

内生解耦-v20230208

内生解耦

对原子特征做进一步思考，一次搜索请求，排序需要决策是item的排序位置，item是treatment，response是是否点击，质量模型中使用双塔结构捕捉item和query之间关系。针对这个设置，我们有两个判断：

1. 原子特征item id等内生的

2. 非原子特征可以作为instrument variable

根据之前针对支付宝搜索bias分析，马太效应严重，item对应原子特征item id等在这个场景下是内生的，举个极端例子，两个都非常相关的itemA和B，A有足够的曝光，B没有曝光，模型从样本中学习很容易得出A相关B不相关的结论，而非原子特征内生性相对弱很多，可以作为instrument variable，关于内生和instrument variable详细见[5]

经济领域一种借助instrumental variable解决变量内生方法是two-stage least squares，先将内生变量投影到instrumental variable，再用response对投影后的内生变量做回归

Stage 1: Regress each column of \mathbf{X} on \mathbf{Z} , ($\mathbf{X} = \mathbf{Z}\delta + \text{errors}$):

$$\hat{\delta} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{X},$$

and save the predicted values:

$$\hat{\mathbf{X}} = \mathbf{Z}\hat{\delta} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{X} = \mathbf{P}_Z \mathbf{X}.$$

In the second stage, the regression of interest is estimated as usual, except that in this stage each endogenous covariate is replaced with the predicted values from the first stage:

Stage 2: Regress \mathbf{Y} on the predicted values from the first stage:

$$\mathbf{Y} = \hat{\mathbf{X}}\beta + \text{noise},$$

which gives

$$\beta_{2SLS} = (\mathbf{X}^T \mathbf{P}_Z \mathbf{X})^{-1} \mathbf{X}^T \mathbf{P}_Z \mathbf{Y}.$$

stage1投影非常有意思，对应我们场景，可以将每个塔表征层拆成两部分，两部分相互对立，第一部分原子特征（内生特征）表征 \mathbf{x} ，第二部分非原子特征（instrumental特征）表征 \mathbf{z} ， \mathbf{x} 可以解

偶成两部分 $\mathbf{x} = \mathbf{x}^\perp + \hat{\mathbf{x}}$ ， $\hat{\mathbf{x}} = \frac{\mathbf{x}^T \mathbf{z}}{\|\mathbf{x}\| \|\mathbf{z}\|} \mathbf{z}$ 是 \mathbf{x} 在 \mathbf{z} 上的投影， $\mathbf{x}^\perp = \mathbf{x} - \hat{\mathbf{x}}$ 与 \mathbf{z} 正交，通过对表征层的在计算相似度时候只保留 \mathbf{x}^\perp ，这种方式有几个好处

1. \mathbf{x}^\perp 更多关注记忆效应， \mathbf{z} 用来做泛化， \mathbf{x} 丢掉 $\hat{\mathbf{x}}$ ，减少高频交互pair影响，从而使用 \mathbf{z} 学习到更重要权重，增强模型泛化性
2. 可以单独对 \mathbf{x}^\perp 加L2正则，约束低频原子特征的噪音影响，在长尾query/item交互上更加稳定
3. instrumental表征 \mathbf{z} 更加稳定，构造ipw，修正z相似损失函数权重修正后loss如下，通过只估计给定非原子特征下该item出现概率，简化ipw估计，减少其方差，同时也有非常好的物理意义，即非原子特征的泛化能力是通过item体现的，而不是通过pv体现，这样就和instrumental variable的定义也联系了起来（instrumental variable 通过变量x起作用）

, , 其中 $g(x,z)$ 表示ipwe,

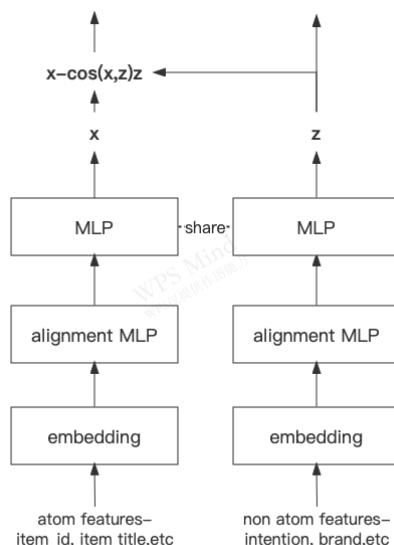


图-9

算法方案

单塔自监督表征学习

在query/item每一侧塔，模拟instrumental regression stage1过程，在 \mathbf{x} 上做回归，同时通过batch采样构造负样本，对原子特征表征向量和side表征向量各自增加对比损失，控制不同表征均匀分布在向量空间内，整体损失函数由三部分构成：

