

# You Only evaluate Once: A Tree-based Rerank Method at Meituan

Shuli Wang\*  
Meituan  
Chengdu, China  
wangshuli03@meituan.com

Yinqiu Huang  
Meituan  
Chengdu, China  
huangyinqiu@meituan.com

Changhao Li  
Meituan  
Chengdu, China  
lichanghao@meituan.com

Yuan Zhou  
Meituan  
Chengdu, China  
zhouyuan22@meituan.com

Yonggang Liu  
Meituan  
Chengdu, China  
liuyonggang02@meituan.com

Yongqiang Zhang  
Meituan  
Chengdu, China  
zhangyongqiang08@meituan.com

Yinhua Zhu  
Meituan  
Chengdu, China  
zhuyinhua@meituan.com

Haitao Wang  
Meituan  
Chengdu, China  
wanghaitao13@meituan.com

Xingxing Wang  
Meituan  
Beijing, China  
wangxingxing04@meituan.com

## Abstract

Reranking plays a crucial role in modern recommender systems by capturing the mutual influences within the list. Due to the inherent challenges of combinatorial search spaces, most methods adopt a two-stage search paradigm: a simple General Search Unit (GSU) efficiently reduces the candidate space, and an Exact Search Unit (ESU) effectively selects the optimal sequence. These methods essentially involve making trade-offs between effectiveness and efficiency, while suffering from a severe **inconsistency problem**, that is, the GSU often misses high-value lists from ESU. To address this problem, we propose YOLOR, a one-stage reranking method that removes the GSU while retaining only the ESU. Specifically, YOLOR includes: (1) a Tree-based Context Extraction Module (TCM) that hierarchically aggregates multi-scale contextual features to achieve "list-level effectiveness", and (2) a Context Cache Module (CCM) that enables efficient feature reuse across candidate permutations to achieve "permutation-level efficiency". Extensive experiments across public and industry datasets validate YOLOR's performance and we have successfully deployed YOLOR on the Meituan food delivery platform.

## CCS Concepts

• Information systems → Retrieval models and ranking; Computational advertising.

\* Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

Conference acronym 'XX, Woodstock, NY

© 2018 ACM.

ACM ISBN 978-1-4503-XXXX-X/18/06

<https://doi.org/XXXXXXX.XXXXXXX>

## Keywords

Recommender Systems, E-commerce, Reranking

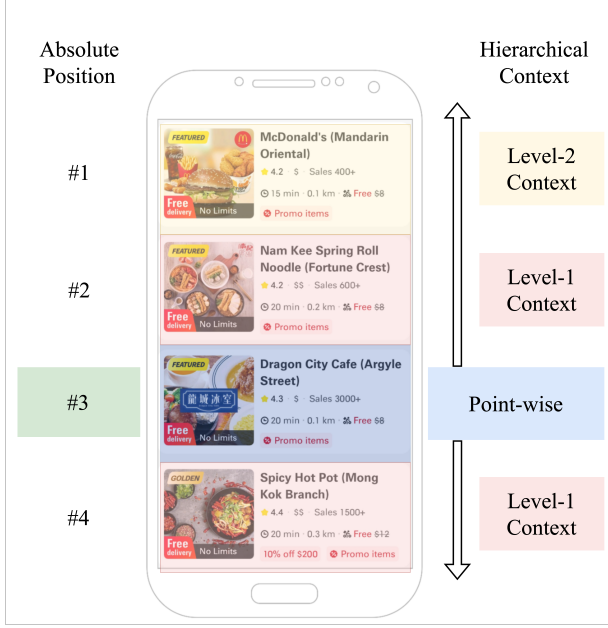
### ACM Reference Format:

Shuli Wang, Yinqiu Huang, Changhao Li, Yuan Zhou, Yonggang Liu, Yongqiang Zhang, Yinhua Zhu, Haitao Wang, and Xingxing Wang. 2018. You Only evaluate Once: A Tree-based Rerank Method at Meituan. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 8 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 Introduction

E-commerce platforms, such as Meituan and Taobao, need to provide users with personalized services from millions of items. As shown in Figure 1, list recommendation is the main display form on Meituan food delivery platform. To improve recommendation efficiency, personalized recommendation systems generally include three stages: matching, ranking, and reranking. The ranking models evaluate the recommended items respectively, focusing on feature interactions [10, 15, 17], user preference modeling [23, 31, 32], and so on. However, ranking methods ignore the crucial mutual influence among contextual items. Research [1, 4, 5, 21] indicates that optimizing a listwise utility during the reranking stage is a more advantageous strategy, as it capitalizes on the mutual influences between items within the list to enhance overall performance.

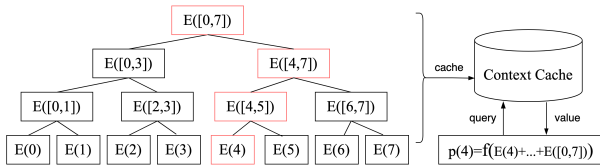
The key challenge in reranking is exploring optimal lists within the vast permutation space. Existing reranking methods can be classified into two categories [13, 26]. The first category is generator-based methods. These methods generate the list by some heuristic strategy [3, 14, 33], or generate a suboptimal list with greedy local sight [13]. Despite considering the context, these methods fail to ensure the quality of the generated lists and obtain optimal results, due to issues such as evaluation-before-reranking problems [29]. The second category is evaluator-based methods [12, 26, 29]. These methods try to evaluate every possible permutation to get the optimal list. However, due to the strict time constraints in online systems, most existing evaluator-based methods use a two-stage architecture, which adds a filtering submodel [12, 26] or generating



**Figure 1: List recommendations on Meituan food delivery platform.**

submodel [24] as the General Search Unit (GSU) before the Exact Search Unit (ESU) to reduce the size of the candidate space.

