

美团-IntEL：如何通过“用户意图建模”赋能个性化推荐？



SmartMindAI

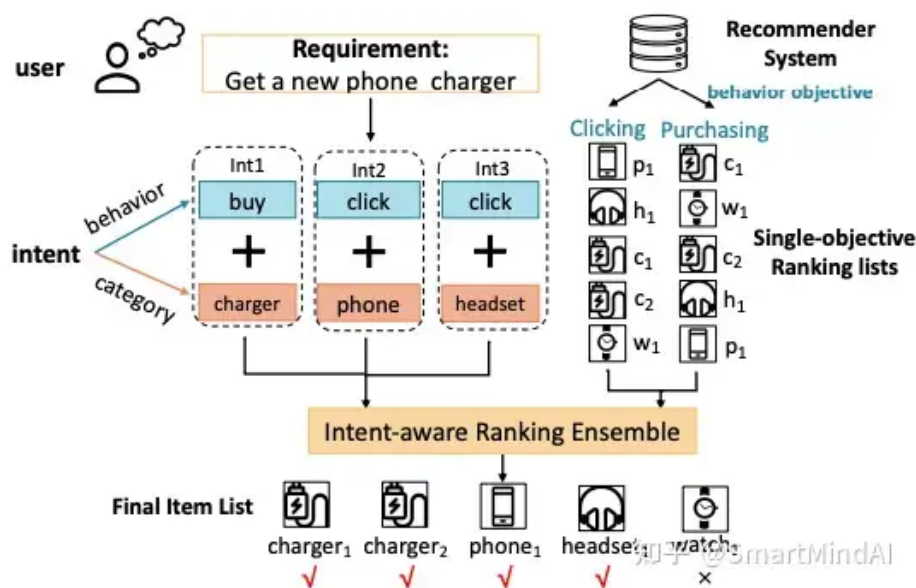
专注搜索、广告、推荐、大模型和人工智能最新技术，欢迎关注

已关注

24 人赞同了该文章

Introduction

用户在使用推荐系统时可能有多种意图，如购买零食、浏览衣服等，这称为用户行为意图和类别意图的组合。现有的推荐模型主要针对单个行为目标，如点击率或转化率等。然而，用户在一个访问中可能会有多种意图，所以重要的是要识别用户的当前意图并将多个单一目标的排序列表聚合起来。为了实现这一目标，可以使用面向意图的排序集合模型。该模型将用户行为意图和类别意图结合起来，将基础排序列表和用户意图相融合，以便更好地满足用户需求。因此，面向意图的排序集合对提高推荐性能非常关键。



本文研究如何将多个单一目标的排序列表组合成个性化的排序集合，同时考虑用户意图。文中提出一种名为Intent-aware Ranking Ensemble Learning (IntEL)的方法，通过预测用户意图并结合基

个单一目标的排序列表以考虑用户意图的方法。

- 这是我们第一次尝试在多类异构项目列表的项目级权重上一般化排序集合学习。在新的设置下，我们理论上证明了排序集合的有效性。
- IntEL是一种新型的意图感知排序集学习模型，能融合多个针对用户意图的单目标推荐列表。其中提出了模糊损失、排序损失和意图损失的集成概念。
- 本研究证实，在多目标情况下，IntEL超越了先前的多任务集成模型。

Related Work

Ranking Ensemble

排序聚合是将多个排序融合为一个共识的过程，已被证明为NP-hard。主要有无监督和监督两种方法，前者如Borda计数和MRA，后者如Outrank的对称相似度和RRA的中心排序距离，以及最近的ERA通过遗传编程进行优化等。这些方法只考虑了项目或分数，而忽视了项目的内容和用户的角色。另一种观点源于集成学习，其基本理论是错误模糊性分解分析，证明了通过组合不同性能的基本模型可以获得更好的性能。这种方法适用于分类和回归等多种任务，并可以在多种损失函数下获得更好的结果。一是基于项目类别和用户行为意向分配项级别的权重，而非计算每个基本模型的一般权重；二是为不同的行为目标生成的异质列表并同时优化多个目标表现。

Multi-Intent Recommendation

本文简要介绍了一种针对[推荐系统](#)⁺中的多意图和多兴趣的新方法。现有研究侧重于捕捉用户在历史交互中的动态意图，如AIR预测用户历史交互中的意图及其迁移。Wang等人（2020）在单个物品级别上模型了用户的动态隐含意图，以捕获物品间的关联性。MIND和ComiRec则采用从历史交互中提取[动态路由](#)⁺的方法来捕获用户的多意图和多样性。TiniRec通过预测用户的多兴趣分布来提炼目标用户的兴趣。通过对比学习，隐含意图表示也被应用于对比损失的约束条件。与先前的研究不同，本研究将意图和兴趣分开处理，并关注于多个行为场景下的意图，同时避免学习每个意图对用户的偏爱，而是利用意图作为融合不同行为目标的指导。

Multi-Objective Recommendation

关于多目标推荐的相关工作有两个研究方向。一组提供多个目标间共享信息的推荐列表，如MMOE和PLE，各列表针对特定目标进行评估。另一组则借助其他目标提高模型性能，如MB-STR预测用户点击偏好。然而，我们并非生成多个列表或指定目标行为，而是融合所有列表，并在该单一列表上评价多个目标。此外，还有一些研究试图同时优化排序准确性和其它目标（如公平性、多样性），但这些研究主要关注维持某一行为的实用性和提升某项指标。相比之下，我们的目标是在提高各项指标的同时提升整个系统的性能。

