

跟onerec的区别

主要针对query设计了一些任务

传统的搜索采用多级级联结构英文缩写是MCA: Multi-stage Cascading Architecture

背景

现在的生成式推荐系统面临在电子商务搜索领域面临的挑战：

商品信息冗长

搜索查询与商品之间具有更强的相关性约束。查询通常由 2–3 个简短关键词组成，任何属性上的不匹配都可能引发严重的相关性问题。

挖掘用户的潜在搜索意图也是一项重要挑战。

由此出现的新的解决办法：OneSearch

OneSearch 是一个面向电子商务搜索的端到端生成式框架。

介绍

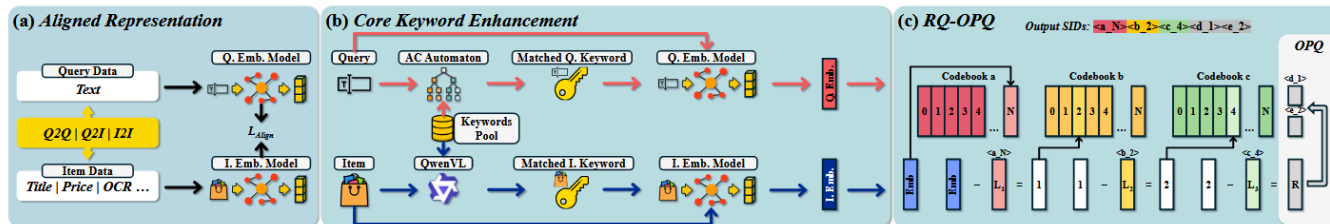
OneSearch 三项关键创新：

- 关键词增强的层次量化编码（KHQE）模块：**在保持强查询-商品相关性约束的前提下，同时保留层次化语义和商品的区分性属性，利用的是RQ-OPQ（残差量化+优化乘积量化）的组合方案。
 - RQ负责层次化语义：第一层编码大类（服装），第二层编码子类（连衣裙），第三层编码细节（颜色、款式）
 - OPQ负责独特特征：保留每个商品的独特属性，避免信息损失
 - 关键词增强：通过NER识别18种结构化属性（品牌、材质、风格等），确保核心属性不被淹没
- 多视角用户行为序列注入策略：**构建基于行为驱动的用户 ID，并融合显式的短期序列（直接作为prompt输入）和隐性的长期序列（注入模型建模），实现用户偏好的全面建模。
- 偏好感知奖励系统（PARS）：**结合多阶段的监督微调和自适应奖励加权排序，以捕捉更加细粒度的用户偏好。

架构

关键词增强的层次量化编码

Keyword-enhanced Hierarchical Quantization Encoding



传统的语义ID编码会把相似商品编成相同ID，导致"红色连衣裙"和"蓝色连衣裙"可能共享同一个编码。

协同语义表示对齐

Aligned collaborative and semantic representation

通过对历史查询-商品交互对的表示进行对齐，将语义知识与协同信号融合在一起。

核心关键词增强

Core Keyword Enhancement

商品文本信息通常存在大量冗余，包含许多无关词甚至相互矛盾的属性。虽然这些属性能提升商品曝光度，但由于堆叠属性过多导致语义顺序极弱，使编码器难以捕捉关键信息。为此提出利用**核心关键词特征**来增强文本表示

首先通过命名实体识别（NER）抽取了 18 类结构化属性（品牌，价格，颜色,...），并从过去一年的点击 query-item 配对中挖掘标注数据。接着，针对每一类属性构建关键词列表，并按照页面访问量从高到低排序，将这些高频关键词选为核心关键词。

RQ-OPQ分层量化标记

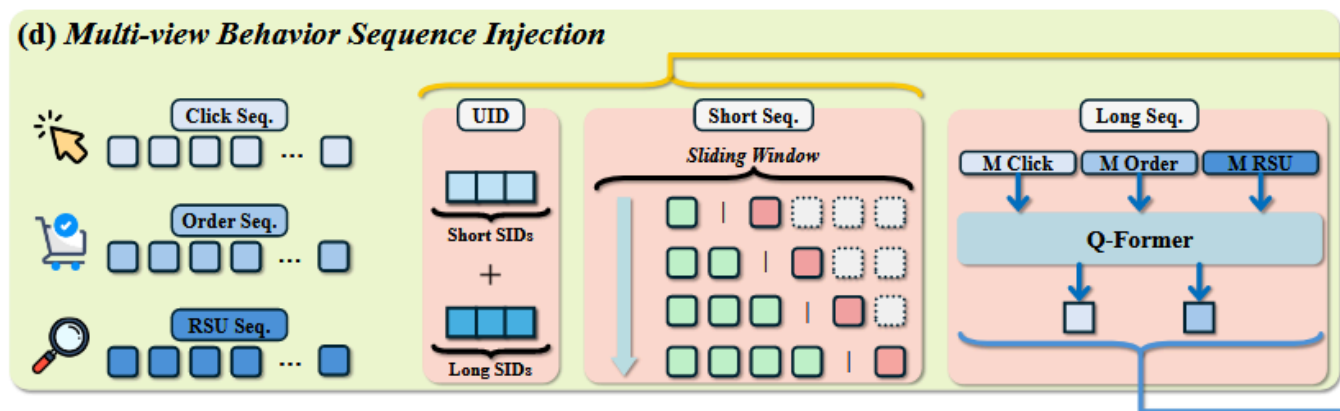
RQ-OPQ Hierarchical Quantization Tokenization