1 stage1回归loss，形式上类似hinge loss，要求原子特征表征向量 \mathbf{z} 在side表征向量 \mathbf{x} 周围，采样hinge loss而并未采用mse主要是考虑到，1) 通常有一些unobserved confounders，side不能完全确定一个atom 2) 同一个side下的不同atom的表征要有差异空间，实践中取margin为0.5

$$\max(0, \text{margin} - \cos(\mathbf{x}, \mathbf{z}))^2$$

2 side表征向量对比损失，其中为batch随机负采样，损失函数形式选择nce (noise cross entropy)，label y 如果是同一个样本则为1，否则为0，ltr训练要求同一次搜索的样本尽可能在一个batch内，方便计算pairwise/listwise loss，从而同一次batch不同负采样样本有较大概率是false negative，我们增加一个loss权重，如果和过于相似，则置该样本权重为0

$$I\{\cos(\mathbf{z}, \mathbf{z}') < 0.99\} \cdot \log_loss(\cos(\mathbf{z}, \mathbf{z}')/\tau, y)$$

3 atom表征向量对比损失，类似side表征向量对比损失，考虑到stage1回归loss已经要求要求原子特征表征向量 \mathbf{z} 在side表征向量 \mathbf{x} 周围，直接计算不能达到最优（各自彼此垂直时对比loss最优），因此改为计算向量对比损失loss

$$l' = l([x_q^\perp, z_q], [x_i^\perp, z_i]) + \beta \cdot \frac{1}{g(x, z)} \cdot l(z_q, z_i)$$

对齐q-i两边表征

利用监督信号，对齐q-i两边表征，损失函数如下

$$l' = l([x_q^\perp, z_q], [x_i^\perp, z_i]) + \beta \cdot \frac{1}{g(x, z)} \cdot l(z_q, z_i)$$

其中 $g(x,z)$ 表示ipwe, $g(x, z) = P(x_i, x_q | z_i, z_q)$

其中ipw计算是难点，估计得到的ipw本身variance很大，误差大ipw样本会对整体模型学习造成比较大的负面影响，同时考虑到ipw小的样本（频次大）提供的信息更置信（大数定律），因此考虑用一种近似的可以减少variance方式计算ipw。特别的，考虑到side表征在向量空间更靠近频次大的atom表征（stage1 loss导致），在item侧做验证，计算side表征和atom表征的相似度，观察同一个意图下二者的关系和全局规一化之后的关系，可以看到单意图和全局都是正相关（如下图），说明高频atom特征确实影响了side特征表征，这启发我们可以用这个相似度来替换ipw。新的ipw计算公式如下，近似ipw偏差会较大，但误差不会直接影响优化目标稳定性，实现上监督loss对ipw部分的变量做梯度截断，避免ipw学习到q-i交互的监督信号

$$g(x, z) = \max(1.0 - (1.0 + \cos(x_i, z_i))/2, 1.0 - (1.0 + \cos(x_q, z_q))/2)$$

item_multi_level_intent_names	cnt_range	avg_sim
二手/租赁,二手交易意图	0	0.34619879722595215
二手/租赁,二手交易意图	1	0.669051965077718
二手/租赁,二手交易意图	3	0.8806126752169803
二手/租赁,二手交易意图	4	0.9466642733686459
二手/租赁,二手交易意图	5	0.8903079820721024
二手/租赁,二手交易意图,二手买卖,买二手手机	0	0.13445790112018585
二手/租赁,二手交易意图,二手买卖,买二手手机	1	0.7458608150482178
二手/租赁,二手交易意图,二手买卖,买二手手机	2	0.5589359700679779
二手/租赁,二手交易意图,二手买卖,买二手手机	3	0.6579642889596888
二手/租赁,二手交易意图,二手买卖,买二手手机	4	0.8982285261154175

normalized_pv	avg_sim
0.0	0.4091351641006239
0.1	0.3973881934770399
0.2	0.37666743684192516
0.3	0.48101296089377316
0.4	0.5464381232869758
0.5	0.6560790050205234
0.6	0.5683531185971943
0.7	0.6748465985526619
0.8	0.5663765467194689
0.9	0.6460382789345626