Preliminaries

Ranking Ensemble Learning Definition

给定条件下的基础模型 k 对于项目 n 的预测得分 $S_n^k(u, c)$ ；学习每个项目的加权集成评分 $S_n^{ens}(u, c)$ 为目标的排序集成学习。

$$S_n^{ens}(u, c) = \sum_{k=1}^K w_k^u(u, c) \cdot S_n^k(u, c)$$

其中 $w_k^u(u, c)$ 表示第 k 个基本模型的权重。权重可以通过辅助信息来学习，例如用户意图和商品类别。 \mathcal{I} 中的项按照 S_n^{ens} 排序，并与排序为

$$\pi^{u,c} = \{\pi_1, \pi_2, \dots, \pi_N\}$$

的预定优先级进行比较，该优先级由业务现实定义，并且不会影响模型的学习策略。排序集合学习的定义与以往的工作类似，但是我们在具有不同目标的不同基本模型上进行集合并使问题更加困难。

定义用户在访问中的意图为一个项目的类别概率分布和行为分布。

$$Int \sim P_{int}(I, B), \sum_{I \times B} P_{int}(I, B) = 1$$

在推荐场景中，类别和行为的类型各异，例如在[在线购物](#)⁺时，商品类用 I 表示，点击和购买等行为用 B 表示；在音乐推荐系统中，音乐流派用 I 表示，听歌和购买专辑等行为用 B 表示。用户意图 Int 则通过用户交互历史和环境上下文来预测。

Ranking Losses

损失定义为： $\mathbf{S}(u, c)$ 其中： $-\pi_1, \pi_2, \dots, \pi_N$ 是用户 u 在会话上下文 c 下的多级实际排序

$$\mathcal{I} = \{i_1, i_2, \dots, i_N\}$$

是一个项目集 $-S_1, S_2, \dots, S_N$ 是用户 u 在会话上下文 c 下的得分列表

π 是一个多级反馈模型，其中MSE作为点对点损失被使用。

$$l_m(\pi, \mathbf{S}) = \frac{1}{N} \sum_{n=1}^N l_m(\pi_n, S_n) := \frac{1}{N} \sum_{n=1}^N (S_n - \pi_n)^2$$

使用一对多损失的贝叶斯个性化排序损失公式：[。](#)根据负采样策略，在每层都随机选择一个正项目和一个负项目配对。

$$l_b(\pi, \mathbf{S}) := \frac{1}{N^+} \sum_{l=1}^L \sum_{n, m \in I_l^+, I_l^-} l_b(S_n, S_m)$$

$$l_b(S_n, S_m) = -\log \sigma(S_n - S_m)$$

N^+ 表示所有的正项物品数量 I_l^+ 和 I_l^- 分别代表第 l 层的正项和一阶负项组 σ 是[sigmoid函数](#)⁺。

使用P-L模型进行排序预测。

$$P_{p-l}(\pi|\mathbf{S}) = \frac{1}{N} \prod_{n=1}^N \frac{\exp(S_{\pi_n})}{\sum_{m=n}^N \exp(S_{\pi_m})}$$

在 π_n 中，定义了排序的第 n 个元素。对应的列表损失函数是。

$$l_{p-l}(\pi, \mathbf{S}) := -\log[P_{p-l}(\pi|\mathbf{S})]$$

Theoretical effectiveness of ranking ensemble learning

证明项级排序集合学习的有效性基于先前的集合学习研究，误差模糊分解（EA）提供了一个对联合损失的上界，用于支持上述证明。基本损失列表

$$\{\ell(\pi, \mathbf{S}^k)\}$$

中的每个元素都被用来计算总损失

$$\ell(\pi, \mathbf{S}^{ens})$$

通过添加一个正的模糊项 A ，并将其作为加权的形式计算基本模型损失（对应于每个 w_k ），我们可以获得此等式。在本文中，我们旨在验证这种分解在具有不同权重的基本列表（即 $w_n^k = w_m^k$ ，对所有 $n = m$ ）的情况下的适用性。为此，我们将尝试证明损失函数可以重新表示为

$$\ell(\pi, \mathbf{S}^{ens}) \leq \sum_k \sum_n w_n^k \ell_n^k(\pi, \mathbf{S}_n^k) - A, \forall w_n^k$$

其中 w_n^k 是点式、[对偶式](#)⁺和列表式的损失函数。

Point-wise Loss

知乎

$$\{S_n^k | k \in \{1, 2, \dots, K\}, n \in \{1, \dots, N\}\}$$

以及一个加权集成模型

$$S_n^{ens} = \sum_{k=1}^K w_n^k S_n^k$$

其中 $w_n^k \geq 0$ 且

$$\sum_{k=1}^K w_n^k = 1$$

第 n 个集成分数 S_n^{ens} 的MSE损失可以分解为两部分

$$l_m(\pi_n, S_n^{ens}) = \sum_{k=1}^K w_n^k l_m(\pi_n, S_n^k) - \sum_{k=1}^K w_n^k A_n^k$$

其中, A_n^k 代表模糊性项 $(S_n^k - S_n^{ens})^2$ 。对每一个基础模型得分 S_n^k , 利用带有拉格朗日型余项的泰勒展开⁺, 将MSE损失 $l_m(\pi_n, S_n^k)$ 在点 S_n^{ens} 附近进行扩展 (若无歧义可省略 π_n)

$$l_m(S_n^k) = l_m(S_n^{ens}) + \frac{\partial l_m(\tilde{S}_n^{ens})}{\partial \tilde{S}_n^{ens}} (S_n^k - S_n^{ens}) + \frac{1}{2!} \frac{\partial^2 l_m(\tilde{S}_n^{ens})}{(\partial \tilde{S}_n^{ens})^2} (S_n^k - S_n^{ens})^2$$

