancing Non-overlapping Cross-domain CTR Prediction via Profile Retrieval. In *20th ACM Conference on Recommender Systems (RecSys '26), September 27-October 02, 2026, Minneapolis, MN, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3773078.3831834

1 Introduction

Click-through rate (CTR) prediction is a fundamental task in industrial recommender systems. These models are designed to estimate the probability that a user will click on a candidate item, based on multi-faceted input information gathered from the relevant single domain [3, 28, 31]. With the growing complexity of industrial applications, cross-domain CTR prediction (CDCTR) has emerged as a prominent strategy to enhance target-domain performance by leveraging auxiliary data from a source domain [20, 29, 41]. CDCTR critically relies on overlapping users and items, which serve as bridges for knowledge transfer under the closed-world assumption [39]. For example, in an Amazon shopping scenario illustrated in Figure 1, a user may be interested in a specific topic in both the book and movie domains. By leveraging overlapping users, movie recommendations can draw on knowledge transferred from the book domain. However, this closed-world assumption breaks down in non-overlapping business scenarios with no shared users or items. As illustrated in Figure 1, in the coupon recommendation scenario, users from diverse countries are recommended local service coupons. Since each country's users and

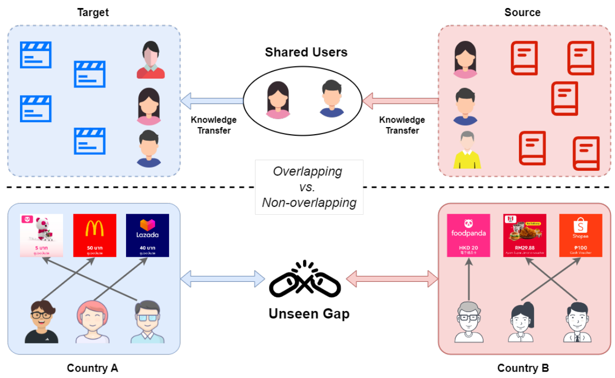

Figure 1: Illustration of overlapping versus non-overlapping cross-domain CTR prediction. (a) Overlapping scenario (top): Users are shared across domains, enabling direct knowledge transfer through common users. (b) Non-overlapping scenario (bottom): Domains are fully isolated (Country A vs. Country B in our coupon recommendation), with no shared users or items, creating an "unseen gap" that existing methods cannot bridge.

coupons are distinct, there are effectively no overlapping users or items across domains.

The efficacy of cross-domain CTR prediction stems from transfer learning via preference learning across the shared data [5]. Existing methods generally follow two prevalent paradigms. The first adopts a multi-task learning framework, jointly training on mixed samples from both domains to implicitly capture shared patterns across tasks [32, 43]. The second follows a pre-training-then-fine-tuning pipeline. A model is first pre-trained on source-domain data and subsequently fine-tuned on target-domain data [25, 40]. A fundamental limitation of both strategies is their strong dependence on overlapping users or items to align data between domains. This reliance restricts their applicability in non-overlapping cross-domain scenarios, which are common in practice but lack direct data correspondences. More concretely, multi-task learning assumes shared input features or labels to enable parameter sharing and implicit knowledge transfer. Likewise, pre-training followed by fine-tuning presumes that the feature distributions of the two domains are sufficiently aligned so that pre-trained representations remain meaningful and adaptable. When users and items are completely disjoint across domains, this core assumption of data alignment no longer holds. Without overlapping entities to serve as anchors, conventional methods lack a mechanism to transfer learned patterns, rendering them ineffective in such open-world settings.

Recent efforts have focused on overcoming the non-overlapping scenario by adopting prompt learning [30] and adversarial learning [26], and by utilizing domain-specific signals such as behavioral sequences, textual descriptions, or knowledge graphs to construct implicit bridges across isolated domains [14, 15, 16, 24, 36]. While these methods represent a significant step forward by relaxing the requirement for identical users or items, they nevertheless face three persistent and fundamental challenges when applied to non-overlapping CDCTR prediction. First, the alignment remains confined to a closed-world assumption: the bridging information, whether IDs, text, or knowledge-graph entities, is inherently derived from and semantically anchored within each isolated domain. This lack of a truly universal, domain-agnostic semantic anchor limits robust generalization. Second, the alignment process is largely implicit and unsupervised, making it susceptible to learning spurious correlations and suffering from negative transfer, especially when domains are genuinely disjoint. Third, these methods primarily focus on representation alignment, while modern CTR models derive their power from modeling high-order feature interactions [13, 35]. This objective misalignment means that even a well-aligned representation may not translate into effective predictive performance within an off-the-shelf CTR architecture.

In this paper, we propose a novel framework, **L**arge **L**anguage **M**odel-generated **P**rofile **R**etrieval & **I**nstance **T**ransfer (**LLM-PRIT**), to tackle the challenge of non-overlapping cross-domain CTR (CDCTR) prediction. The framework operates in three coherent stages: (1) generating domain-agnostic semantic profiles via an LLM, (2) retrieving relevant source-domain instances based on profile similarity, and (3) efficiently fine-tuning the target-domain CTR model with the retrieved instances. To infuse open-world knowledge into isolated domains, we first employ a large language model (LLM) to generate descriptive profiles for users and items based on their within-domain behavioral histories [37]. The LLM's inherent reasoning capability and factual knowledge allow it to act as a universal semantic interpreter [12], producing profiles that serve as domain-agnostic anchors for cross-domain alignment. However, directly integrating these raw textual profiles into the CTR model is suboptimal due to textual noise and the modality gap between natural language and the dense feature representations used in recommender systems [27]. To circumvent this issue, we adopt a retrieval-augmented strategy [33]. Instead of using the semantic profiles as direct model inputs, we leverage them to retrieve the most relevant historical instances (e.g., user-item interactions) from the source domain. This approach offers a dual advantage: it avoids the complexity of explicit modality alignment, and more importantly, the retrieved instances themselves serve as information-rich, structured carriers of transferable knowledge [23]. These instances act as high-quality bridges that encapsulate not only semantic similarity but also the contextual and behavioral patterns necessary for effective CTR prediction. Finally, to transfer knowledge without compromising the base model's capabilities, we fine-tune the target-domain CTR model using the retrieved source-domain instances. This is achieved via our designed multiple Low-Rank Adaptation (LoRA) [17], which updates a small set of low-rank matrices associated with the source domains. This allows the model to adaptively absorb knowledge from the retrieved instances while preserving its original feature-interaction modeling power, thereby enhancing CTR prediction performance in the target domain. We summarize the contributions of this work as follows:

- We identify the limitations of existing methods in non-overlapping CDCTR prediction and introduce a paradigm shift from implicit representation alignment to explicit semantic information retrieval-based instance transfer.