RQ 专注于分层语义：第一层编码大类（服装），第二层编码子类（连衣裙），第三层编码细节（颜色、款式）

OPQ 负责横向特征表示：保留每个商品的独特属性，避免信息损失

多视角用户行为序列注入

Multi-view Behavior Sequence Injection



三个视角：首先提出一种基于行为序列构建的用户 ID 方法，用于生成具有区分性的用户表示；然后，在**提示文本（prompt）**中显式加入短行为序列，以学习用户的近期偏好；并且隐式融入长行为序列，用于建模用户的长期画像

基于行为序列构建的用户 ID

Behavior Sequence Constructed User IDs.

随机分配的 ID 并不能有效表达用户个性化信息，因为固定大小的词表可能会把行为完全不同的用户映射成相同的 ID。因此，提出了一种基于用户行为序列构建的用户 ID 方法，以获得更具区分度的用户表示。

显示短期行为序列

Explicit Short Behavior Sequence

短期行为序列：short behavior sequence 由用户最近点击的商品构成
将这些查询和项目的SID输入到提示符中

隐式长期行为序列

Implicit Long Behavior Sequence.

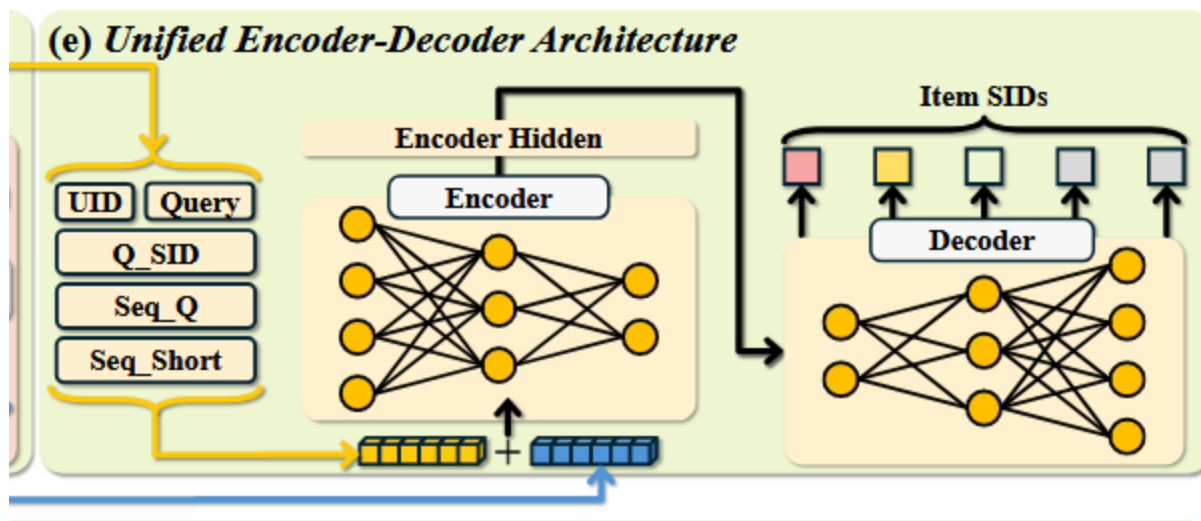
长期行为序列：long behavior sequence 长期行为序列记录用户按时间顺序排列的历史行为
对于新用户或冷启动用户，我们统计