令 \tilde{S}_n^{ens} 为 S_n^{ens} 与 S_n^k 间的插值点, 并定义 A_n^k (公式缺失)。接下来, 我们对所有基本模型的损失进行加权求和表示如下

$$\begin{aligned} \sum_{k=1}^K w_n^k l_m(S_n^k) &= \sum_{k=1}^K [w_n^k l_m(S_n^{ens}) + \frac{\partial l_m(\tilde{S}_n^{ens})}{\partial \tilde{S}_n^{ens}} w_n^k (S_n^k - S_n^{ens}) + w_n^k A_n^k] \\ &= l_m(S_n^{ens}) + \frac{\partial l_m(\tilde{S}_n^{ens})}{\partial \tilde{S}_n^{ens}} \left(\sum_{k=1}^K w_n^k S_n^k - S_n^{ens} \right) + \sum_{k=1}^K w_n^k A_n^k \\ &= l_m(S_n^{ens}) + \sum_{k=1}^K w_n^k A_n^k \end{aligned}$$

第一条等式是因为 $\sum_{k=1}^K w_n^k = 1$ 以及

$$\frac{\partial^2 l_m}{\partial S^2} = 2$$

第二条等式则是基于公式 Eq. **eq : ens**{reference-type="ref" reference="eq:ens"}。

因此,

$$l_m(S_n^{ens}) = \sum_{k=1}^K w_n^k l_m(S_n^k) - \sum_{k=1}^K w_n^k A_n^k$$

证明完毕。

根据 A_n^k 和 w_n^k 的特性, 该方程可表示为EA分解形式。使用点对点损失作为评价标准的排序集合与顶级权重有效性已在理论上得到证明, 这是因为集合损失小于基本模型损失的加权和。为了避免冗长, 我们在以下定理中未提及得分列表和集合公式。

Pair-wise Loss

当满足 $w_n^k \geq 0$, $\sum_{k=1}^K w_n^k = 1$ 以及

$$|w_m^k - w_n^k| \leq \delta, \forall m, n$$

一对集成得分 S_n^{ens} 和 S_m^{ens} 的BPR损失可以分解为:

$$l_b(S_n^{ens}, S_m^{ens}) < \sum_{k=1}^K w_n^k l_b(S_n^k, S_m^k) + \delta \sum_{k=1}^K S_m^k - \sum_{k=1}^K w_n^k A_{nm}^k$$

其中 A_{nm}^k 表示由基础模型⁺生成的得分的模糊性

$$A_{nm}^k = \sigma(\tilde{z}^{ens})(1 - \sigma(\tilde{z}^{ens})) \sum_{k=1}^K w_n^k (z_{nm}^k - z_{nm}^{ens})^2$$

知乎

$$\text{令 } z_{nm}^k = S_n^k - S_m^k,$$

$$l_b(z_{nm}^*) = l_b(S_n^*, S_m^*)$$

在公式中, 我们通过泰勒展开将 $l_b(z_{nm}^k)$ 在 z_{nm}^{ens} 附近展开

$$\begin{aligned} l_b(z_{nm}^k) &= l_b(z_{nm}^{ens}) + \frac{\partial l_b(z_{nm}^{ens})}{\partial z_{nm}^{ens}} (z_{nm}^k - z_{nm}^{ens}) + \frac{1}{2!} \frac{\partial^2 l_b(z_{nm}^{ens})}{(\partial z_{nm}^{ens})^2} \\ &:= l_b(z_{nm}^{ens}) - B_{nm}^k + A_{nm}^k \end{aligned}$$

其中 z_{nm}^{ens} 是 z_{nm}^{ens} 和 z_{nm}^k 之间的插值点。在满足条件 $\sum_{k=1}^K w_n^k = 1$ 及 $|w_n^k - w_m^k| \leq \delta$ 的前提下, 对 B_{nm}^k 的加权求和有界为

$$\begin{aligned} \sum_{k=1}^K w_n^k B_{nm}^k &= [1 - \sigma(z_{nm}^{ens})] \sum_{k=1}^K w_n^k (z_{nm}^k - z_{nm}^{ens}) \\ &\leq \sum_{k=1}^K |w_n^k - w_m^k| |S_m^k| \leq \sigma \sum_{k=1}^K S_m^k \end{aligned}$$

对公式两边进行加权求和, 我们得到

$$l_b(z_{nm}^{ens}) < \sum_{k=1}^K w_n^k l_b(z_{nm}^k) + \sigma \sum_{k=1}^K S_m^k - \sum_{k=1}^K w_n^k A_{nm}^k$$