However, existing two-stage reranking methods fall into trade-off and face two significant issues. **Firstly, the inherent inconsistency problem between the two stages limits the effectiveness of the GSU.** GSU is expected to search high-value lists to ESU. However, due to strict time constraints, the GSU can only be designed simply but not precisely. Although some recent generative reranking methods [13, 18, 24] leverage the ESU as a teacher to guide the GSU, GSU's inconsistency issues remain serious [26, 28] and lead to unsatisfactory hit ratio due to the vast permutation space. The inconsistency problem arises from the trade-offs inherent in the two-stage model [6] which is challenging to solve. **Secondly, lacking of multi-scale contextual information limits the effectiveness of ESU.** As shown in Figure 1, in the list recommendation tasks, the context information between items is complex, and the impact windows between items may span 2, 4 or even the entire list.



**Figure 2: Demo of YOLOR. When evaluating the 4th item, we can obtain its precise score by combining its multi-scale contextual information and achieve efficient reuse through caching.**

To resolve the aforementioned issues, we propose a tree-based Rerank framework that You Only evaluate Once, termed YOLOR, which removes the GSU while retaining only the ESU. We resolve the efficiency-effectiveness trade-off by simultaneously addressing "permutation-level efficiency" and "list-level effectiveness". As shown in Figure 2, for effectiveness at the list level, we designed a tree-based context extraction module to capture contextual information at different scales. For efficiency at the permutation level, we developed a context caching module to enable efficient reuse of multi-scale contextual information across lists. It is an extremely time-saving method that allows you to predict all permutations.

The main contributions of our work are summarized as follows:

- We propose a novel framework that resolves the efficiency-effectiveness trade-off through simultaneous optimization of "permutation-level efficiency" and "list-level effectiveness". To the best of our knowledge, we are the first to introduce this concept.
- Based on this idea, we design the YOLOR model, which contains the TCEM module to extract multi-scale contextual information and the CCM module to reduce repeated calculations.
- We conduct extensive experiments on both offline and real-world industrial datasets from Meituan. Experimental results demonstrate the effectiveness of YOLOR. It is notable that YOLOR has been deployed in Meituan food delivery platform and has achieved significant improvement under various metrics.

## 2 Related Work

In recommendation systems, the core of the reranking stage lies in modeling the context and selecting the optimal list from the permutation space. Existing research on reranking can be systematically classified into two principal categories[19]: generator-based methods [13, 22, 30] and evaluator-based methods [12, 26].

Generator-based methods directly generate one list as output by capturing the mutual influence among items. Therefore, these generator-based methods are also called one-stage methods [24, 26]. Initially, these methods directly used behavioral logs as training guides. For instance, Seq2slate [3] utilizes pointer-network and MIRNN [33] utilizes GRU to determine the item order one-by-one. Methods such as PRM [22] and DLCM [1] take the initial ranking list as input, use RNN or self-attention to model the context-wise signal, and output the predicted value of each item. Such methods bring an evaluation-before-reranking problem [29] and lead to sub-optimum. Similarly, methods such as EXTR [8] estimate the predicted Click-Through Rate (pCTR) of each candidate item on each candidate position, which are substantially point-wise models and thus limited in extracting exact context. MIR [30] capturing the set2list interactions by a permutation-equivariant module. Since it is difficult to achieve permutation space optimization in supervised training, some generator-based methods using evaluators [13, 16] have become popular in recent years. For instance, ListCVAE [16] utilizes conditional variational autoencoders (CVAE) to capture the positional biases of items and the interdependencies within the list distribution. GRN [13] proposes an evaluator-generator framework

to replace the greedy strategy, but it can't avoid the evaluation-before-reranking problem [29] because it takes the rank list as input to the generator. DCDR [18] introduces diffusion models into the reranking stage and presents a discrete conditional diffusion reranking framework. NAR4Rec [24] uses a non-autoregressive generative model to speed up sequence generation. However, this paradigm heavily depends on the accuracy of the evaluator which makes it less promising in industrial recommendation tasks.

Evaluator-based methods try to evaluate every possible permutation through a well-designed context-wise model. Due to the strict time constraints in online systems, most existing evaluator-based methods use a two-stage architecture, which adds a filtering stage [12, 26] or generating stage [24] before the evaluating stage to reduce the size of the candidate set. For instance, PRS [12] adopts beam-search to generate a few candidate permutations first, and score each permutation through a permutation-wise ranking model. PIER [26] applies SimHash [7, 9, 20] to select top-K candidates from the full permutation. In industrial recommendation systems, the concept of evaluator-based methods is broader, as any generator-based method can be utilized in the first stage and form a multi-channel retrieval framework [24]. However, the inconsistency problem between the two stages limits the effectiveness of the model.

### 3 Problem Definition

Let  $U = \{u_1, u_2, \dots, u_{|U|}\}$  represent a set of  $|U|$  users which consist of some profile features (e.g. user ID, gender, age) and recent interaction history. For each user  $u$ , given a candidate set with  $n$  items  $X = \{x_1, x_2, \dots, x_n\}$ , the goal of reranking is to propose an ordered list with  $m$  ( $m \leq n$ ) items  $L = \{x_1, x_2, \dots, x_m\}$  from  $O(A_n^m)$  candidate space  $\mathcal{L}$  and maximizes the listwise reward  $\mathcal{R}(u, L)$ :

$$L^* = \arg \max_L \mathcal{R}(u, \mathcal{L}), \quad (1)$$

where  $\mathcal{R}(u, L) = \sum_{i=1}^m \mathcal{R}(u, x_i)$ .

### 4 Proposed Method

In this section, we will introduce the structure of YOLOR in detail. As shown in Figure 3, YOLOR mainly includes three modules, namely the Item-level Representation Module (IRM), the Tree-based Context Extraction Module (TCEM) for modeling contextual information, and Context Cache Module (CCM). We will introduce them in detail in the following subsections.