- We design and implement LLM-PRIT, a framework that realizes this paradigm through three cohesive stages: generating transferable semantic profiles via LLMs, retrieving relevant source-domain instances through profile-based matching, and effectively adapting the target CTR model via fine-tuning.
- We conduct extensive experiments on public and real-world industrial datasets. Furthermore, we also deploy our LLM-PRIT on our cross-region coupon recommendation scenario. Both offline and online results demonstrate that LLM-PRIT significantly enhances the non-overlapping CDCTR prediction.

2 Related Work

This section provides a concise review of related work, including cross-domain CTR prediction and the adoption of large language models for CTR prediction.

2.1 Cross-domain CTR prediction

Cross-domain recommendation (CDR) aims to improve performance in a target domain by leveraging knowledge from a related source domain [5, 45]. Within this field, a line of work closely aligned with ours focuses on enhancing CTR prediction across domains. Methods such as MiNet [29] use mixed-interest representations, while RAL-CDNet [33] and Gist [38] incorporate lifelong user behavior and contextual behavior patterns, respectively. DASL [20] employs a dual-embedding and dual-attention mechanism for information fusion, and CDANet [6] designs sequential translation and augmentation networks to enable explicit knowledge sharing. For time-evolving domains, CTLNet [25] proposes a continual transfer learning approach. Particularly relevant to our setting are methods developed for non-overlapping scenarios, where no shared users or items exist across domains. Representative works include KG-Bridge [36], MCRPL [24], PLCR [16], and TCPLP [15]. Although these are primarily designed for sequential recommendation, they commonly rely on prompt learning with various inputs (e.g., knowledge graphs, behavioral sequences) to align domains. Our work distinguishes itself by specifically targeting non-overlapping cross-domain CTR prediction, moving beyond prompt-based alignment to a retrieval-augmented transfer paradigm grounded in LLM-generated semantic profiles.

2.2 LLM-enhanced CTR prediction

With the increasing integration of large language models in recommendation systems [22], recent studies have explored leveraging LLMs to enhance CTR prediction [8, 10, 21, 37]. The primary advantage of LLMs lies in their rich semantic knowledge, which can provide a deeper understanding of user preferences and item characteristics. For instance, KAR [37] generates reasoning-enhanced knowledge and factual item descriptions to enrich feature representations, while UniCTR [10] extracts hierarchical semantic features to capture domain commonalities. However, directly deploying LLMs in industrial recommender systems faces challenges such as high inference latency and complex system integration. To mitigate these issues, another line of work focuses on using LLMs to generate user or item profiles as auxiliary semantic signals. Representative methods include UserIP [27], which employs expectation–maximization to infer latent user profiles, and LettinGO [34], which aligns profile generation with downstream preference optimization through direct feedback from recommendation tasks. Our work takes a distinct and complementary direction. Rather than using LLM-generated profiles as direct input features or for representation alignment, we employ them as semantic anchors in a retrieval stage to identify transferable instances across non-overlapping domains.

3 Methodology

In this section, we first formally formulate the non-overlapping cross-domain CTR (CDCTR) problem. Subsequently, we provide a detailed elaboration of our proposed LLM-PRIT framework, whose overall architecture is illustrated in Figure 2. The framework operates through three coherent stages. First, we design specialized prompts at the user and item levels to elicit rich, transferable semantic profiles from a large language model (LLM). We then encode these textual profiles into dense vector representations and perform instance retrieval from the source domain based on profile similarity. Finally, we adapt the target-domain CTR model via parameter-efficient fine-tuning using Low-Rank Adaptation (LoRA), leveraging the retrieved instances as augmented training data.

3.1 Problem Formulation

Given a dataset from the target domain $\mathcal{D}^t = \{(\mathbf{X}_i^t, y_i^t)\}_{i=1}^{N_t}$ and a dataset from the source domain $\mathcal{D}^s = \{(\mathbf{X}_j^s, y_j^s)\}_{j=1}^{N_s}$. Let $\mathcal{U}^t$ and $\mathcal{V}^t$ denote the sets of users and items in the target domain, and $\mathcal{U}^s$ and $\mathcal{I}^s$ denote those in the source domain. Each instance in the target domain is represented as $(\mathbf{X}_i^t, y_i^t)$, where $\mathbf{X}_i^t \in \mathcal{X}^t$ denotes the input feature vector. This vector is composed of the concatenation of three feature groups: the user features $\mathbf{x}_u^t$ for $u \in \mathcal{U}^t$, the item features $\mathbf{x}_v^t$ for $v \in \mathcal{V}^t$, and the contextual features. The label $y_i^t \in \{0, 1\}$ is a binary indicator of whether user u clicked item v ($y = 1$) or not ($y = 0$). The source domain data $\mathbf{X}_j^s \in \mathcal{X}^s$ is structured analogously, with $y_j^s \in \{0, 1\}$.

The general goal of cross-domain CTR prediction is to train a model that leverages instances from both the source and target domains to accurately predict click probability in the target domain. Formally, the objective can be expressed as learning a prediction function:

$$\hat{y}_{uv}^t = f(\mathbf{X}_{uv}^t; W \mid \mathcal{D}^t, \mathcal{D}^s), \tag{1}$$

where $u \in \mathcal{U}^t$ and $v \in \mathcal{V}^t$ denote a target user and item respectively, $\mathbf{X}_{uv}^t$ denotes the feature vector for the user-item pair (u, v), W denotes the model parameters, and the model gives the prediction $\hat{y}_{uv}^t$ by leveraging knowledge from both domains.

Conventional cross-domain CTR typically assumes the existence of overlapping entities across domains, facilitating knowledge transfer through shared representations or parameters. In contrast, we focus on the non-overlapping scenario where:

$$\mathcal{U}^t \cap \mathcal{U}^s = \emptyset \quad \text{and} \quad \mathcal{V}^t \cap \mathcal{V}^s = \emptyset. \tag{2}$$

Under this stricter condition, conventional methods that rely on explicit entity alignment are inapplicable. Our objective remains

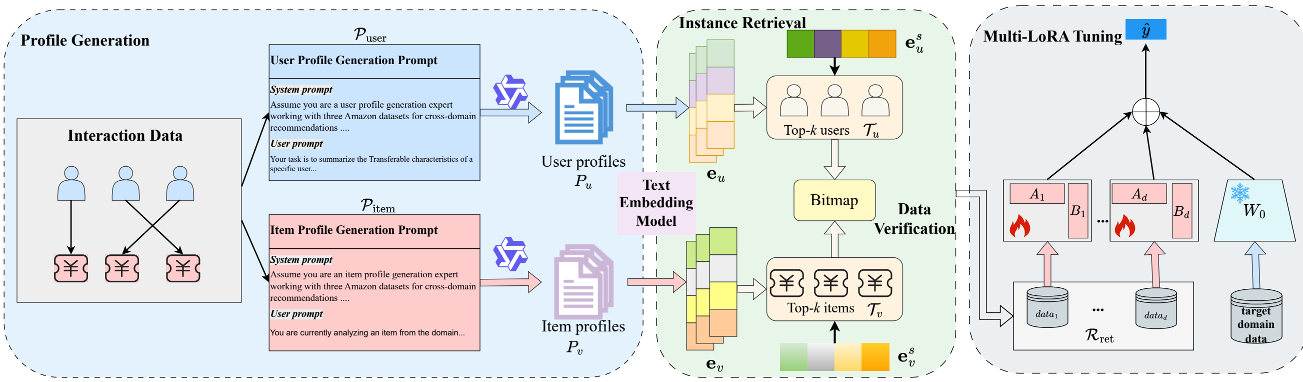

Figure 2: Overview of our LLM-PRIT framework. (1) Profile Generation: The LLM generates item (P_v) and user (P_u) profiles via designed prompts. (2) Instance Retrieval: Profiles are encoded (e_v, e_u) to retrieve top-k source users/items ($\mathcal{T}_u$, $\mathcal{T}_v$), followed by bitmap-based interaction verification. (3) Multi-LoRA Tuning: Verified instances train domain-specific low-rank adapters (A_d, B_d), aggregated during inference for prediction.