证明完毕。

对于范围限制的 w_n^k , 我们有 $\delta \leq 1$ 的条件。此外, 在对称损失下, 顺序的重要性大于得分值。因此, 第二个项可以在

$$(\delta \sum_{k=1}^K S_m^k < \sum_{k=1}^K S_m^k)$$

中尽可能地小。另外, 模糊性的 A_{nm}^k 是半正定的。因此, 我们的算法遵循EA分解的形式, 并且我们采用的是有效的理论上的对称损失排序集合方法。

List-wise Loss

当满足 $w_n^k \geq 0$ 、 $\sum_{k=1}^K w_n^k = 1$ 及 $|w_n^k - w_m^k| \leq \delta$ 对任意的 m 和 n 时, 排序集合 (按 π 排序) 的集成得分

$$\mathbf{S}^{ens} = \{S_1^{ens}, S_2^{ens}, \dots, S_n^{ens}\}$$

的列表式损失可以分解为:

$$l_{p-l}(\pi, \mathbf{S}^{ens}) < \sum_{k=1}^K w_{\max}^k l_{p-l}(\pi, \mathbf{S}^k) + \delta N S_{\text{sum}}^{\max} - \sum_{k=1}^K \sum_{n=1}^N w_n^k A_n^k$$

其中, w_{\max}^k 表示列表 k 中所有权重的最大值, A_n^k 表示位置 n 上的模糊性, 其定义为:

$$A_n^k = \frac{\left[\sum_{m=n+1}^N \exp(-z_{nm}^{ens}) (z_{nm}^k - z_{nm}^{ens}) \right]^2}{\left(1 + \sum_{m=n+1}^N \exp(z_{nm}^{ens}) \right)^2}$$

S_{sum}^{\max} 被定义为:

$$S_{\text{sum}}^{\max} = \max_{m=1}^N \sum_{k=1}^K S_m^k$$

z_{nm}^* 表示两个评分之间的差异。

我们定义分数差分向量:

$$z_{n:N} = [z_{n+1}, \dots, z_N] = [S_n - S_{n+1}, \dots, S_n - S_N]$$

知乎

$$g_n(z_{n:N}) = \log\left(1 + \sum_{m=n+1}^N \exp(-z_{nm})\right)$$

对于每个基础模型给出的一组物品得分

$$\mathbf{S}^k = \{S_n^k | n \in \{1, 2, \dots, N\}\}$$

其PL损失为：

$$l_{p-l}(\mathbf{S}^k) = \sum_{n=1}^N g_n(z_{n:N}^k)$$

我们通过带有拉格朗日型余项的泰勒展开将 $g_n(z_{n:N}^k)$ 在点 $z_{n:N}^{ens}$ 处展开得到：

$$\begin{aligned} \sum_{k=1}^K w_n^k B_n^k &= \frac{\sum_{m=n+1}^N \exp(-z_{nm}^{ens}) \sum_{k=1}^K w_n^k [z_{nm}^k - z_{nm}^{ens}]}{1 + \sum_{m=n+1}^N \exp(-z_{nm}^{ens})} \\ &< \delta \cdot S_{sum}^{max} \end{aligned}$$

故将上式从 $n = 1$ 累积到 $n = N$ ，可得

$$l_{p-l}(\mathbf{S}^{ens}) < \sum_{k=1}^K w_{max}^k l_{p-l}(\mathbf{S}^k) + \delta N S_{sum}^{max} - \sum_{n=1}^N \sum_{k=1}^K w_n^k A_n^k$$

Proof done S_n^{ens} 最小可任意小。 A_n^k 为半正数，故满足EA分解理论。排序集合作为列表式损失有效。

Ensemble Loss for Model Training

定理证明了我们的排序集合学习方法的有效性。使用EA分解理论，我们推导出：当列表索引为 π 时，基础列表的总损失小于任何权重求和组合的总损失：

$$l_{ens}(\pi, \mathbf{S}^{ens}) \leq \sum_k w_k l_k(\pi, \mathbf{S}^k) - A + \Delta, \forall w_k \leq 0, \sum_{k=1}^K w_k = 1$$