#### 4.1 Item-level Representation Module

We use the Item-level Representation Module (IRM) to generate the semantic embedding of each candidate item from the raw input. First, we use an embedding layer to get the embedding of the original input. We denote the embeddings of the original input user profile features, context features, user behaviors sequence and candidate items features as  $\mathbf{e}^u$ ,  $\mathbf{e}^c$ ,  $\mathbf{E}^b \in \mathbb{R}^{N_b \times D}$  and  $\mathbf{X} \in \mathbb{R}^{n \times D}$  respectively, where  $N_b$  and  $n$  are the number of user behaviors and candidate items,  $D$  is the dimension of the embedding layer. Then, we use a target attention unit to encode the interaction between the historical behaviors of the user and the corresponding item:

$$\mathbf{x}'_i = \text{Attention}(\mathbf{X}_i, \{\mathbf{E}^b_j\}_{j=1}^{N_b}), \forall i \in [n], \quad (2)$$

where  $\mathbf{X}_i \in \mathbb{R}^D$  is the  $i$ -th candidate item in  $\mathbf{X}$ .

Then, we use MLP as a simple feature crosses unit to extract the semantic embedding of each candidate item:

$$\mathbf{x}^s_i = \text{MLP}(\mathbf{x}'_i || \mathbf{e}^u || \mathbf{e}^c), \forall i \in [n], \quad (3)$$

where  $||$  represents concatenate operate.

For ease of notation, we can also write the semantic embeddings for items in matrix form, each row of which represents one item in the sequence, i.e.,

$$\mathbf{X}^s = [\mathbf{x}^s_1; \mathbf{x}^s_2; \dots; \mathbf{x}^s_n]^\top. \quad (4)$$

Note that this is just the simplest implementation. IRM is only  $O(n)$  complexity rather than  $O(A_n^m)$ , so it can be more complex.

Unlike other reranking methods that solely rely on the predicted scores from ranking models, we directly deploy the ranking model as IRM in our online system, which enhances the consistency of the recommendation pipeline. The details will be discussed in Section 5.6.

#### 4.2 Tree-based Context Extraction Module

We design the Tree-based Context Extraction Module (TCEM) to model multi-scale contextual relationships. For the  $t$ -th item  $x_t$  in the candidate list  $L$ , we construct multi-scale subsequences to extract  $x_t$ 's local and global contextual information. First, We split the list  $L$  constantly until there are only two items left in the subsequence, then we obtain the set of all subsequence denoted as:

$$C = \{L, L_{1,m/2}, L_{m/2+1,m}, L_{1,m/4}, \dots, L_{t,t+1}, \dots, L_{m-1,m}\}, \quad (5)$$

where  $L_{l,r}$  represents the subsequence of list  $L$  from index  $l$  to index  $r$ , denoted as  $L_{l,r} = [\mathbf{x}^s_l; \mathbf{x}^s_{l+1}; \dots; \mathbf{x}^s_r]$ .

Then, we collect all subsequences containing  $x_t$ , denoted as  $C_t = \{L, \dots, L_{t,t+1}\}$ , and perform contextual information extraction for each subsequence in  $C_t$ : And then we input each subsequence in  $C_t$  into a Self-Attention layer (SA) [27] to extract contextual information  $\mathbf{e}_{l,r} \in \mathbb{R}^D$ :

$$\begin{aligned} \mathbf{e}_{(1)} &= \mathbf{e}_{1,m} = \text{SA}(\mathbf{x}^s_1 || \mathbf{x}^s_2 || \dots || \mathbf{x}^s_m), \\ \mathbf{e}_{(2)} &= \mathbf{e}_{l,r} = \text{SA}(\mathbf{x}^s_l || \mathbf{x}^s_{l+1} || \dots || \mathbf{x}^s_r), \\ &\dots \\ \mathbf{e}_{(\log_2 m)} &= \mathbf{e}_{t,t+1} = \text{SA}(\mathbf{x}^s_t || \mathbf{x}^s_{t+1}), \end{aligned} \quad (6)$$

where  $||$  represents concatenate operate. Note that the SA layer in the above formula does not have position encoding, which can increase reusability and reduce calculations, also termed Set Attention. Through Eq. 6, we can obtain multi-scale contextual relationships associated with  $x_t$ , denoted as  $\mathbf{X}^C_t = [\mathbf{e}_{(1)}; \mathbf{e}_{(2)}; \dots; \mathbf{e}_{(\log_2 m)}] \in \mathbb{R}^{D \cdot \log_2 m}$ .

Finally, we use a parameter-sharing fully connected layer to predict the list-wise pCTR of each item in each permutation. Taking the  $t$ -th item in permutation as an example still, the inputs consist of three parts: absolute position representation of  $t$ -th position  $\mathbf{E}^p_t \in \mathbb{R}^D$ , item representation of  $t$ -th position  $\mathbf{X}^s_t \in \mathbb{R}^D$ , and context representation of  $t$ -th position  $\mathbf{X}^C_t$ . Then the list-wise pCTR of the  $t$ -th item is predicted as follows:

$$\hat{y}_t = \sigma \left( \text{FC}(\mathbf{E}^p_t || \mathbf{X}^s_t || \mathbf{X}^C_t) \right), \quad (7)$$