to learn an effective prediction function $f(\cdot)$ that transfers knowledge from $\mathcal{D}^s$ to improve target-domain performance, despite the absence of shared users or items.

3.2 Dual Semantic Profile Generation

Existing methods for cross-domain retrieval predominantly use pre-trained language models (PLMs) to encode textual information directly from item titles or descriptions, as demonstrated in previous work [15, 33]. While effective at capturing surface-level semantics, these approaches are inherently limited by their reliance on domain-specific vocabulary. The generated embeddings are tightly bound to the lexical space of the training data, lacking the abstraction necessary to bridge truly non-overlapping domains. This limitation stems from their reliance on the source data's lexicon, which fails to provide a universal semantic anchor. To overcome this, we propose an LLM-generated profile method that leverages the open-world knowledge and reasoning capabilities of Large Language Models (LLMs). Our core insight is to shift from encoding raw text to eliciting structured and transferable characterizations of users and items. This is achieved through two carefully designed prompt templates that guide the LLM to act as a domain-agnostic semantic interpreter.

Formally, let a user's interaction history be denoted as $h = \{i_1, i_2, \cdots, i_{|h|}\}$, where each i_j is an interacted item. The corresponding metadata sequence is $M(h) = [m_{i_1}, \cdots, m_{i_{|h|}}]$, containing titles and descriptions. We construct a composite prompt $\mathcal{P}_{\text{user}}(h, M(h))$ that includes: (1) a *System Prompt* defining the expert role, and (2) a *User Prompt* containing the interaction history, formatting rules, and in-context examples. A critical component of the user prompt is the structured requirement set as illustrated in Figure 3, which explicitly instructs the LLM how to output traits. Similarly, for an item v with metadata m_v, i.e., title, description, we employ a separate prompt $\mathcal{P}_{\text{item}}(m_v)$. This prompt guides the LLM to generate a single, succinct sentence that describes the item's core functionality or essence in a domain-agnostic manner, following the requirements illustrated in Figure 3.

User Profile Generator

...

Requirement:

- Please conclude it not beyond 50 words.
- Do not only evaluate one specific commodity, but illustrate the interests overall.
- Instead of answering a fluent paragraph, return a concise list of 5–7 abstract, transferable user traits.
- Each trait should be 2–5 words, capturing the underlying value, cognitive pattern, or emotional driver.
- Return only the list, one per line.

...

Item Profile Generator

...

Requirement:

- Write one sentence (≤50 words) that summarizes the item's core content or genre.
- Describe the nature of the item rather than making subjective claims about its quality.
- Do not mention the specific item title, author, director, or brand names in your description.
- Strictly base your description on the information inferred from the title alone.

...

Figure 3: The requirements for generating user/item profiles.

This dual-prompt design captures generalizable user preferences and item semantics that go beyond the surface-level information in raw metadata, while ensuring greater specificity and consistency than a single-prompt alternative. Specifically, the LLM's reasoning capabilities yield coherent psychological profiles for users, while its open-world knowledge provides precise, domain-agnostic descriptions of items. As a result, the generated profiles reside in a shared conceptual semantic space, enabling meaningful similarity computation across otherwise isolated domains. Unlike previous approaches that fine-tune LLMs for recommendation [11, 34], we utilize the pre-trained LLM without task-specific tuning. Notice that this aligns with our goal to leverage the model's general semantic understanding rather than narrow task adaptation.

3.3 Cross-domain Instance Retrieval

In this stage, we leverage the generated semantic profiles to retrieve relevant cross-domain instances from the source domain. Specifically, for a given target user-item pair (u^t, v^t), we first encode their textual profiles $\mathcal{P}_{u^t}$ and $\mathcal{P}_{v^t}$ into dense vectors $\mathbf{e}_u^t$ and $\mathbf{e}_v^t$ using an embedding model [42]. We independently retrieve the top-K most similar source-domain entities for both the user and the item based on cosine similarity in the profile embedding space. For the user side, we calculate:

$$\cos(\mathbf{e}_u^t, \mathbf{e}_{u_i}^s) = \frac{\mathbf{e}_u^t \cdot \mathbf{e}_{u_i}^s}{\|\mathbf{e}_u^t\| \|\mathbf{e}_{u_i}^s\|}, \tag{3}$$

where $\mathbf{e}_{u_i}^s$ is the embedding of the i-th source user profile. The set of retrieved candidate source users is:

$$\mathcal{T}_u = \mathrm{Top}_K \left(\left\{ \cos(\mathbf{e}_u^t, \mathbf{e}_{u_i}^s) \right\}_{i=1}^{|\mathcal{U}^s|} \right). \tag{4}$$

Analogously, we retrieve a set of candidate source items $\mathcal{T}_v$ using the target item profile embedding $\mathbf{e}_v^t$. A key distinction from prior retrieval-based methods is our explicit interaction verification step. Since we independently retrieve users and items, we must verify that each candidate pair (u^s, v^s) has an actual interaction record in the source domain, ensuring that transferred knowledge is grounded in real user behavior rather than arbitrary user-item combinations. Formally, the final set of verified transferable instances is:

$$\mathcal{R}_{\text{ret}} = \{(u^s, v^s) \mid u^s \in \mathcal{T}_u,\ v^s \in \mathcal{T}_v,\ (u^s, v^s) \in \mathcal{D}^s\}. \tag{5}$$

For each verified pair (u^s, v^s), we fetch the complete interaction record from the source domain data, including the feature vector $\mathbf{X}^s$ and label y^s. This ensures that transferred instances are not only semantically relevant but also grounded in actual user behavior.

Technically, we implement several optimizations to improve the efficiency. First, we use Faiss [9] to build two separate indices for source user and source item profile embeddings, respectively. This allows sub-linear search complexity for similarity retrieval. Second, we utilize a compressed Bitmap [2] index to store the source-domain interactions, enabling $O(1)$ membership queries during the interaction verification step.

