

赞同 82

分享

Figure 2: Transfer knowledge among models by soft-label

KDD'24 [Airbnb] | 模型蒸馏多目标学习: Multi-objective Learning to Rank by Model Distillation



tiantianheheaa
欢迎评论交流~~

关注

82 人赞同了该文章

收起

文章链接: [Multi-objective Learning to Rank by Model Distillation](#)

- 个人评价
- 优点:
 - 背景贴近工业生产实际业务中的多目标问题。将后链路目标作为约束，将前链路的短期收益目标作为主要目标。
 - 解决思路清晰:
 - 多目标约束问题用蒸馏学习来解决: teacher模型预估后链路目标，产出soft-label。student模型预估主要目标，用hard-label训练。并可以引入临时业务需求等不可微目标。
 - 通过自蒸馏解决蒸馏学习中的训练开销和模型不可再现性。
 - 很喜欢读Airbnb的文章，贴近实际业务，就像同行在交流。
- 缺点:
 - 自蒸馏的缺点是不能及时引入更多数据信息，长期在线效果会衰减。最终模式会像在线学习一样，固定周期的离线训练来更新线上模型。

关键词

- 搜索排序、多目标优化、learn to rank、[模型蒸馏](#)

摘要

- 搜索排名的目标是购买（主要目标），也包括购买结果（次要目标），例如订单取消、评论评级、客户服务查询、平台长期增长等。多目标学习用来平衡主要目标和次要目标。
- 挑战
 - 参数调节，导致次优。
 - 数据稀疏。和ad-hoc目标不兼容。
- 基于蒸馏的多目标排序
 - 大大提高了主要目标，满足次要目标约束，提高了模型稳定性。
 - 提出的系统可以通过模型自蒸馏进一步简化。
 - 可以有效的将ad-hoc **不可微的业务目标** 注入到rank系统中，使我们能够平衡我们的优化目标。

1 介绍

工作

题描述

多目标优化到模型蒸馏

系统的多目标学习

系统概述

标签编码是什么?

Ad-hoc业务目标

训练开销和不可再现性

实验设置

模型性能

蒸馏测试

模型再生性测试

Ad-hoc业务目标测试

工作



要,例如订单取消、客户服务查询、评论评级和长期增长。尽管更高的转化率可以增加收入,但是取消、退单、客户服务的相关成本也会吃掉利润。过度强调短期转化收益,可能偏向已经建立的item,而不是新的item,从而危及长期增长,并限制新的item曝光机会。

本质是考虑了更多后链路目标。对于业务长期发展来讲,确实都需要考虑。但是目标也有主次之分,通常前链路目标作为主要目标,更关注短期收益。后链路目标作为约束,更关注长期价值。