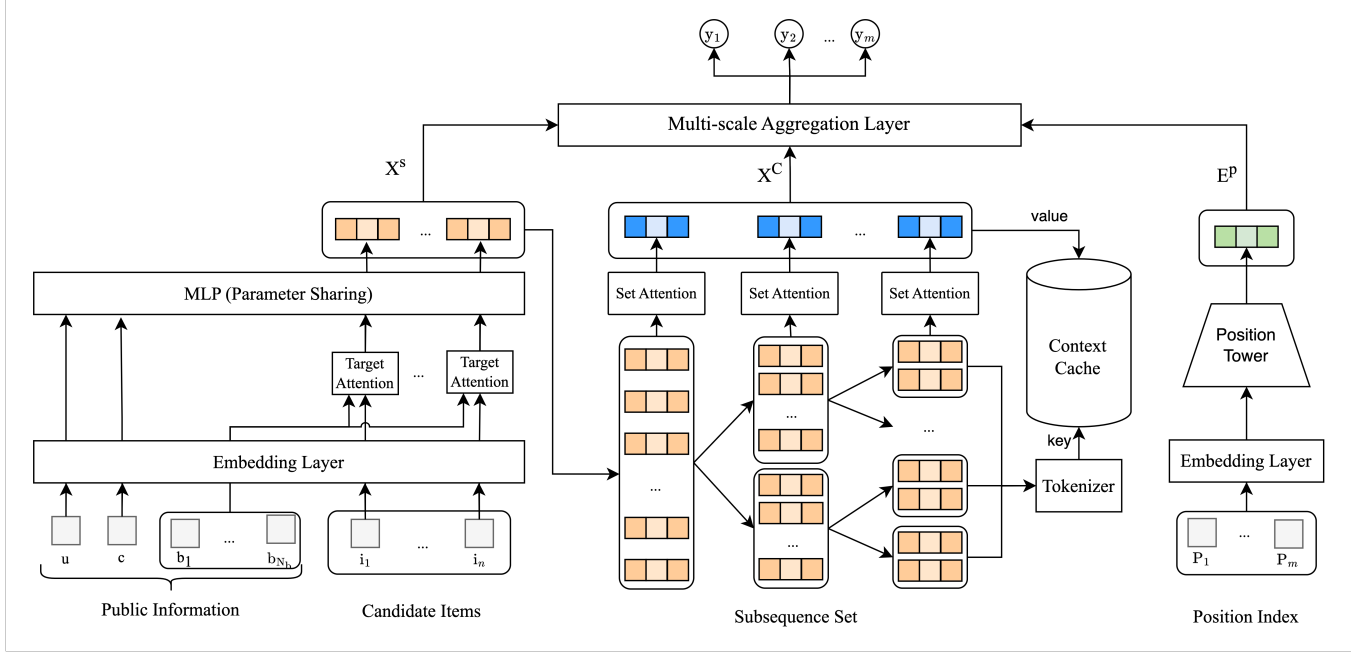


Figure 3: The overall architecture of YOLOR.

where  $\sigma$  is the Sigmoid Function. The score of each list is easily obtained by summing the output list-wise pCTR:

$$\hat{y}_L = \sum (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_t, \dots, \hat{y}_m). \quad (8)$$

### 4.3 Context Cache Module

We have obtained the item-level representation and contextual representation of the  $t$ -th item  $x_t$  through the aforementioned IRM and TCEM modules, respectively. However, due to strict time constraints, such a complex method cannot be applied directly to the permutation space. Therefore, we further design a Context Cache Module (CCM) to enable efficient reuse, which only requires simple matrix operation.

First, we generate the set of all subsequences (including candidate items  $X$ ) in the permutation space. To avoid notational confusion, we still denote this set as  $C$ . Clearly, the size of  $C$  is  $|C| = C_n^m + C_n^{m/2} + \dots + C_n^1$ . Then, based on Eq. 6, we extract contextual information to obtain the multi-scale context representation matrix  $X^C \in \mathbb{R}^{|C| \times D}$ .

Next, to evaluate all candidate lists in the permutation space  $\mathcal{L} \in \mathcal{R}^{A_n^m \times m}$ , we gather the required contextual information for  $\mathcal{L}$  from  $X^C$ , denoted as:

$$X_{\mathcal{L}}^C = \text{tf.gather}(X^C, M_{\text{indices}}), \quad (9)$$

where  $\text{tf.gather}^1$  is Tensorflow's gather operator. And  $M_{\text{indices}} \in \mathcal{R}^{A_n^m \times m \times \log_2 m}$  is a request-independent matrix as long as  $m$  and  $n$  are fixed.

Then, based on Eq. 7 and Eq. 8, we calculate the list-wise score for each candidate list  $\hat{y}_{\mathcal{L}} \in \mathcal{R}^{A_n^m}$ :

$$\hat{y}_{\mathcal{L}} = \text{reduce\_sum} \left( \sigma \left( \text{FC}(\text{tile}(E^P) \| X_{\mathcal{L}}^C) \right), \text{axis} = -1 \right). \quad (10)$$

Finally, based on Eq. 1, we select the optimal list:

$$L^* = \arg \max_L \hat{y}_{\mathcal{L}}. \quad (11)$$

It should be noted that the score of each list can be conveniently adjusted according to business needs, such as CVR (Conversion Rate) and GMV (Gross Merchandise Volume).

### 4.4 Model Complexity

We perform a model complexity analysis of YOLOR to illustrate that our model meets the standards for online deployment. As mentioned, reranking models face serious challenges of  $O(A_n^m)$  candidate space. YOLOR uses the CCM module to reduce calculations. The computational complexity of LRM is  $O(n)$ . The original computational complexity of TCEM is  $O(A_n^m * \log_2 m)$ . But through the CCM module, we can fully reuse the context information. The additional space complexity required by CCM is  $O(C_n^m, C_n^{m/2}, \dots, C_n^2, C_n^1)$ . Take  $n = m = 8$  as an example, there are a total of  $O(A_8^8) = 40320$  lists in the permutation space, and CCM requires storing only  $O(C_8^8, C_8^4, C_8^2, C_8^1) = 107$  context embeddings, a few amount of storage is sufficient to enable rapid computations at the permutation level. The remaining only  $O(A_n^m)$  complexity in YOLOR is Eq. 10, and the computational cost of Eq. 9 is negligible compared to that of Eq. 10. We design the prediction layer as a single FC layer, resulting in minimal computational complexity. Therefore, the complexity of YOLOR is acceptable for online serving.