3.4 Target-domain Model Tuning

The final stage of the LLM-PRIT framework is the adaptive fine-tuning of the target-domain model. In our non-overlapping scenario, the absence of shared users or items means each domain occupies a mutually exclusive feature space. Given that the feature spaces remain strictly disjoint, we avoid explicit alignment across them. Instead, we transfer knowledge directly at the instance level through $\mathcal{R}_{\text{ret}}$, as the retrieved samples encode domain-agnostic interaction patterns that provide a supervisory signal without requiring the latent spaces to coincide. Consequently, we maintain separate, domain-specific embedding layers and keep them frozen during the tuning phase. This design anchors each representation within its native semantic space, preserving the fidelity of pre-trained knowledge while avoiding spurious feature entanglement across domains. With the embedding layers and backbone parameters fixed, we inject lightweight, domain-specific Low-Rank Adaptation (LoRA) modules into the model and update only these adapters. These modules capture the behavioral patterns in the retrieved examples without disrupting the pre-trained representations.

3.4.1 *Multi-LoRA Fine-tuning Process.*

To accommodate the distinct distributional characteristics of different source domains, we adopt a Multi-LoRA architecture.

Let the base CTR model pre-trained on the target domain be parameterized by weights W_0. We employ a Multi-LoRA architecture to bridge the distribution gap between domains without modifying the pre-trained backbone parameters. For each source domain $d \in \{1, \ldots, D\}$, we allocate a separate low-rank adapter (B_d, A_d), where $B_d \in \mathbb{R}^{m \times r}$, $A_d \in \mathbb{R}^{r \times n}$ and $r \ll \min(m, n)$.

During tuning, a retrieved instance $(\mathbf{X}^s, y^s)$ originating from source domain d is first embedded using the corresponding pre-trained domain-specific embedding table $\mathbf{E}_{\mathrm{s}}^{(d)}$. The prediction is then computed by the shared backbone augmented with the domain adapter:

$$\hat{y}^s = f\Big(\mathbf{E}_{\mathrm{s}}^{(d)}(\mathbf{X}^s);\ W_0 + \frac{\alpha}{r} B_d A_d\Big), \tag{6}$$

where α is a scaling hyper-parameter. The adapter is optimized by minimizing the standard binary cross-entropy loss on the retrieved instances:

$$\mathcal{L}_d = -\frac{1}{|\mathcal{R}_{\text{ret}}^{(d)}|} \sum_{(\mathbf{X}^s, y^s) \in \mathcal{R}_{\text{ret}}^{(d)}} \left[y^s \log \hat{y}^s + (1 - y^s) \log(1 - \hat{y}^s) \right]. \tag{7}$$

By updating only the low-rank matrices B_d and A_d, the model learns to capture the transferable interaction patterns from each source domain, without distorting the pre-trained target-domain representations.

3.4.2 *Knowledge Integration and Inference.*

Once tuning is complete, each domain-specific adapter encapsulates transferable structural knowledge from its respective source domain. During inference on a target-domain instance $\mathbf{X}^t$, we use the frozen target-domain embedding table $\mathbf{E}_{\mathrm{t}}$ and perform an ensemble over all adapters:

$$\hat{y}^t = \frac{1}{D} \sum_{d=1}^{D} f\Big(\mathbf{E}_{\mathrm{t}}(\mathbf{X}^t);\ W_0 + \frac{\alpha}{r} B_d A_d\Big). \tag{8}$$

This strategy effectively translates the behavioral patterns learned from source-domain interactions to the target domain, enabled by the semantic proximity established in the profile retrieval stage. Notably, no LLM inference, profile retrieval, or external database access is required during serving, ensuring minimal latency and full compatibility with industrial deployment constraints.

4 Experiments

In this section, we conduct comprehensive experiments to validate the effectiveness and efficiency of LLM-PRIT. Specifically, we aim to answer the following research questions:

- **RQ1:** How effectively can LLM-PRIT bridge non-overlapping domains for CTR prediction, compared to baselines?
- **RQ2:** What is the contribution of each key component in LLM-PRIT to the overall performance?
- **RQ3:** What characterizes the retrieved instances, and how do the semantic profiles influence retrieval quality?
- **RQ4:** How does LLM-PRIT perform under real-world, large-scale industrial deployment in terms of both accuracy and efficiency?

4.1 Experiment Setup

4.1.1 Datasets. Our experiments are conducted on two datasets: the XMarket public benchmark and a private industrial dataset. Both consist of multiple domains with disjoint user and item sets.

XMarket Dataset[1]**.** XMarket is a large-scale real-world product recommendation dataset constructed by global Amazon marketplaces, covering 18 markets and 11 languages. In this study, we selected three representative cross-domain settings: Canada Books(ca-Book), UK Movies & TV(uk-Movie), and US Electronics(us-Elec), for non-overlapping CDCTR prediction tasks. During preprocessing, we removed overlapping entities to ensure complete domain non-overlap and retained only users/items with $\geq$ 5 interactions for data quality. We converted ratings of 4 or 5 to positive samples and other ratings to negative samples. To ensure experimental fairness, we randomly split all samples into training, validation, and test sets with a ratio of 8:1:1.

Industrial Dataset. This dataset was collected from the marketing platform of Ant International Alipay, containing coupon recommendation data from multiple leading e-wallets in overseas markets, recording user click behavior and actual usage conversion after coupon exposure. The users and items across different market domains are completely non-overlapping. We formulate the prediction task as modeling coupon conversion behavior, where a successful conversion is labeled 1 and an unsuccessful one 0. The dataset comprises 1.7 million users and 16 million interaction records collected over 33 consecutive days. We allocate the first 30 days to training and the last 3 days to testing.

4.1.2 Metrics. For offline experiments, we introduce three metrics commonly used in previous work [13, 44]: AUC (Area Under the Curve), logarithmic loss (logloss), and GAUC (Group AUC). For online experiments, we evaluate performance using four business metrics, computed at both the sample level (PV: Page View) and the user level (UV: User View): CTR-PV = $N_{\text{click}}/N_{\text{imp}}$, CTR-UV = $U_{\text{click}}/U_{\text{imp}}$, CVR-PV = $N_{\text{conv}}/N_{\text{imp}}$, and CVR-UV = $U_{\text{conv}}/U_{\text{imp}}$, where N and U denote sample and user counts, with subscripts indicating impressions (imp), clicks (click), and conversions (conv).

4.1.3 Baselines Methods. To comprehensively evaluate LLM-PRIT, we compare it with a wide range of baselines from three categories: single-domain CTR models, cross-domain methods, and non-overlapping cross-domain methods. The selected baselines are summarized below.