- **[29]回顾了多目标推荐系统和排序算法**,它将工作分为两类:标量化方法[11]、多目标进化算法MOEAs[81]。工业实践中,标量化方法比MOEA更受欢迎,因为简单有效,易于扩展到大型训练数据。本文提出的方法和标量化技术相关,因此我们将专注和当前的标量化方法进行比较。
- 最流行的直观方法是模型融合[6],为每个目标单独训练一个模型,通过加权和组合起来。缺点:目标之间相互干扰,简单的结合可能导致次优解决方案。
- 更好的方法是为每个目标的loss函数分配权重,然后相加为单个目标,从而多目标优化转化为单目标优化。也被称为标量化方法。在[11,24,10,9,21]中得到广泛应用。
- 另一种方法是e-constraint方法[12],将一个目标作为优化目标,次要目标作为约束。
- 出了目标cost函数,一些研究将多个label聚合为一个label,从而将多目标cost function构建为单个cost function。仍然被看做一个特殊的标量化方法。
- 这些方法在讨论如何更好的优化,如何更好的权衡。但是忽略了一个重要的事实:不同目标之间的数据是不平衡的。例如点击和转化数据可能很多,订单取消和客户服务查询的数据要少很多。
- 标量化方法通常是一个多任务学习方案[20,21,29]。多任务学习的一个优点是所有任务共享底层,这样数据较少的任务从数据较多的任务中受益。在搜索中,目标通常相反,长期增长和短期利益之间会有冲突。模型倾向于将质量或价格更高的成熟item排到头部。因此,虽然越来越多的冲突目标被添加到模型中,会有效率较低的学习和次优解决方案。
- 虽然目标不相关时,多任务学习方法可能是次优的,但是由于共享emb的力量,仍然比简单模型的融合表现的更好。这也是**Airbnb实践中获得的经验[21]**,该多任务系统为所有目标训练模型,并取得显著的业务指标改进。该系统的一个学习是调整两组参数,具有挑战性:一是训练损失权重,另一个是服务阶段的在线分数融合权重。第二个参数,在下网格搜索具有成本效益,即使找到权重也不能保证是最优的。在线分数融合是不稳定的,针对一组模型调节的权重可能不适用另一组模型。
- 除了上述挑战,还有一个问题:现在的排序模型大多是深度模型,标量化方法假设每个成本函数是可微的,因此聚合的单个目标函数可以通过反向传播进行训练。但有时一个目标可能只是一个规则:例如新item曝光的百分比,它有利于长期增长。这个目标不能直接包含在优化项中,必须在模型训练和部署后进行一些手动调节,手动调节会降低模型性能。
- 最近,另一个趋势是将模型蒸馏应用于rank和推荐系统[22,26,16,23],该领域所以方法都侧重2个主题:一是将大而复杂的模型提炼成紧凑高效的模型。2是将集成模型提炼成一个单一的模型。目的是在保持模型性能的同时降低推理成本。模型蒸馏的损失函数和多目标优化的标量化方法具有相似的结构,我们将模型蒸馏和多目标学习排序集成到一起,从而解决上述挑战。
- 本文,一个通用的模型蒸馏方法来优化多目标。这是**第一次将模型蒸馏应用到多目标rank模型**中。本文贡献:
 - 将多目标学习重新定义为模型蒸馏问题,从而缓解数据不平衡,拜托在线分数融合的权重调整。在满足次要目标约束的同时,实现更好的主要目标。
 - 将soft-label概念引入MO-LTR,来扩展基于蒸馏的多目标排序。证明它可以帮助减少模型的不可再现性,并通过自蒸馏**软标签**+简化提出的rank系统。
 - 通过修改soft-label,可以很容易的将不可微的目标加入到模型中。

2 相关工作

- LTR是目前的一个重要分支是进化loss
 - point-wise loss: 预测每个item的动作概率,例如pctr、pcvr。
 - pair-wise loss: 每次查看两个目标,将rank转化为二元分类问题, item a是否由于item b。
 - list-wise loss: 考虑整个item列表,并试图近似最优顺序。
- 早期,只优化单目标:**NDCG⁺**。早期是web search,现在online market是一个双边的市场:一个典型的用户是浏览大量item,将物品进行比较,购买的物品还可能被用户取消或退回。如果用户不满意,可能会抱怨客户服务。商家也将经历类似的过程。所以,仅优化NDCG是不够的,开始将多目标优化MOO应用于rank:[19]将MOO用于人才匹配系统,[6]将MOO用于平衡收入和购买,[15]优化了相关性、购买、质量、评级、回报在内的多目标,[11]提出了一种平衡GMV和CTR的成熟方法,[21]通过针对pre rank和rank的一致性,优化点击、预定、取消、拒