这个定理表明，我们可以通过选择较小的集合损失函数和较大的模糊项来优化排序集合学习的任务。此外，随着模糊项的增加，我们可以使集合损失函数变得更加小。

$$l_{el} = l_{ens}(\pi, \mathbf{S}^{ens}) - \alpha A$$

哪种 l_{ens} 可以插入 A ？可能是 l_m 、 l_b 或 l_{p-l} 中的哪一个。 A 表示歧义项。对于BPR和P-L损失，存在一个插值

$$\tilde{S}_n^{ens} = S_n^{ens} + \theta(S_n^k - S_n^{ens})$$

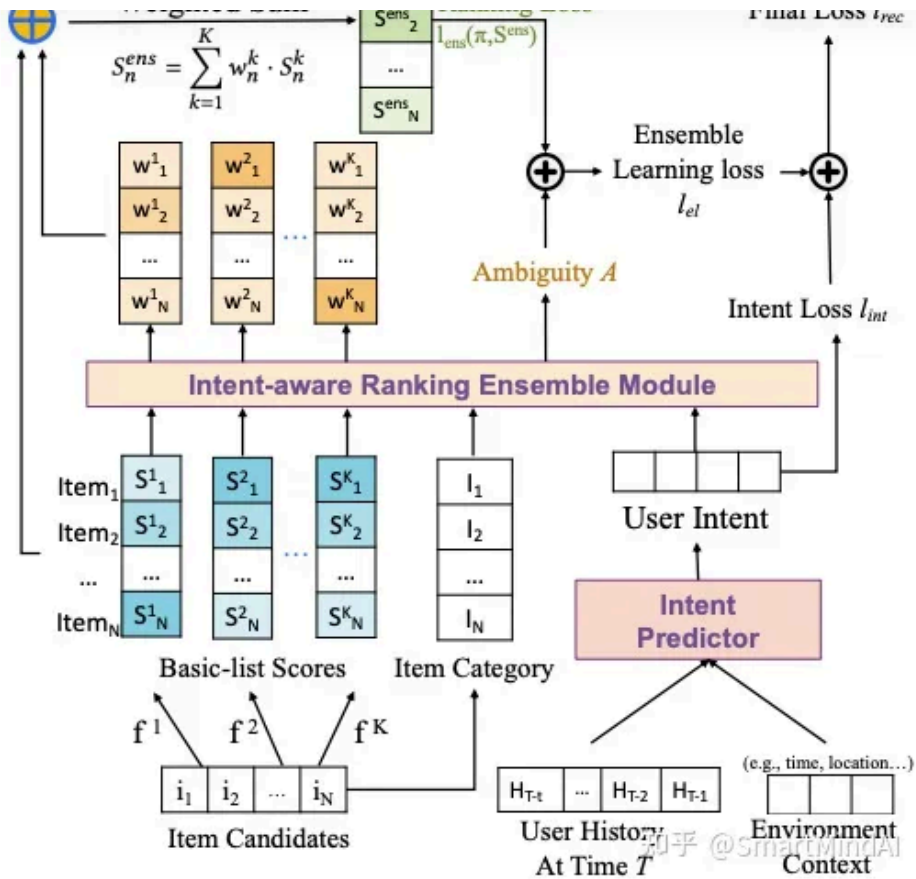
为了简化计算，我们可以假设 $\theta \rightarrow 0$ 而不失一般性。

Intent-aware Ranking Ensemble Method

Overall Framework

利用不同损失函数对排序组合进行权重的有效性已被证实。为了学习这些权重，我们需要设计一个神经网络。在图示中，用户意图通过行为和类别帮助聚合基本列表。然而，意图无法预先获取。为此，我们设计了一个意图预测器和意图感知的排序组合模块。我们的IntEL方法的主要框架如下：首先，通过预测用户的意图和从历史交互和当前环境中获取的意图感知模块，可以计算出每个项目的级别权重。最后，使用线性序列模型和综合得分优化不同的损失函数：排序损失，模糊性损失和意图预测损失。

知乎



User Intent Predictor

用户意图由多维概率分布 Int 定义，该分布模拟了用户在每个访问期间可能的行为和类别。预测器的目标是为每个用户访问生成意图概率分布。用户的历史交互和当前环境被用来预测意图。时间窗口内的用户交互历史和环境上下文（如时间戳和位置）用于预测在特定时间点 T 的意图，其中 t 是一个预定义的时间窗口。环境上下文通过线性环境编码器以向量形式表示。用户意图和上下文通过两个序列编码器分别嵌入，然后将两个嵌入向量组合并用一个序列编码器进行编码，形成会话级别嵌入矢量 $h_s(u, T)$ 。商品级别的“意图”被等效于会话级别的意图编码器进行编码，并且输入到一个序列编码器中，形成商品级别的历史 $h_i(u, T)$ 。可以使用任何序列模型，例如GRU、Transformer等。最后，环境上下文 $c(u, T)$ 、会话级别 $h_s(u, T)$ 和商品级别 $h_i(u, T)$ 被串联起来，通过一个线性层进行预测，得到意图 $\hat{\text{Int}}(u, T)$ 。

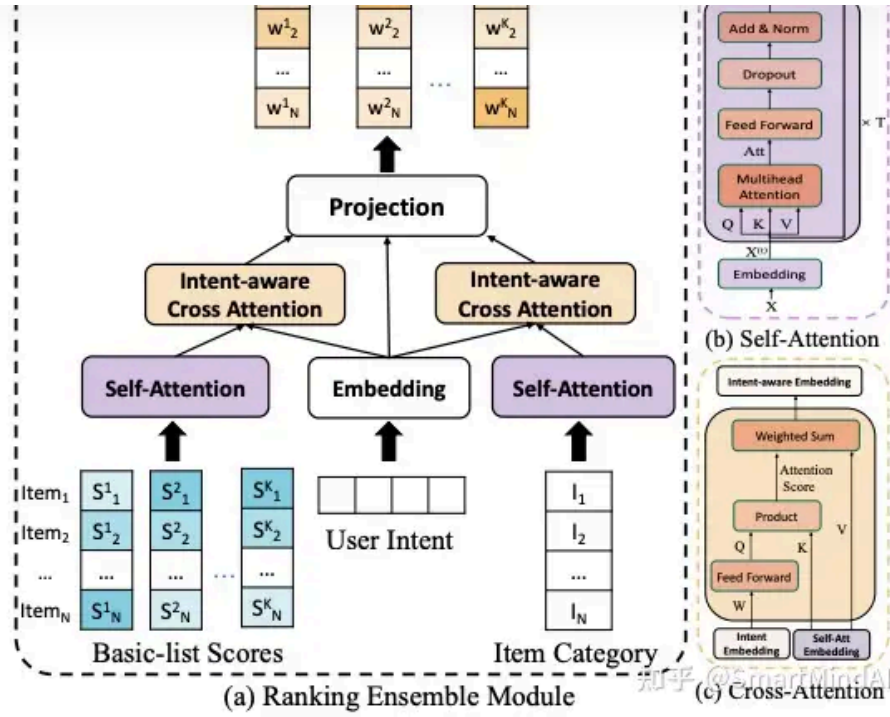
$$\hat{\text{Int}} = \text{Softmax}(\mathbf{W}^I[c, h_s, h_i] + b^I)$$