<sup>1</sup> [https://www.tensorflow.org/api\\_docs/python/tf/gather](https://www.tensorflow.org/api_docs/python/tf/gather)



## 4.5 Model Training

We train YOLOR using real data collected from online logs. The input is the features of the recommended advertisement lists exposed in reality online, and the advertising return situation, including exposure, click, conversion, and other performance indicators, is used as the label to supervise the training of YOLOR. First, we use cross-entropy loss to train YOLOR, the loss of each list is calculated as follows:

$$\mathcal{L}_{ce} = -\frac{1}{m} \sum_{t=1}^m (y_t \log(\hat{y}_t) + (1 - y_t) \log(1 - \hat{y}_t)), \quad (12)$$

where subscript  $t$  is the index of displayed items,  $y_t$  represents the real label,  $\hat{y}_t$  represents the predicted value,  $m$  is the length of exposed list. Additionally, we set a dropout rate to mask contextual information randomly.

Then, to enhance the context extraction ability, we propose GBPR loss (Group BPR [25] loss) to enhance the contrast of the positive and negative samples within each exposure list:

$$\mathcal{L}_{gbpr} = -\frac{1}{m} \sum_{i=1}^m \log(\text{sgn}(y_i - y_j)(\hat{y}_i - \hat{y}_j)), \quad (13)$$

where subscript  $i$  and  $j$  are the indexes of displayed items. And  $\text{sgn}(x)$  represents sign function,  $\text{sgn}(x) = 0$  if  $x = 0$ ,  $\text{sgn}(x) = 1$  if  $x > 0$  and  $\text{sgn}(x) = -1$  if  $x < 0$ .

Finally, we sample a batch of samples  $\mathcal{B}$  from the dataset and update YOLOR using gradient back-propagation w.r.t. the loss:

$$\mathcal{L} = \frac{1}{|\mathcal{B}|} \sum_{\mathcal{B}} (\mathcal{L}_{ce} + \alpha \cdot \mathcal{L}_{gbpr}), \quad (14)$$

where  $\alpha$  is the coefficient to balance the two losses.

## 5 Experiments

To validate the superior performance of YOLOR, we conducted extensive offline experiments on the Meituan dataset and verified the superiority of YOLOR in online A/B tests. In this section, we first introduce the experimental setup, including the dataset and baseline. Then, in Section 5.2, we present the results and analysis of various reranking methods in both offline and online A/B tests.

### 5.1 Experimental Setup

**5.1.1 Dataset.** In order to verify the effectiveness of YOLOR, we conduct sufficient experiments on both public dataset and industrial dataset. For public dataset, we choose Taobao Ad dataset. For industrial dataset, we use real-world data collected from Meituan food delivery platform. Table 1 gives a brief introduction to the datasets.

**Table 1: Statistics of datasets.**

Dataset	#Users	#Items	#Records
Taobao Ad	1,141,729	99,815	26,557,961
Meituan	5,648,310	14,054,691	161,247,488

- **Taobao Ad**<sup>2</sup>. It is a public dataset collected from the display advertising system of Taobao. This dataset contains more than 26 million interaction records of 1.14 million users within 8 days. Each sample comprises five features: user ID, timestamp, behavior type, item brand ID, and category ID. It includes four behavior types: browse, cart, like, and buy, and each behavior is timestamped. We use the first 7 days as training samples (20170506-20170512), and the 8th day as test samples (20170513).
- **Meituan**. It is an industrial dataset collected from the Meituan food delivery platform during August 2024, which contains 161 million interaction records of 5.6 million users within 15 days. The dataset includes 239 features, two labels: click and conversion, and collects all items on the same page as one record. We use the data of the first 14 days as the training set, and the data of the last 1 day as the test set.

Note that all samples are list-level, that is, each sample contains all items in an exposed list. We filter out samples whose labels are all 0 or all 1.

**5.1.2 Baseline.** The following six state-of-the-art reranking methods are chosen for comparative experiments and divided into three groups. We select DNN and DeepFM as point-wise baselines (Group I), PRM and MIR as one-stage generator-based baselines (Group II), and Edge-Rerank and PIER as two-stage evaluator-based baseline methods (Group III). A brief introduction of these methods is as follows:

- **DNN**[11] is a basic deep learning method for CTR prediction, which applies MLP for high-order feature interaction.
- **DeepFM**[15] is a general deep model for recommendation, which combines a factorization machine component and a deep neural network component.
- **PRM**[22] adjusts an initial list by applying the self-attention mechanism to capture the mutual influence between items.
- **MIR**[30] learns permutation-equivariant representations for the inputted items via self-attention mechanism to capture the mutual influence between items.
- **Edge-Rerank**[14] generates the context-aware sequence with adaptive beam search on estimate scores.
- **PIER**[26] applies hashing algorithm to select top-k candidates from the full permutation based on user interests.

**5.1.3 Evaluation Metrics.** We adopt several metrics, i.e., **AUC** (Area Under ROC Curve) and **HR** (Hit Ratio) to evaluate YOLOR in offline experiments. To make AUC more suitable for reranking models, we adapt AUC as **GAUC**[32] (average intra-list AUC) in our experiments. AUC measures the global estimated accuracy, and GAUC measures the estimated accuracy within the list. High values in both AUC and GAUC indicate that the model excels at ranking positive samples in front of negative samples, demonstrating strong discriminative power across different evaluation contexts. In online experiments, we adopt CTR and GMV as evaluation metrics.

We use **HR (Hit Ratio)** [2] to evaluate the consistency of the two stages. For each data, HR is 1 only when the permutations selected by GSU contain the best permutation. Obviously, the HR metrics are only meaningful with evaluator-based reranking methods. The results of HR can be seen in Section 5.3.