- **DNN** [7]: A deep neural network with embedding, fully-connected, and output layers.
- **DeepFM** [13]: Combines factorization machines and deep networks to capture both low- and high-order feature interactions.
- **DCN_V2** [35]: Uses a mixture of low-rank architecture to enhance feature crossing.
- **CoNet** [18]: A classical cross-domain model that transfers knowledge via cross-connections between domain-specific networks.
- **STAR** [32]: A star-topology adaptive recommender that trains a unified model to serve all domains with domain-specific parameters.
- **MLoRA** [40]: Extends LoRA to multi-domain CTR by introducing a dedicated low-rank adapter for each domain.
- **SSIM** [23]: Uses an adaptive selection network to identify informative and shareable instances across domains.
- **TCPLP** [15]: Employs text-based item representations to enable semantic transfer in non-overlapping domains.

4.1.4 Parameter Settings. We implement all models using PyTorch with the Adam optimizer [19]. For the XMarket and Industrial datasets, the batch sizes are set to 128 and 8,196, respectively. The embedding dimension is consistently set to 32, and we employ a three-layer MLP with hidden dimensions of 256, 128, and 64. In the tuning stage, the learning rate is searched within {1e-3, 5e-4, 1e-4, 5e-5}, the L2 regularization coefficient is searched within {1e-4, 1e-5, 5e-6, 1e-6, 1e-7}, and the LoRA rank and alpha are searched within {4, 8, 16, 32} and {4, 8, 16, 32, 64}, respectively. The semantic profiles for users and items are generated by the Qwen-8B [1], and subsequently converted into dense vector representations using the BGE-M3 [4] embedding model for retrieval. All experiments are conducted on NVIDIA H100 GPUs. For each target domain, all remaining domains serve as source domains for knowledge transfer. All baseline models are implemented based on their official open-source code, with TCPLP adapted from its original sequential recommendation task to the CTR prediction task.

4.2 RQ1: Overall Performance

In this section, we compare our proposed LLM-PRIT with three categories of baseline methods: single-domain, cross-domain, and non-overlapping methods. Experiments are conducted on two datasets, and the overall performance comparison is demonstrated in Table 1.

The experimental results lead to several key observations.

- LLM-PRIT consistently outperforms all baselines across nearly all domains and metrics, which demonstrates its superior performance in non-overlapping cross-domain scenarios.
- TCPLP exhibits limited gains in performance, even compared to single-domain models, and is consistently outperformed by LLM-PRIT. This highlights that methods designed for sequential recommendation are not inherently suited for CTR prediction. Its reliance on domain-specific textual information alone further hinders effective cross-domain knowledge transfer.
- Cross-domain models generally underperform single-domain models in non-overlapping scenarios. This suggests that methods relying on implicit representation transfer struggle to effectively bridge domains when there are no overlapping entities.

[1]https://xmrec.github.io/

Table 1: Overall performance comparison across datasets. Bold numbers indicate the best performance, and underlined numbers indicate the second best. Note that $*$ indicates a significance level of $p \leq 0.05$ based on a two-sample t-test.

Dataset	Domain	Metric	Single-Domain			Cross-Domain				Non-Overlapping	
			DNN	DeepFM	DCN_V2	CoNet	STAR	MLoRA	SSIM	TCPLP	LLM-PRIT
XMarket	ca-Book	AUC	0.7798	0.7818	0.7806	0.7687	0.7753	0.7660	0.7788	0.7806	**0.7849***
		GAUC	0.6012	0.6025	0.6021	0.5833	0.5851	0.6013	0.6010	0.6013	**0.6034***
		LogLoss	0.3553	0.3522	0.3536	0.3616	0.3596	0.3841	0.3726	0.3539	**0.3510***
	uk-Movie	AUC	0.8313	0.8317	0.8318	0.8259	0.8288	0.8315	0.8311	0.8306	**0.8339***
		GAUC	0.6729	0.6722	0.6737	0.6627	0.6672	0.6723	0.6724	0.6727	**0.6758***
		LogLoss	0.3674	0.3669	0.3672	0.3825	0.3701	0.3866	0.3724	0.3672	**0.3657***
	us-Elec	AUC	0.7534	0.7545	0.7516	0.7399	0.7437	0.7526	0.7505	0.7494	**0.7564***
		GAUC	0.6593	0.6611	0.6601	0.6435	0.6458	0.6600	0.6589	0.6592	**0.6658***
		LogLoss	0.3841	0.3830	0.3861	0.3954	0.3900	0.3866	0.3908	0.3888	**0.3811***
Industrial	Domain1	AUC	0.9207	0.9318	0.9570	0.9463	0.9582	0.9522	0.9602	0.9546	**0.9619***
		GAUC	0.8286	0.8485	0.8754	0.8345	0.8823	0.8748	0.8742	0.8734	**0.8825***
		LogLoss	0.3138	0.3048	0.2998	0.3069	0.2961	0.3075	0.3022	0.3026	**0.2912***
	Domain2	AUC	0.9430	0.9503	0.9755	0.9656	0.9792	0.9760	0.9700	0.9683	**0.9844***
		GAUC	0.8406	0.8412	0.8745	0.8646	0.8736	0.8638	0.8709	0.8610	**0.8816***
		LogLoss	0.3131	0.3024	0.2891	0.2934	0.2913	0.2967	0.2996	0.2990	**0.2885***
	Domain3	AUC	0.8845	0.8859	0.9072	0.8914	0.9126	0.9022	0.9089	0.8861	**0.9143***
		GAUC	0.8101	0.8120	0.8316	0.8130	0.8328	0.8235	0.8349	0.8122	**0.8389***
		LogLoss	0.3734	0.3698	0.3688	0.3715	0.3664	0.3703	0.3677	0.3709	**0.3606***
	Domain4	AUC	0.9685	0.9716	0.9813	0.9743	0.9794	0.9717	0.9777	0.9737	**0.9850***
		GAUC	0.9008	0.9203	0.9368	0.9185	0.9370	0.9361	0.9358	0.9248	**0.9405***
		LogLoss	0.3164	0.3136	0.2982	0.3098	0.2990	0.3073	0.3050	0.3110	**0.2954***
	Domain5	AUC	0.9666	0.9684	0.9831	0.9750	0.9836	0.9820	0.9790	0.9767	**0.9844***
		GAUC	0.9201	0.9271	0.9438	0.9322	0.9463	0.9416	0.9445	0.9429	**0.9526***
		LogLoss	0.3455	0.3327	0.3226	0.3294	0.3207	0.3255	0.3247	0.3260	**0.3178***

- Among cross-domain methods, STAR achieves the best performance, followed by SSIM and then MLoRA, while CoNet is the least effective. STAR benefits from its star-topology design, offering better adaptability to distribution shifts. SSIM and MLoRA partially mitigate negative transfer through instance selection and parameter-efficient tuning, yet remain limited by implicit representation alignment. CoNet's dependence on explicit inter-domain connections renders it ineffective in the absence of overlapping entities. Overall, while these models incorporate useful techniques, they do not fundamentally resolve the semantic gap in non-overlapping settings.
- Single-domain models remain strong baselines. This highlights the importance of feature interaction modeling within a single domain, though it is still insufficient compared to LLM-PRIT's cross-domain semantic transfer.