在WI和BI是线性映射+参数时。

Design of Ensemble Module

首先，将每个基本列表的项得分向量馈送到自注意力层，以表示同一基本列表中项得分之间的关系。此外，还使用自注意力层对项类别进行编码，以捕捉同一基本列表内部类+分布。

知乎



自注意力结构由将评分（或类别）嵌入到 d_e 维表示的线性层组成，该表示为

$$\mathbf{S} \in \mathbb{R}^{N \times d_e}$$

（或

$$\mathbf{I} \in \mathbb{R}^{N \times d_e}$$

），并且包含 T 个多头注意力层，这些注意力层遵循由@wang2020toward提出的交叉关系注意力层。然后，用户意图通过线性投影映射到维度为 d_{int} 的向量

$$\mathbf{Int}_d = \mathbf{W}^i \mathbf{Int} \in \mathbb{R}^{N \times d_{int}}$$

最后，使用交叉注意力层获取用户意图对分数和特征的影响。

$$\mathbf{A}_s = \text{Attention}(Q = \mathbf{W}^Q \mathbf{Int}_d, K = \mathbf{S}, V = \mathbf{S})$$

$$\mathbf{A}_i = \text{Attention}(Q = \mathbf{W}^Q \mathbf{Int}_d, K = \mathbf{I}, V = \mathbf{I})$$

在一个包含投影矩阵

$$\mathbf{W}_Q \in \mathbb{R}^{d_e \times d_{int}}$$

的两个意图感知模块间，通过使用全局用户意图来指导基本列表得分和项目类别的聚合，而非分别处理用户和推荐器交互时的行为意图和类别意图。最终生成所有信息的权重 $\{w_n^k\}$ 。接着，使用意图感知得分嵌入 \mathbf{A}_s ，意图感知项目类别嵌入 \mathbf{A}_i ，以及意图向量 \mathbf{Int}_d 被连接并投影到 \mathbb{R}^K 空间中的权重矩阵

$$\mathbf{W} \in \mathbb{R}^{N \times K}$$

$$\mathbf{W} = \mathbf{W}^w \cdot ([\mathbf{A}_s, \mathbf{A}_i, \mathbf{Int}_d])$$

投影矩阵 \mathbf{W}^w 大小为 $K \times (2d_e + d_{int})$ ，用于训练。输出矩阵 \mathbf{W} 是基本模型分数的加权和，其中权重为 w_n^k 。

Model Learning Strategy

端到端框架通过联合学习两个模块优化模型：意图预测模块和意图感知排序模块。EA分解优化了排序集合的结果。排序学习损失由排序元组和排序集构成，其中包括



和 A 。真实用户意图引导了排序集合的生成，所以也用意图预测损失进行训练。用户意图的多维分布被用KL-散度损失衡量与预测意图的距离。最后推荐损失是加权和。

$$l_{rec} = l_{el} + \gamma l_{int} = l_{ens}(\pi, \mathbf{S}^{ens}) - \alpha A + \gamma l_{int}$$

$$l_{ens}(\pi, \mathbf{S}^{ens})$$

、 A 、 l_{int} 分别表示排序集合损失、模糊项和意图预测损失。 α 和 γ 是调节权重的超参数。

Experiments

Experimental Setup

Dataset

实验在天猫和LifeData两个数据集上进行。

本文研究了天猫和LifeData两个多行为数据集。在天猫数据集中，用户的点击、收藏和购买交互被用于集成学习和基本模型生成。而在LifeData数据集中，用户的点击和购买行为被用于会话级别的排序融合和基本模型优化。

- 一、方法：
- 1.天猫数据集：-排除物品和用户少于3个正向交互的情况。-将用户的交互视为一个会话。-在前三个星期生成基本模型分数。
 - 2.LifeData数据集：-从一组匿名用户中随机抽取一月的交互数据。-过滤掉用户和物品的交互次数小于3次的情况。-每个用户的每次访问定义为一个会话。-按会话级别进行排序融合。

Dataset	#Intent	Avg. Session Length	Avg. Pos./Session
Tmall	1,071	68.37	3.73
LifeData	12	32.78	1.47

二、结果： 1. 集成学习数据集中，只有不包含用于基本模型生成的数据。基本模型训练数据与集成学习数据无交集。 2. 表示性分析：-所有数据都包括行为特征、用户特征和物品特征。

三、讨论： 1. 基本模型生成阶段将使用九月份收集的天猫数据集。 2. 排序融合将考虑大部分历史会话中的20个交互作为意图预测器。

四、结论： 通过上述方法，我们可以有效地处理天猫和LifeData数据集，并在后续的研究中使用它们。[*subsubsec : base_model*] 京东使用固定的积分排序集合， 淘宝采用DeepFM模型训练，每个会话选择三个模型预测的前30项，然后组合成基本项列表进行重排序。LifeData提供两个基本分数列表，分别是点击概率和购买概率排序后的类别。淘宝Intent维度为1071，LifeData为12。Ground Truth Int的概率从所有正交互计算得出。

Baseline Methods

单击 (Single Click)、随机化随机方法 (Randomized Randomization Approach)、博尔达计分法 (Borda Voting Method)、基于 λ 优化的排名学习(λ Rank), [进化算法](#)+排序方法 (ERA), 以及自适应加权集成学习Lambda投票法 (aWELv) 和其扩展版 aWELv+Intention/IntentEL。其中，IntentEL是一种群体模型，而aWELv是一种针对个性化需求的排序集合技术。