<sup>2</sup> <https://tianchi.aliyun.com/dataset/56>

It is worth noting that AUC/GAUC and HR can measure two aspects of two-stage methods. AUC/GAUC measures the model's ability to evaluate an ordered list, while HR measures the consistency of the two stages. The deficiency of any indicator will reduce the recommendation effect. For example, when the reranking model directly returns the result of the ranking stage, the HR will be up to 1 but the AUC will decrease. When the reranking model is very complex, the AUC will increase but the increased time-cost will lead to a decrease in the HR.

**5.1.4 Implementation Details.** We implement all the deep learning baselines and YOLOR with TensorFlow 1.15.0 using NVIDIA A100-80GB GPU. For all comparison models and our YOLOR model, we adopt Adam as the optimizer with the learning rate fixed to 0.001 and initialize the model parameters with normal distribution by setting the mean and standard deviation to 0 and 0.01, respectively. The batch size is 1024, the embedding size is 8. The hidden layer sizes of MLP in Eq. 3 are (1024, 256, 128). For the Taobao Ad dataset, the length of the ranking list and reranking list are both 5, thus the length of full permutation is 120. For Meituan dataset page, we select 8 items from the initial ranking list which contains 8 items, thus the length of full permutation is  $A_8^8 = 40,320$ . For the baseline methods, we follow the settings in PIER and set the number of candidate lists to 100. Similarly, we follow the edge settings and set the beam size to 3. All experiments are repeated 5 times and the averaged results are reported.

## 5.2 Overall Performance

Here we show the results of our proposed method YOLOR. All results are averaged from 5 experiments. As can be seen in Table 2 and Table 3, YOLOR outperforms baselines including recent two-stage evaluator-based reranking methods. We have the following observations from the experimental results: i) All re-ranking listwise model (e.g. PRM, MIR) makes great improvements over point-wise model (e.g. DNN, DeepFM) by modeling the mutual influence among contextual items, which verifies the impact of context on user clicks behavior. ii) Compared with generator-based methods (e.g. PRM, MIR), evaluator-based methods also improve the CTR prediction because they evaluate more candidate lists. iii) Our proposed YOLOR brings 0.0035/0.0047 absolute AUC and 0.0113/0.0111 absolute GAUC on Taobao/Meituan dataset gains over the state-of-the-art independent baseline which is a significant improvement in industrial recommendation system.

**Table 2: Comparison between YOLOR and baseline methods on the Taobao Ad dataset. The best and second-best results in each column are in bold and underlined.**

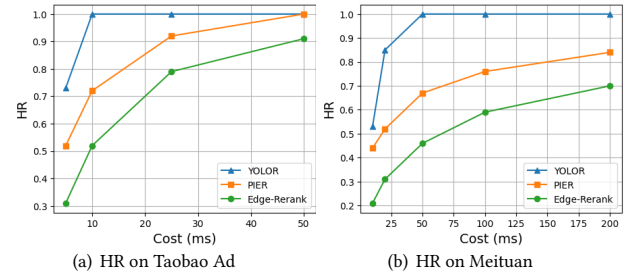
Model	AUC	GAUC	Loss
DNN	0.5869	0.8130	0.1878
DeepFM	0.5891	0.8132	0.1866
PRM	0.6152	0.8163	0.1842
MIR	0.6147	0.8169	0.1853
Edge-Rerank	0.6286	0.8201	0.1781
PIER	<u>0.6316</u>	<u>0.8210</u>	<u>0.1758</u>
YOLOR	<b>0.6351</b>	<b>0.8323</b>	<b>0.1743</b>

**Table 3: Comparison between YOLOR and baseline methods on Meituan. The best and second-best results in each column are in bold and underlined.**

Model	AUC	GAUC	Loss
DNN	0.7347	0.7418	0.1162
DeepFM	0.7392	0.7442	0.1154
PRM	0.7573	0.7595	0.1108
MIR	0.7598	0.7603	0.1101
Edge-Rerank	0.7586	0.7605	0.1097
PIER	<u>0.7622</u>	<u>0.7638</u>	<u>0.1068</u>
YOLOR	<b>0.7669</b>	<b>0.7749</b>	<b>0.1032</b>

## 5.3 Consistency Analysis

We compared the HR (Hit Ratio) of Edge-Rerank, PIER and YOLOR under different inference times on Taobao Ad and Meituan datasets (only evaluator-based methods have HR metrics). All experiments were performed on NVIDIA A100-80GB GPU, with the batch size set to 1024 and averaged 100 times. In the Meituan data set, there are a total of  $A_8^8 = 40,320$  candidate lists. We carefully adjust the number of candidate lists of the model to make the model time-consuming equal, and keep the error within 1 ms. For YOLOR, we use random generation to conduct experiments. As shown in Figure 4, we have the following observations: i) YOLOR performs well across all periods. This is because YOLOR is a very efficient method that is capable of retrieving more candidate lists at the same time. ii) As time increases, the HR of all methods improves, since the amount of retrieval lists for all methods becomes larger. iii) YOLOR can traverse all candidate lists within 50ms, achieving an HR of 1.



**Figure 4: HR results of YOLOR and baseline methods under different cost conditions on Taobao Ad and Meituan datasets.**

## 5.4 Ablation Study

To assess the effectiveness of each component in YOLOR, we conducted a series of ablation studies using the Taobao Ad and Meituan datasets. Specifically, we build several variants of the YOLOR:

- w/o IRM. A variant of YOLOR without the IRM, which means only using candidate items' raw embedding  $X$  instead of the semantic embedding  $X^s$ .

- w/o TCEM. A variant of YOLOR without the TCEM, which is replaced by a single global self-attention.
- w/o GBPR. A variant of YOLOR without the GBPR loss in Eq. 13.