- 每个 query 下最常被点击的商品（query-item 共现次数）
- 按浏览量倒序排序
作为默认行为序列。

比如一个用户最近搜索了"情人节礼物"、"情侣款"，再搜"银戒指"时，模型就能推断出用户可能在找情侣对戒。

统一编码-解码器架构

Unified Encoder-Decoder Architecture



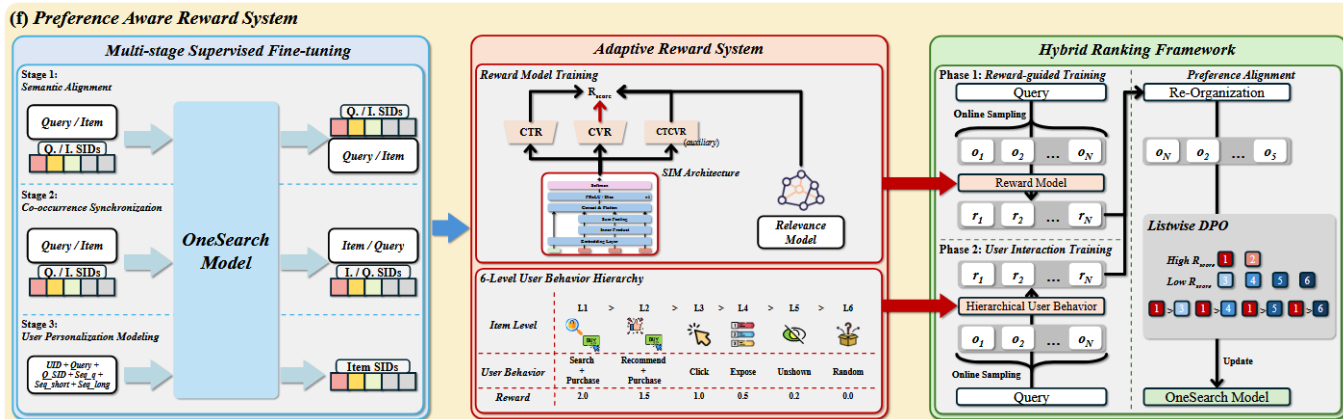
这里与OneRec对比：OneRec使用的是decoder-only，OneSearch使用了统一编码—解码架构，如果采用decoder-only会不会更好呢？

推荐（OneRec、OneSug）用 decoder-only 合适(输入结构复杂)，但搜索场景更难融合结构化特征，所以生成式搜索更不适合 decoder-only。

Encoder-Decoder由用于建模 **用户、查询、行为序列** 的编码器，以及一个用于 **生成商品** 的解码器组成

偏好感知奖励系统

Preference Aware Reward System



对于生成式检索（GR）模型来说，不仅要实现 SID（商品 ID）与查询/商品文本描述之间的语义对齐，还必须基于用户的历史行为序列，直接生成满足查询相关性约束与用户偏好的商品列表。偏好感知奖励系统（preference aware reward system），包括多阶段监督微调（SFT）与自适应奖励系统

多阶段监督微调

Multi-stage Supervised Fine-tuning

语义内容对齐

Semantic Content Alignment

设置三个子任务：

- (a) 输入查询/商品文本，输出对应 SID
- (b) 输入 SID，生成原始查询/商品文本
- (c) 输入查询/商品文本，输出对应的类目信息

其中前两个任务用于实现 $SID \leftrightarrow$ 文本内容 的双向对齐，类目预测用于增强相关性理解, 让模型理解商品ID和文本描述的对应关系

共现同步

Co-occurrence Synchronization

- 查询与商品的相互预测
- 查询 SID 与商品 SID 的相互预测

此阶段不引入用户特征，旨在利用大量在线交互语料学习查询与商品之间的内在语义与协同关系。

用户个性化建模

User Personalization Modeling

引入用户特征，学习个性化偏好

在前两个阶段完成后，引入实际线上预测所需的用户信息，包括：

- 用户 ID
- 查询文本
- 查询 SID (SIDq)
- 搜索查询序列 (Seqq)
- 点击行为序列 (Seqshort)
- 隐式长行为序列 (SeqEmb_long)

自适应奖励系统

自适应加权奖励

Adaptive Reward System

将搜索系统中的用户行为划分为六级：

- L1: 搜索后购买 (权重2.0)
- L2: 推荐场景购买同类商品 (权重1.5)
- L3: 点击 (权重1.0)

- L4: 曝光未点击（权重0.5）
 - L5: 同类目未展示商品（权重0.2）
 - L6: 随机负样本（权重0.0）
- 为每一层赋予不同的基础奖励

奖励模型训练

Reward Model Training

这里我们构建的奖励模型：

- 基于 SIM（Search-based Interest Model）
- **三塔结构：分别学习 CTR、CVR 和 CTCVR**
- 用二元交叉熵训练

CTR: 点击率，用户看到商品后实际点击的比例

CVR: 转化率，点击商品后最终发生购买或转化的比例

CTCVR: 点击到转化率，综合 CTR 与 CVR 的指标

混合排序框架

Hybrid Ranking Framework

收集真实搜索日志中的查询，用奖励模型对 OneSearch 输出结果进行重排序

将排名发生变化的样本用于 **list-wise DPO 训练**

- 升序/点击的商品 → 正样本
- 降序/排名靠后的商品 → 负样本

- 前三类交互 → 正样本
- 后三类交互 → 负样本
- 使用同一损失继续训练