Experimental settings

数据集按时间切分，分为训练集和验证集⁺。多级ground truth π 的优先顺序为Tmall购买 > 喜爱 > 点击 > 查看和LifeData购买 > 点击 > 查看。评估指标为NDCG@3、5和10，评估对象为ensemble列表 \mathbf{S}^{ens} 在所有行为目标上的multi-level ground truth π 。

四部分评估基本模型得分，无个人化基准模型集，个性化基线：aWELv及其两个变体用户意图，IntEL三个损失函数。观察结果。最后，awelv在LifeData上的表现优于Tmall，因为在Tmall的会话列表更长。awelv忽略了有用的内部列表信息。所以，商品类别意图的项级别权重是有必要的。awelv在两个数据集上优于基本模型，且awelv+Int/IntEL在大多数指标上优于awelv，说明用户意图有助于排序联合学习。

Table 4: Results of IntEL with three different loss functions and baseline methods on Tmall. Boldface shows the best result. Underline indicates the best baseline. Notation **/* demonstrates significantly better than the best baseline with p<0.05/0.01.

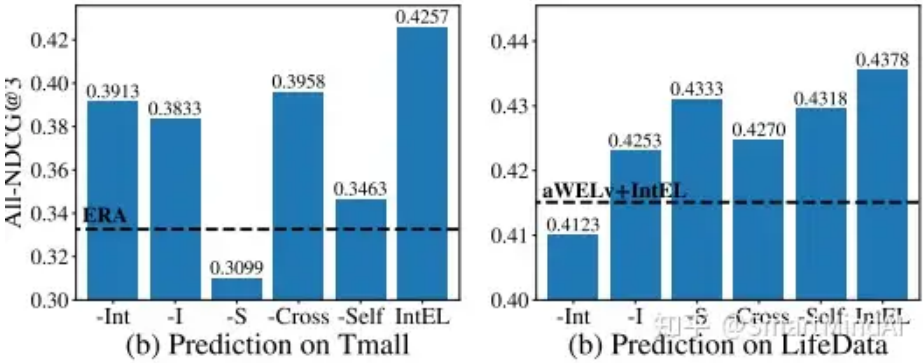
Model	All-NDCCG@K			Click-NDCCG@K			Fav-NDCCG@K			Buy-NDCCG@K		
	K=3	K=5	K=10	K=3	K=5	K=10	K=3	K=5	K=10	K=3	K=5	K=10
Single Click	0.1356	0.1473	0.1673	0.1435	0.1532	0.1721	0.0701	0.0829	0.1014	0.0918	0.1057	0.1243
Single Fav.	0.0752	0.0874	0.1066	0.0779	0.0894	0.1083	0.0630	0.0748	0.0920	0.0492	0.0607	0.0765
Single Buy	0.0572	0.0689	0.087	0.0587	0.0699	0.0878	0.0393	0.0489	0.0638	0.0632	0.0776	0.0974
RRA	0.0960	0.1093	0.1317	0.0998	0.1120	0.1341	0.0683	0.0813	0.1014	0.0753	0.0907	0.1122
Borda	0.1258	0.1398	0.1626	0.1317	0.1440	0.1660	0.0741	0.0880	0.1081	0.0830	0.0989	0.1218
λ Rank	0.2742	0.2797	0.3003	<u>0.3104</u>	<u>0.3064</u>	<u>0.3189</u>	0.1878	0.2122	0.2472	0.1376	0.1586	0.1913
ERA	0.3325	0.3378	0.3623	0.2301	0.2420	0.2716	0.1933	0.2156	0.2502	0.1921	0.2163	0.2504
aWELv	0.1387	0.1533	0.1770	0.1469	0.1584	0.1811	0.0837	0.0986	0.1197	0.1025	0.1198	0.1436
aWELv+Int	0.1398	0.1574	0.1784	0.1484	0.1592	0.1822	0.0903	0.1016	0.1183	0.1030	0.1259	0.1445
aWELv+IntEL	0.1427	0.1556	0.1774	0.1535	0.1620	0.1821	0.0906	0.0934	0.1120	0.1042	0.1263	0.1451
IntEL-MSE	0.4257**	0.4364**	0.4676**	0.4693**	0.4680**	0.4712**	0.2943**	0.3271**	0.3721**	0.2433*	0.2760**	0.3100**
IntEL-BPR	0.3992*	0.3859*	0.3755	0.4417**	0.4157**	0.3960*	0.2791**	0.2943**	0.3168*	0.2310*	0.2510*	0.2740
IntEL-PL	0.4041**	0.3865*	0.3678	0.4367**	0.4060**	0.3829**	0.2811**	0.2934**	0.3032*	0.2355*	0.2472*	0.2594

Table 5: Results of IntEL with three different loss functions and baseline methods on LifeData. Boldface shows the best result. Underline indicates the best baseline. Notation **/* demonstrates significantly better than the best baseline with p<0.05/0.01.