**Table 4: The contributions of different components of YOLOR.**

Model	Taobao Ad		Meituan	
	AUC	GAUC	AUC	GAUC
w/o IRM	0.5712	0.7940	0.7332	0.7454
w/o TCEM	0.6236	0.8225	0.7574	0.7621
w/o GBPR	0.6305	0.8207	0.7629	0.7632
YOLOR	<b>0.6351</b>	<b>0.8323</b>	<b>0.7669</b>	<b>0.7749</b>

Table 4 shows the results of the ablation study, and we can draw the following conclusions: 1) without the IRM, the model’s performance decreases significantly because accurate point-wise predictions are the basis of YOLOR; 2) without TCEM, the performance of the model declines, indicating that TCEM has stronger context extraction capabilities comparing with single self-attention; 3) without GBPR, the performance of the model declines, indicating that the intra-list loss is valid for reranking.

To assess the efficiency of CCM, we remove the CCM module and randomly sample  $K$  candidate lists to evaluate their HR and time-consuming. And we set different  $K$  values for detailed comparison.

**Table 5: HR and Cost of different sample numbers without CCM.**

Settings	Taobao Ad		Settings	Meituan	
	HR	Cost (ms)		HR	Cost (ms)
K=5	0.0417	4.7	K=100	0.0025	56.1
K=10	0.0833	7.1	K=200	0.0050	103.4
K=20	0.1667	11.8	K=300	0.0074	169.8
K=50	0.4167	28.6	K=400	0.0099	217.7

As shown in Table 5, after the CCM module is removed, all vectors in both IRM and TCEM must be recalculated, which significantly increases the average time cost. Furthermore, due to random sampling constraints, the Hit Ratio (HR) shows a marked decline as it is sensitive to the number of evaluated candidate lists.

## 5.5 Hyperparameter Analysis

We analyze the impact of weight  $\alpha$  in GBPR loss. Table 6 shows the results of our experiments and we can find that  $\alpha$  significantly affects YOLOR’s AUC/GAUC metric. As  $\alpha$  increases, AUC first increases and then decreases while GAUC increases and stabilizes at a high level.

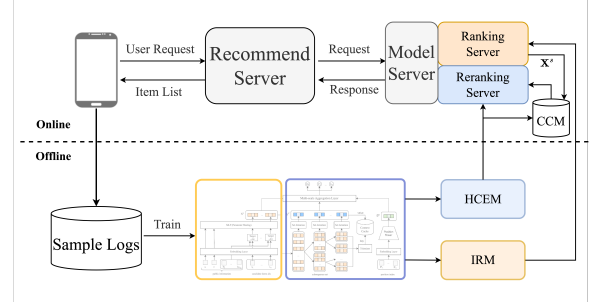
## 5.6 Performance on Online System

To evaluate the online performance of YOLOR, we deployed YOLOR on the Meituan Shichiguan business as shown in Figure 5.

We also conducted a rigorous A/B test for three weeks, from March 2025 to April 2025. Specifically, we assigned YOLOR with

**Table 6: The effect of parameter weight  $\alpha$  in GBPR loss.**

Settings	Taobao Ad		Meituan	
	AUC	GAUC	AUC	GAUC
$\alpha=0$	0.6305	0.8207	0.7629	0.7632
$\alpha=0.01$	0.6334	0.8297	0.7641	0.7704
$\alpha=0.05$	<b>0.6351</b>	<b>0.8323</b>	<b>0.7669</b>	0.7749
$\alpha=0.1$	0.6345	0.8323	0.7664	0.7749
$\alpha=0.5$	0.6341	0.8323	0.7658	<b>0.7750</b>

**Figure 5: Architecture of the online deployment with YOLOR.**

30% traffic, while the remaining 70% traffic was assigned to baseline (PIER). Table 7 shows the online performance of YOLOR. Compared to the baseline model (PIER), YOLOR has increased the CTR by 5.13% and the GMV by 7.64%, which are very significant growth to business. Besides, we find that time consumption has not increased at all, which is an important indicator to determine whether it can be applied to large-scale industrial scenarios. Now, YOLOR is deployed on the Meituan food delivery platform and serves millions of users.

**Table 7: Online A/B test result.**

Method	CTR	GMV	Cost (ms)	Time-out
YOLOR	+5.13%	+7.64%	-0.003	-0.001%

## 6 Conclusion

In this paper, we identify the inherent trade-off issues in two-stage re-ranking methods, which cannot simultaneously balance efficiency and effectiveness. To overcome this limitation, we propose a tree-based approach, named YOLOR, designed to achieve both "permutation-level efficiency" and "list-level effectiveness." Specifically, we design a Tree-based Context Extraction Module (TCEM) that integrates multi-scale contextual information to accurately evaluate the list. Additionally, we develop a Context Cache Module (CCM) to enable efficient reuse of multi-scale contextual information across lists. The efficiency of YOLOR allows you to only evaluate once. Both offline experiments and online A/B tests show that YOLOR significantly outperformed other existing reranking baselines. We have deployed YOLOR on the Meituan food delivery platform.