These results collectively indicate that successful knowledge transfer under non-overlapping constraints necessitates an explicit open-world semantic information paradigm. LLM-PRIT addresses this by leveraging LLM-generated user and item profiles, which encapsulate open-world knowledge, to retrieve relevant cross-domain instances, thereby enabling effective transfer while preserving the capability of off-the-shelf CTR models.

4.3 RQ2: Ablation Study

To investigate the contribution of each key component in LLM-PRIT, we conduct comprehensive ablation studies on the XMarket dataset. We compare: (1) **LLM-PRIT**: complete model; (2) **w/o Domain**: single shared LoRA instead of domain-specific modules; (3) **w/o Profile**: direct text similarity instead of semantic profiles; (4) **w/o Retrieval**: random instance transfer instead of retrieval-based selection; (5) **w/o Selection**: transferring all source instances without filtering. The results are presented in Table 2.

Table 2: Performance of LLM-PRIT variants on the XMarket dataset. Removing any component degrades performance, confirming the necessity of each module in our framework.

Datasets	ca-Book		uk-Movie		us-Elec	
	AUC	GAUC	AUC	GAUC	AUC	GAUC
LLM-PRIT	**0.7849**	**0.6034**	**0.8339**	**0.6758**	**0.7564**	**0.6658**
w/o Domain	0.7821	0.6022	0.8323	0.6740	0.7544	0.6635
w/o Profile	0.7767	0.5985	0.8278	0.6694	0.7501	0.6562
w/o Retrieval	0.7732	0.5969	0.8235	0.6675	0.7498	0.6568
w/o Selection	0.7506	0.5812	0.7871	0.6302	0.7405	0.6493

We can observe that performance degrades consistently across all ablated settings, confirming that each component contributes

User Profile Transfer

Target User Profile (ca-Book): Classic literature enthusiast; Appreciates dramatic narratives; Values emotional depth; Enjoys character-driven stories; Prefers timeless storytelling.
Interaction History: *Othello* (4.6), *A Woman of Substance* (4.7), *Dear and Glorious Physician* (4.6), ...
Top-1 Retrieved User Profile (uk-Movie, Similarity: 0.82): Period drama aficionado; Drawn to compelling performances; Values authentic storytelling; Enjoys historical narratives; Appreciates cinematic adaptations.
Interaction History: *Taxi Driver [DVD]* (4.5), *My Left Foot [DVD]* (4.6), *Chaplin [DVD]* (4.6), ...

Item Profile Transfer

Target Item Profile (ca-Book): A classic war novel depicting human courage and sacrifice during wartime, exploring themes of duty, honor, and the psychological impact of combat through richly developed characters.
Item Title: *For Whom the Bell Tolls.*
Top-1 Retrieved Item Profile (uk-Movie, Similarity: 0.75): A powerful war film showcasing heroism and brotherhood in battle, featuring intense combat sequences and emotional character arcs that examine the cost of warfare.
Item Title: *Sands of Iwo Jima [DVD].*

Figure 4: Case study illustrating LLM-PRIT's capability to establish semantic bridges across non-overlapping domains. The alignment of user and item profiles with high similarity scores confirms the effectiveness of profile-based retrieval for knowledge transfer.

positively to the final performance. The removal of semantic profiles (**w/o Profile**) results in a notable decline, underscoring the importance of LLM-generated profiles for capturing transferable user preferences and item semantics beyond closed-world information. The most severe performance degradation occurs when removing instance selection (**w/o Selection**), followed closely by removing the retrieval mechanism (**w/o Retrieval**). This highlights the need for semantic filtering to prevent negative transfer from irrelevant cross-domain data. The domain-specific LoRA design (**w/o Domain**) also shows a clear decline, indicating that preserving domain-aware adaptation helps in fine-tuning the retrieved knowledge. These results further validate our core design of explicit, retrieval-based instance transfer over implicit representation alignment.

4.4 RQ3: Case Study

To illustrate the effectiveness of LLM-PRIT in establishing semantic correlations across non-overlapping domains, we present a case of cross-domain knowledge transfer from the *uk-Movie* to the *ca-Book* domain. As a case in Figure 4, the target user in *ca-Book* is profiled by the LLM as a "classic literature enthusiast who appreciates dramatic narratives and emotional depth". Based on this profile, our method retrieves a highly similar user from *uk-Movie*, described as a "period drama aficionado drawn to compelling performances and authentic storytelling". This demonstrates that LLM-generated user profiles successfully capture high-level, transferable user preferences beyond domain-specific behaviors. Similarly, for the target book "For Whom the Bell Tolls," its LLM-generated item profile emphasizes "themes of courage, sacrifice, and the psychological impact of war." The top retrieved movie "Sands of Iwo Jima [DVD]" is correspondingly described as "a war film showcasing heroism and brotherhood in battle". This alignment at the conceptual and thematic level validates that LLM-PRIT can bridge domain gaps by focusing on semantic essence rather than surface-level features. These examples underscore a key advantage of LLM-PRIT: it enables conceptually grounded cross-domain transfer, moving beyond implicit embedding alignment to explicit semantic matching.

4.5 RQ4: Online Deployments

To validate the real-world effectiveness of LLM-PRIT, we deployed it on the marketing platform within Ant International Alipay, which serves tens of millions of users monthly across multiple leading e-wallets in different markets. The platform focuses on personalized coupon recommendations in information feeds, aiming to achieve precise user acquisition and efficient merchant customer conversion.

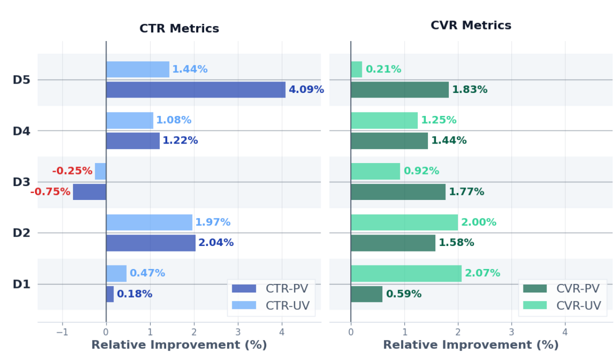

Figure 5: Online A/B test results across five e-wallet domains. LLM-PRIT achieves consistent improvements across four business metrics, with D5 showing the highest CTR-PV gain of 4.09% and D1 demonstrating the strongest CVR-UV improvement of 2.07%.

We conducted 14-day A/B tests across five e-wallet domains, with each target domain paired with the remaining four as source domains. The experiment was allocated 20% of total traffic, exposing the system to millions of real-time user interactions. As summarized in Figure 5, LLM-PRIT achieved consistent and statistically significant improvements across all key business metrics compared to the baseline, which is a production-optimized cross-domain CTR model. Notably, LLM-PRIT achieved average improvements of **1.44%** in CVR-PV and **1.29%** in CVR-UV compared to the baseline, further validating the practical value of our approach in large-scale industrial deployment.

