

# 内生解耦去偏模型tuning记录2

## 目标

解决搜推场景有偏样本引起的模型学习有偏问题，特别针对原子特征效应，提出内生解耦，更好学习side info缓解有偏，尝试在召回，个性化序列建模，迁移学习等场景做验证

## 迭代版本记录

模型初版设计见[文档](#)

实验配置：item端，atom特征：bizId, side特征：item意图

## 自监督学习表征

目标1：自监督学习出atom repr，要求：

- a. 不同side repr uniform distributed
- b. 同一side对比atom repr整体聚集在对应side repr周围，且atom repr互相uniform distributed

版本号	problem	action	work	result
1	side ssl loss下降不显著	在计算side自监督loss时，mask掉similarity异常大（大于0.99，fake negative）的in batch负样本	yes	side自监督loss从1.13降低到0.55
2	atom自监督loss依然比较大，1.2左右，怀疑主要是因为projection mlp共享原因	projection mlp不再共享	partially	atom自监督loss从1.2降低到1.0257859，结果不是很显著
3	怀疑stage1_loss权重过大，stage1_loss和atom自监督loss互为跷跷板	stage1_loss权重从1.0 变为0.25	yes	1. atom自监督loss从1.2降低到0.7455813，结果可以接受 2. stage1_loss从0.05增长到0.16009676，符合预期
4		去除低频特征过滤	not	还是存在item atom表征和side表征cosine similarity分数小

	低频item未被充分学习，造成这部分item stage1_loss 较大			
5	没有对atom 表征和side 表征的关系做强约束	1. 引入非对称损失，变更 stage1_loss 为 truncate cosine loss, $\max(0, \text{margin} - \text{cosine\_sim}(\text{atom\_repr}, \text{side\_repr}))^2$ , margin 取 0.5 2. atom ssl loss 变更 atom_orthogonal ssl loss, 因为 stage1 变更后, side 对应 atom repr 大部分在 side repr 周围, 再直接要求 atom repr 在全空间均匀分布比较困难, 变更为 atom_orthogonal repr 是比较自然的	yes	在两个item意图下表现都较好
6		进一步验证query侧效果	partially	stage1_loss 和 side_ssl_loss 符合预期, atom_ssl_loss 较高且 batch 不 aogeny_decomposition_query_endogeny_decomposition-ia [worker-0] [2022-12-16 13:34:05,902] [INFO] [basic_hooks] ss=0.5245103240013123 config_dipn_assembler_loss=0.1 <b>endogeny_decomposition_atom_repr_ssl_loss</b> =[[1.1240 _cosine_sim_side=[0.999999 0.99999905 0.999999 0.999 o_loss=[[9.997383]] quality_query_endogeny_decomposition ogeny_decomposition_query_endogeny_decomposition-lab [worker-0] [2022-12-16 13:35:16,653] [INFO] [basic_hooks]
7	query 密度比 item 高, in batch 内存在更多	atom 侧引入 fake negative mask	work	1 query atom_ssl_loss 下降到 0.7 左右 2 比较有趣的是, side_ssl_loss 同样有一定小幅度下降

fake negative			rain'  1 1 composition_cosine_sim_side_min=[[0.04763628]] quality_q ple_weight_sum=[[512.]] quality_query_endogeny_decompo 2.]] <b>quality_query_endogeny_decomposition_atom repr</b> age1_loss=[[0.03977511]] quality_query_endogeny_decompo
------------------	--	--	---

## 对齐q-i两边表征

目标2：利用监督信号，对齐q-i两边表征，验证/实验以下假设

- a. a property is expected that difference between atom repr and side repr can be used to replace ipw weight

版本号	problem	action	work /hold	result
1	假设：side表征在向量空间更靠近频次大的atom表征	在item侧做验证 1) 一个意图下二者的关系 2) 全局规一化之后的关系	yes	单意图和全局都是正相关性，说明高频atom特征确实影响了side特征表征 1) 2)
2	根据上面1的分析，能否利用 side_repr和 atom_repr_orthogonal的相似度替代ipw，越相似则权重越小，这里有几问题要考虑：  1 在stage1_loss已经约束二者相似度要在margin (0.5) 范围内，可能需要一种手段做归一化， batch norm?  2 cosine的范围是-1到1，需要需要归一化到0到1之间，且和atom的频次能联系起来，额外做个回归?  3 权重 (ipw) 非静态，loss传导梯度时候同样也会影响，要不要做rank loss的梯度阻	1 近似 ipw weight 做 batch norm  2 近似 ipw weight 简单线性归一化，暂时不施加非线性变换  3 近似 ipw weight 对 rank loss做	partially	item ssl loss较大，其他在loss可接受范围内，原因可能是在rank层，同一个search id下side的区分度没有atom大，例如query=“租手机”，下面item侧大部分side都是租赁相关性的  0.9999989 0.9999989 0.9999989 0.9999989 0.9999989] quality_ite 9]] <b>quality_item_side_repr_ssl_loss</b> =[[1.1463524]] quality_ite git=[-2.0161653 -2.674533 -2.6631305 2.7902017 -2.950419 y_decomposition_similarity-sample_weights=[1. 1. 1. 1. 1.] quali 6 0.17049667 0.4118396 0.17049667 0.17049667] quality_qi_ 7874 ] quality_decomposition_similarity-label_input=[[1.]  1 1 y_item_item-sample_weights=[1. 1. 1. 1. 1.] quality_item_item-m 0.9999989 0.9999989 0.9999989 0.9999989 0.9999989] quali 72]] <b>quality_item_side_repr_ssl_loss</b> =[[1.1106796]] quality_it ogit=[-2.2108662 -3.172333 -2.4784548 -2.7587478 -3.1117 ty_decomposition_similarity-sample_weights=[1. 1. 1. 1. 1.] qual 13 0.18297952 0.22662267 0.18297952 0.18314886] quality_ 0846213] quality_decomposition_similarity-label_input=[[1.]  根据case看，side相似度学习的还不错

	断，同时做 batch norm?  4 rank部分要不要额外接一个mlp，避免和ssl目标冲突引起的梯度冲突  最后一个问题是怎么验证ipw影响	梯度阻隔  4 rank模块额外增加一个mlp		
3				

## 模型对比

数据集1: 支付宝主搜意图搜类query启发式分档, bpr loss

	ndcg	bpr
双塔 (concat (side_emb + atom_emb) -> mlp)	0.9282	
双塔 (concat ((side_emb-> mlp) , (atom_emb-> mlp))	0.9325 (+0.0043)	0.2280
内生解耦	<b>0.9439 (+0.114)</b>	<b>0.0402</b>

数据集2: movielen-100m, rating >= 3, label=2, rating < 3, label=1, 全局随机负采样, label=0, bpr loss, item侧atom特征: [moveid, title], item侧side特征[genres], query侧atom特征[userid], query侧side特征[gender, age, occupation, zip\_code],

	ndcg	bpr
双塔 (concat (side_emb + atom_emb) -> mlp)	0.942	0.428
双塔 (concat ((side_emb-> mlp) , (atom_emb-> mlp))	0.941	0.431
内生解耦	<b>0.950</b>	<b>0.350</b>