## References

- [1] Qingyao Ai, Keping Bi, Jiafeng Guo, and W Bruce Croft. 2018. Learning a deep listwise context model for ranking refinement. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 135–144.
- [2] Areej Alsini, Du Q Huynh, and Amitava Datta. 2020. Hit ratio: An evaluation metric for hashtag recommendation. *arXiv preprint arXiv:2010.01258* (2020).
- [3] Irwan Bello, Sayali Kulkarni, Sagar Jain, Craig Boutilier, Ed Chi, Elad Eban, Xiyang Luo, Alan Mackey, and Ofer Meshi. 2018. Seq2Slate: Re-ranking and slate optimization with RNNs. *arXiv preprint arXiv:1810.02019* (2018).
- [4] Christopher JC Burges. 2010. From ranknet to lambdarank to lambdamart: An overview. *Learning* 11, 23–581 (2010), 81.
- [5] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*. 129–136.
- [6] Jianxin Chang, Chenbin Zhang, Zhiyi Fu, Xiaoxue Zang, Lin Guan, Jing Lu, Yiqun Hui, Dewei Leng, Yanan Niu, Yang Song, et al. 2023. TWIN: Two-stage interest network for lifelong user behavior modeling in CTR prediction at kuaishou. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3785–3794.
- [7] Moses S Charikar. 2002. Similarity estimation techniques from rounding algorithms. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*. 380–388.
- [8] Chi Chen, Hui Chen, Kangzhi Zhao, Junsheng Zhou, Li He, Hongbo Deng, Jian Xu, Bo Zheng, Yong Zhang, and Chunxiao Xing. 2022. EXTR: Click-Through Rate Prediction with Externalities in E-Commerce Sponsored Search. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2732–2740.
- [9] Qiwei Chen, Changhua Pei, Shanshan Lv, Chao Li, Junfeng Ge, and Wenwu Ou. 2021. End-to-end user behavior retrieval in click-through rate prediction model. *arXiv preprint arXiv:2108.04468* (2021).
- [10] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*. 7–10.
- [11] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.
- [12] Yufei Feng, Yu Gong, Fei Sun, Junfeng Ge, and Wenwu Ou. 2021. Revisit recommender system in the permutation perspective. *arXiv preprint arXiv:2102.12057* (2021).
- [13] Yufei Feng, Binbin Hu, Yu Gong, Fei Sun, Qingwen Liu, and Wenwu Ou. 2021. GRN: Generative Rerank Network for Context-wise Recommendation. *arXiv preprint arXiv:2104.00860* (2021).
- [14] Xudong Gong, Qinlin Feng, Yuan Zhang, Jiangling Qin, Weijie Ding, Biao Li, Peng Jiang, and Kun Gai. 2022. Real-time Short Video Recommendation on Mobile Devices. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 3103–3112.
- [15] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: a factorization-machine based neural network for CTR prediction. *arXiv preprint arXiv:1703.04247* (2017).
- [16] Ray Jiang, Sven Gowal, Timothy A Mann, and Danilo J Rezende. 2018. Beyond greedy ranking: Slate optimization via list-CVAE. *arXiv preprint arXiv:1803.01682* (2018).
- [17] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. 2018. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 1754–1763.
- [18] Xiao Lin, Xiaokai Chen, Chenyang Wang, Hantao Shu, Linfeng Song, Biao Li, et al. 2023. Discrete Conditional Diffusion for Reranking in Recommendation. *arXiv preprint arXiv:2308.06982* (2023).
- [19] Shuchang Liu, Qingpeng Cai, Zhankui He, Bowen Sun, Julian McAuley, Dong Zheng, Peng Jiang, and Kun Gai. 2023. Generative flow network for listwise recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1524–1534.
- [20] Gurmeet Singh Manku, Arvind Jain, and Anish Das Sarma. 2007. Detecting near-duplicates for web crawling. In *Proceedings of the 16th international conference on World Wide Web*. 141–150.
- [21] Liang Pang, Jun Xu, Qingyao Ai, Yanyan Lan, Xueqi Cheng, and Jirong Wen. 2020. Setrank: Learning a permutation-invariant ranking model for information retrieval. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*. 499–508.
- [22] Changhua Pei, Yi Zhang, Yongfeng Zhang, Fei Sun, Xiao Lin, Hanxiao Sun, Jian Wu, Peng Jiang, Junfeng Ge, Wenwu Ou, et al. 2019. Personalized re-ranking for recommendation. In *Proceedings of the 13th ACM conference on recommender systems*. 3–11.
- [23] Qi Pi, Guorui Zhou, Yujing Zhang, Zhe Wang, Lejian Ren, Ying Fan, Xiaoqiang Zhu, and Kun Gai. 2020. Search-based user interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2685–2692.
- [24] Yuxin Ren, Qiya Yang, Yichun Wu, Wei Xu, Yalong Wang, and Zhiqiang Zhang. 2024. Non-autoregressive generative models for reranking recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 5625–5634.
- [25] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618* (2012).
- [26] Xiaowen Shi, Fan Yang, Ze Wang, Xiaoxu Wu, Muzhi Guan, Guogang Liao, Wang Yongkang, Xingxing Wang, and Dong Wang. 2023. PIER: Permutation-Level Interest-Based End-to-End Re-ranking Framework in E-commerce. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 4823–4831.
- [27] A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems* (2017).
- [28] Shuli Wang, Xue Wei, Senjie Kou, Chi Wang, Wenshuai Chen, Qi Tang, Yinhua Zhu, Xiong Xiao, and Xingxing Wang. 2025. NLGR: Utilizing Neighbor Lists for Generative Rerank in Personalized Recommendation Systems. *arXiv preprint arXiv:2502.06097* (2025).
- [29] Yunjia Xi, Weiwen Liu, Xinyi Dai, Ruiming Tang, Weinan Zhang, Qing Liu, Xiuqiang He, and Yong Yu. 2021. Context-aware reranking with utility maximization for recommendation. *arXiv preprint arXiv:2110.09059* (2021).
- [30] Yunjia Xi, Weiwen Liu, Jieming Zhu, Xilong Zhao, Xinyi Dai, Ruiming Tang, Weinan Zhang, Rui Zhang, and Yong Yu. 2022. Multi-Level Interaction Reranking with User Behavior History. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1336–1346.
- [31] Guorui Zhou, Na Mou, Ying Fan, Qi Pi, Weijie Bian, Chang Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Deep interest evolution network for click-through rate prediction. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 5941–5948.
- [32] 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.
- [33] Tao Zhuang, Wenwu Ou, and Zhirong Wang. 2018. Globally optimized mutual influence aware ranking in e-commerce search. *arXiv preprint arXiv:1805.08524* (2018).