Beyond prediction accuracy, we also evaluated the system's inference efficiency, which is a critical factor for industrial deployment. As shown in Table 3, LLM-PRIT introduces only a modest latency overhead of **+1.13 ms** on average and **+1.80 ms** at P95. The overhead is well within the acceptable range for real-time recommendation services. This efficiency is achieved because no LLM inference or cross-domain retrieval occurs during serving; only the tuned multi-LoRA weights are used, ensuring compatibility with existing low-latency infrastructure. We also analyzed the offline LLM profile generation cost on the industrial dataset (NVIDIA H100 GPUs): 1.7M user profiles take 56.7h (single H100) / 1.8h (32-GPU parallel), 2K item profiles take 1min, with total parallel initial build 2.7h. Weekly user profile updates cost 56.7 GPU-hours (8.1/day), within industrial budgets; all LLM operations are offline, leaving online latency unaffected.

Table 3: Latency comparison between the baseline model and LLM-PRIT in online deployment.

Model	Avg Latency (ms)	P95 Latency (ms)
Baseline	7.12	12.22
LLM-PRIT	8.25	14.02
Overhead	1.13	1.80

5 Conclusion

In this work, we address the challenge of non-overlapping cross-domain CTR prediction by proposing LLM-PRIT, a novel framework that shifts from implicit representation alignment to explicit instance transfer. Our approach leverages large language models to generate domain-agnostic semantic profiles, which serve as anchors for retrieving relevant cross-domain instances, followed by parameter-efficient fine-tuning via multi-LoRA modules. This design effectively circumvents the need for overlapping entities and enables robust knowledge transfer in non-overlapping settings. Extensive experiments conducted on both public and industrial datasets demonstrate that LLM-PRIT achieves state-of-the-art performance. Moreover, large-scale online A/B tests deployed on the coupon recommendation scenario within Ant International Alipay confirm its practical effectiveness and efficiency.

6 Acknowledgement

We thank the support of the Shenzhen Technology University School-level (No.20251061020002), and this work was supported by the Ant Group Research Intern Program.

References

[1] Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, et al. 2023. Qwen technical report. *arXiv preprint arXiv:2309.16609* (2023).

[2] Chee-Yong Chan and Yannis E Ioannidis. 1998. Bitmap index design and evaluation. In *Proceedings of the 1998 ACM SIGMOD international conference on Management of data*. 355–366.

[3] Olivier Chapelle, Eren Manavoglu, and Romer Rosales. 2014. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5, 4 (2014), 1–34.

[4] Jianlv Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, and Zheng Liu. 2024. BGE M3-Embedding: Multi-Lingual, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation. arXiv:2402.03216 [cs.CL]

[5] Shu Chen, Zitao Xu, Weike Pan, Qiang Yang, and Zhong Ming. 2024. A survey on cross-domain sequential recommendation. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*. Article 884, 10 pages.

[6] Xu Chen, Zida Cheng, Jiangchao Yao, Chen Ju, Weilin Huang, Jinsong Lan, Xiaoyi Zeng, and Shuai Xiao. 2024. Enhancing Cross-Domain Click-Through Rate Prediction via Explicit Feature Augmentation. In *Companion Proceedings of the ACM Web Conference 2024 (WWW '24)*. 423–432.

[7] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, Rohan Anil, Zakaria Haque, Lichan Hong, Vihan Jain, Xiaobing Liu, and Hemal Shah. 2016. Wide & Deep Learning for Recommender Systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*. 7–10.

[8] Yu Cui, Feng Liu, Jiawei Chen, Xingyu Lou, Changwang Zhang, Jun Wang, Yuegang Sun, Xiaohu Yang, and Can Wang. 2025. Field Matters: A lightweight LLM-enhanced Method for CTR Prediction. *arXiv preprint arXiv:2505.14057* (2025).

[9] Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2024. The Faiss library. (2024). arXiv:2401.08281 [cs.LG]

[10] Zichuan Fu, Xiangyang Li, Chuhan Wu, Yichao Wang, Kuicai Dong, Xiangyu Zhao, Mengchen Zhao, Huifeng Guo, and Ruiming Tang. 2025. A Unified Framework for Multi-Domain CTR Prediction via Large Language Models. *ACM Trans. Inf. Syst.* 43, 5, Article 117 (July 2025), 33 pages.

[11] Zhaolin Gao, Joyce Zhou, Yijia Dai, and Thorsten Joachims. 2025. LangPTune: Optimizing Language-based User Profiles for Recommendation. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*. 707–717.

[12] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Peiyi Wang, Qihao Zhu, Runxin Xu, Ruoyu Zhang, Shirong Ma, Xiao Bi, et al. 2025. DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature* 645, 8081 (2025), 633–638.

[13] Huifeng Guo, TANG Ruiming, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: A Factorization-Machine based Neural Network for CTR Prediction. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*.

[14] Lei Guo, Hao Liu, Lei Zhu, Weili Guan, and Zhiyong Cheng. 2023. DA-DAN: A Dual Adversarial Domain Adaption Network for Unsupervised Non-overlapping Cross-domain Recommendation. *ACM Trans. Inf. Syst.* 42, 2, Article 48 (Nov. 2023), 27 pages.

[15] Lei Guo, Chenlong Song, Feng Guo, Xiaohui Han, Xiaojun Chang, and Lei Zhu. 2025. Semantic-enhanced Co-attention Prompt Learning for Non-overlapping Cross-domain Recommendation. *ACM Trans. Inf. Syst.* 43, 5, Article 137 (Aug. 2025), 27 pages.

[16] Lei Guo, Chunxiao Wang, Xinhua Wang, Lei Zhu, and Hongzhi Yin. 2025. Automated Prompting for Non-Overlapping Cross-Domain Sequential Recommendation. *IEEE Transactions on Knowledge and Data Engineering* 37, 9 (2025), 4990–5003.

[17] Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *International Conference on Learning Representations*. https://openreview.net/forum?id=nZeVKeeFYf9

[18] Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. CoNet: Collaborative Cross Networks for Cross-Domain Recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 667–676.

[19] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*.

[20] Pan Li, Zhichao Jiang, Maofei Que, Yao Hu, and Alexander Tuzhilin. 2021. Dual Attentive Sequential Learning for Cross-Domain Click-Through Rate Prediction. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21)*. 3172–3180.

[21] Jianghao Lin, Bo Chen, Hangyu Wang, Yunjia Xi, Yanru Qu, Xinyi Dai, Kangning Zhang, Ruiming Tang, Yong Yu, and Weinan Zhang. 2024. ClickPrompt: CTR Models are Strong Prompt Generators for Adapting Language Models to CTR Prediction. In *Proceedings of the ACM Web Conference 2024*. 3319–3330.

[22] Jianghao Lin, Xinyi Dai, Yunjia Xi, Weiwen Liu, Bo Chen, Hao Zhang, Yong Liu, Chuhan Wu, Xiangyang Li, Chenxu Zhu, et al. 2025. How can recommender systems benefit from large language models: A survey. *ACM Transactions on Information Systems* 43, 2 (2025), 1–47.