Model	All-NDCCG@K			Click-NDCCG@K			Buy-NDCCG@K		
	K=3	K=5	K=10	K=3	K=5	K=10	K=3	K=5	K=10
Single Click	0.4004	0.4443	0.4972	0.4009	0.4449	0.4980	0.6365	0.6665	0.6918
Single Buy	0.3102	0.3526	0.4070	0.3102	0.3528	0.4072	0.6893	0.7211	0.7438
RRA	0.3539	0.4020	0.4586	0.3540	0.4022	0.4590	0.6556	0.6865	0.7104
Borda	0.4094	0.4538	0.5061	0.4097	0.4541	0.5066	0.7030	0.7250	0.7447
λ Rank	0.4129	0.4487	0.4830	0.4133	0.4492	0.4835	0.6866	0.7083	0.7225
ERA	0.4063	0.4451	0.5053	0.4181	0.4579	0.5112	0.5782	0.6307	0.6764
aWELv	0.4074	0.4531	0.5033	0.4077	0.4535	0.5041	<u>0.7063</u>	<u>0.7339</u>	0.7466
aWELv+Int	0.4150	0.4607	0.5143	0.4151	0.4610	0.5147	0.6962	0.7271	0.7482
aWELv+IntEL	<u>0.4174</u>	<u>0.4663</u>	<u>0.5176</u>	<u>0.4189</u>	<u>0.4638</u>	<u>0.5171</u>	0.7036	0.7318	<u>0.7503</u>
IntEL-MSE	0.4253**	0.4695*	0.5211**	0.4257**	0.4700**	0.5217*	0.7096	0.7379	0.7498
IntEL-BPR	0.4308**	0.4752**	0.5268**	0.4312**	0.4757**	0.5275**	0.7115**	0.7530**	0.7619**
IntEL-PL	0.4378**	0.4819**	0.5332**	0.4382**	0.4825**	0.5339**	0.7093	0.7382*	0.7604**

Further Analysis

本文探讨了在Tmall和LifeData两个数据集上应用排序联合学习方法，并对其性能进行了切片研究、用户意图分析和超参数分析。结果表明，IntEL-MSE和IntEL-PL是最佳模型。



Ablation Study

IntEL的主要贡献是使用用户意图进行异构排序集和整合项目类别和用户意图。我们对比了IntEL和五个变体：-Int（无意图）、-I（无项目类别）和-S（无基本列表分数）。我们还进行了两种变体的测试：-Cross（去除带意图意识的交叉注意力层）和-Self（将自注意力层替换为直接连接）。结果显示，在一般多级排序列表上，IntEL及其变体的NDCG@3有所下降，这表明所有输入和两个注意力层都对IntEL的性能改进有贡献。从Tmall和LifeData中删除分数和自我注意力层会导致显著的性能降低，显示出基本列表内部信息的重要性。从LifeData中删除用户意图会导致最严重的退化，说明采用用户意图对于多目标排序集合非常重要。然而，删除变体仍优于所有基本列表，这与我们的损失减少通过EA模糊分解证明一致。

Influence of User Intent

知乎

我们研究了损失函数 l_{rec} 对于我们的方法的重要性。为此，我们分析了优化过程中超参数的影响。其中涉及两个超参数： α 表示基本列表歧义 A 的权重 γ 表示意图预测损失 l_{int} 的权重。通过展示NDCG@3在不同超参数下的结果，我们可以看到 α 过大或过小都会导致排序集合性能下降。特别是在 α 大于一定值时，模型过于关注最大化基本模型歧义 A 来最小化 l_{rec} ，而忽视了学习损失 l_{el} 。而且，如果 γ 过高或过低，可能会导致意图预测困难，从而影响整体模型性能。总体来看，调整这两个超参数是非常重要的。在Tmall数据集上 γ 的性能有所波动；而在LifeData数据集上，它的性能相对稳定。这是因为Tmall包含了大量的意图类型（1071种），使得预测器的优化变得更为复杂，因此需要适当的 γ 权重来确保准确性。而相比之下，LifeData只包含12种意图，因此对于意图预测优化来说，较小的 γ 是必需的。

Conclusion

本文介绍了一种新的意向关注单目标排序列表聚合方法IntEL，它是首次将排序集成学习与项级权重扩展到异构项列表上，同时也首次将用户意图整合到推荐中的排序聚合中。IntEL通过排序集成学习和项级权重扩展，最小化排序集合损失并最大化模糊性。为了以用户意图为导向，IntEL设计了一个以用户意图为导向的排序集学习模型。在两个大型数据集上的实验表明，IntEL在多个优化目标的同时获得了显著改善。虽然该方法有局限性，但可以通过多行为方法*和更准确的用户意图预测等方法改进。

原文《Intent-aware Ranking Ensemble for Personalized Recommendation》

发布于 2024-01-29 15:14 · IP 属地北京

美团 个性化推荐 意图识别

▲ 赞同 24 ▼ ● 添加评论 ↗ 分享 ❤ 喜欢 ★ 收藏 📄 申请转载 ...



理性发言，友善互动



还没有评论，发表第一个评论吧

推荐阅读



APP和服务端-架构设计(二)

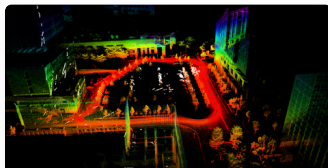
J Strive、允

B 端设计指南 - 帮助体系

“请问你需要帮助吗？”在生活中，我们有任何问题都会请求别人的帮助；而在屏幕世界中，使用不同的软件我们也需要寻求帮助。使用Notion，我会通过B站的视频去寻找教程；管理待办事项，...

CE青年

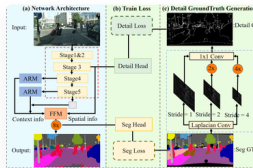
发表于B端设计指...



二.全局定位--开源定位框架livox-relocalization实录数...

goldq...

发表于多传感器融...



美团学者对经典分割架构BiSeNet的Rethinking

我爱计算机...

发表于C