[23] Dugang Liu, Chaohua Yang, Yuwen Fu, Xing Tang, Gongfu Li, Fuyuan Lyu, Xiuqiang He, and Zhong Ming. 2025. Scenario Shared Instance Modeling for Click-through Rate Prediction. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*. 2436–2447.

[24] Hao Liu, Lei Guo, Lei Zhu, Yongqiang Jiang, Min Gao, and Hongzhi Yin. 2024. MCRPL: A Pretrain, Prompt, and Fine-tune Paradigm for Non-overlapping

Many-to-one Cross-domain Recommendation. *ACM Trans. Inf. Syst.* 42, 4, Article 97 (Feb. 2024), 24 pages.

[25] Lixin Liu, Yanling Wang, Tianming Wang, Dong Guan, Jiawei Wu, Jingxu Chen, Rong Xiao, Wenxiang Zhu, and Fei Fang. 2023. Continual Transfer Learning for Cross-Domain Click-Through Rate Prediction at Taobao. In *Companion Proceedings of the ACM Web Conference 2023*. 346–350.

[26] Daniel Lowd and Christopher Meek. 2005. Adversarial learning. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. 641–647.

[27] Yusheng Lu, Zhaocheng Du, Xiangyang Li, Pengyue Jia, Yejing Wang, Weiwen Liu, Yichao Wang, Huifeng Guo, Ruiming Tang, Zhenhua Dong, Yongrui Duan, and Xiangyu Zhao. 2025. Prompt Tuning as User Inherent Profile Inference Machine. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*. 5898–5906.

[28] Fuyuan Lyu, Xing Tang, Dugang Liu, Haolun Wu, Chen Ma, Xiuqiang He, and Xue Liu. 2023. Feature representation learning for click-through rate prediction: A review and new perspectives. *arXiv preprint arXiv:2302.02241* (2023).

[29] Wentao Ouyang, Xiuwu Zhang, Lei Zhao, Jinmei Luo, Yu Zhang, Heng Zou, Zhaojie Liu, and Yanlong Du. 2020. MiNet: Mixed Interest Network for Cross-Domain Click-Through Rate Prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM '20)*. 2669–2676.

[30] Fabio Petroni, Tim Rocktäschel, Sebastian Riedel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander Miller. 2019. Language models as knowledge bases?. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*. 2463–2473.

[31] Matthew Richardson, Ewa Dominowska, and Robert Ragno. 2007. Predicting clicks: Estimating the click-through rate for new ads. In *Proceedings of the 16th International Conference on World Wide Web*. 521–530.

[32] Xiang-Rong Sheng, Liqin Zhao, Guorui Zhou, Xinyao Ding, Binding Dai, Qiang Luo, Siran Yang, Jingshan Lv, Chi Zhang, Hongbo Deng, and Xiaoqiang Zhu. 2021. One Model to Serve All: Star Topology Adaptive Recommender for Multi-Domain CTR Prediction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 4104–4113.

[33] Xing Tang, Chaohua Yang, Yuwen Fu, Dongyang Ao, Shiwei Li, Fuyuan Lyu, Dugang Liu, and Xiuqiang He. 2025. Retrieval Augmented Cross-Domain Life-Long Behavior Modeling for Enhancing Click-through Rate Prediction. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*. 4891–4900.

[34] Lu Wang, Di Zhang, Fangkai Yang, Pu Zhao, Jianfeng Liu, Yuefeng Zhan, Hao Sun, Qingwei Lin, Weiwei Deng, Dongmei Zhang, Feng Sun, and Qi Zhang. 2025. LettinGo: Explore User Profile Generation for Recommendation System. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*. 2985–2995.

[35] Ruoxi Wang, Rakesh Shivanna, Derek Cheng, Sagar Jain, Dong Lin, Lichan Hong, and Ed Chi. 2021. DCN V2: Improved Deep & Cross Network and Practical Lessons for Web-scale Learning to Rank Systems. In *Proceedings of the Web Conference 2021*. 1785–1797.

[36] Yuhan Wang, Qing Xie, Zhifeng Bao, Mengzi Tang, Lin Li, and Yongjian Liu. 2025. KGBridge: Knowledge-Guided Prompt Learning for Non-overlapping Cross-Domain Recommendation. *arXiv preprint arXiv:2511.02181* (2025).

[37] Yunjia Xi, Weiwen Liu, Jianghao Lin, Xiaoling Cai, Hong Zhu, Jieming Zhu, Bo Chen, Ruiming Tang, Weinan Zhang, and Yong Yu. 2024. Towards Open-World Recommendation with Knowledge Augmentation from Large Language Models. In *Proceedings of the 18th ACM Conference on Recommender Systems*. 12–22.

[38] Wei Xu, Haoran Li, Baoyuan Ou, Lai Xu, Yingjie Qin, Ruilong Su, and Ruiwen Xu. 2025. GIST: Cross-Domain Click-Through Rate Prediction via Guided Content-Behavior Distillation. *arXiv preprint arXiv:2507.05142* (2025).

[39] Wujiang Xu, Qitian Wu, Runzhong Wang, Mingming Ha, Qiongxu Ma, Linxun Chen, Bing Han, and Junchi Yan. 2024. Rethinking Cross-Domain Sequential Recommendation under Open-World Assumptions. In *Proceedings of the ACM Web Conference 2024*. 3173–3184.

[40] Zhiming Yang, Haining Gao, Dehong Gao, Luwei Yang, Libin Yang, Xiaoyan Cai, Wei Ning, and Guannan Zhang. 2024. MLoRA: Multi-Domain Low-Rank Adaptive Network for CTR Prediction. In *Proceedings of the 18th ACM Conference on Recommender Systems*. 287–297.

[41] Wei Zhang, Pengye Zhang, Bo Zhang, Xingxing Wang, and Dong Wang. 2023. A Collaborative Transfer Learning Framework for Cross-domain Recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 5576–5585.

[42] Yanzhao Zhang, Mingxin Li, Dingkun Long, Xin Zhang, Huan Lin, Baosong Yang, Pengjun Xie, An Yang, Dayiheng Liu, Junyang Lin, et al. 2025. Qwen3 Embedding: Advancing Text Embedding and Reranking Through Foundation Models. *arXiv preprint arXiv:2506.05176* (2025).

[43] Yu Zhang and Qiang Yang. 2022. A Survey on Multi-Task Learning. *IEEE Transactions on Knowledge and Data Engineering* 34, 12 (2022), 5586–5609. doi:10.1109/TKDE.2021.3070203

[44] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep Interest Network for Click-Through Rate Prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1059–1068.

[45] Feng Zhu, Yan Wang, Chaochao Chen, Jun Zhou, Longfei Li, and Guanfeng Liu. 2021. Cross-domain recommendation: Challenges, progress, and prospects. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*. 4721–4728.