- 一共四种变形
 - 模型不共享任何参数，但是和聚合损失函数联合训练。预测时，所有模型打分融合。
 - 典型的深度学习多任务模型设置。模型共享底层来学习emb。预测时，所有模型打分融合。
[21][20]
 - 一个模型来拟合多目标，预测时，模型分直接用于排序分。例如[6][11]
 - 每个模型没有可训练的参数，每个模型都是经过预训练来优化 c_i 。预测时，所有模型打分融合。
- 变体1和2是多任务学习的设置，2更受欢迎，因为共享底层emb可以学习共享和更好的表征，还有助于缓解数据稀疏。1和2都需要调节两组权重：一是训练损失，二是模型在线预测的融合分。变体3的优点是单个模型的分数直接用于预测，但是单个模型来拟合多目标更具挑战。变体4实际上是模型融合，简单，也很受欢迎，但是表现可能更差。
- 和我们最相关的先前工作是[6]，提出了两种方法
 - 随机label聚合：对于每条训练样本，label从label集合中随机采样，每个label对应一个目标。
 - 两阶段模型组合：一是训练每个模型，优化不同目标。二是模型分数作为另一个模型的特征，该模型使用随机label聚合的方法进行训练。
- [6]显示了它们方法的理论优势，但是跨目标的训练数据不平衡，所以单个模型设置，较小的目标可能会被大数据的目标淹没。在Airbnb的rank系统中，目标的训练数据高度不平衡，最坏情况下，label不平衡率大于10。[6]中的训练数据更加平衡：2个数据集的平衡很好，1个数据集的不平衡比例为3。此外，上面提到的4个标量化方法的变体都不能处理不可微的目标，例如在工业中常见的手动规则。

3 框架

3.1 问题描述

3.2 从多目标优化到模型蒸馏

经过一系列公式推导，得到一个单目标的损失函数，它和模型蒸馏的损失函数非常类似[8][26]。第一项是预测和主目标的ground true之间的损失。第二项是蒸馏损失，推动模型逼近soft-label。

4 排序系统的多目标学习

4.1 系统概述

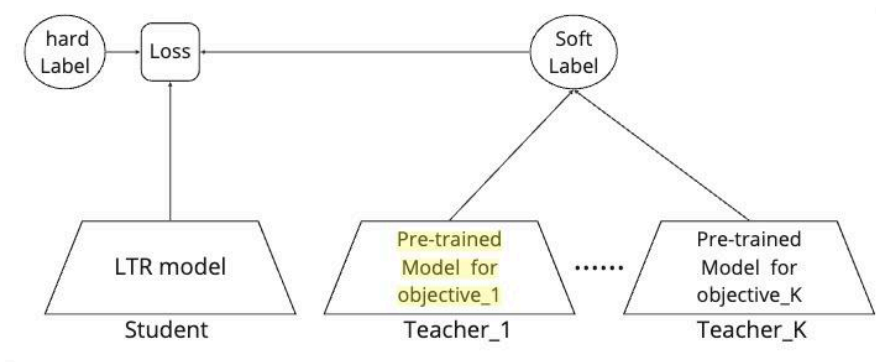


Figure 1: Model training graph @tiantianheheaa

模型训练图

- 在训练时，每个预训练模型被冻结，不可训练。
- 在训练期间，所有的模型（LTR模型和预训练模型）被加载到计算图中。所有的预训练模型（教师模型）权重被冻结，因此预训练模型仅在前向传播时用于soft-label的计算。在infer阶段，所有预训练模型被丢弃，只有LTR模型（学生模型）导出用于服务。

- soft-label的优点
 - 通过引入soft-label, 为每个训练样本分配了一个密集的soft-label向量, 有助于缓解数据不平衡。[6]中的多目标算法, 当数据不平衡时, 不得不进行训练数据中的下采样或上采样, 这可能损坏模型训练效率。
 - soft-label充当了正则化(类似标签平滑正则[25])。soft-label给出了多目标LTR模型的先验知识。
 - soft-label可以很好的处理不可微目标。
 - 可以将rank知识传递给新版本的rank模型。工业中, 模型需要定期重训, 来捕捉新的用户趋势。通常采用冷启的方式, 每次从头开始训练模型, 模型会忘记之前学到的知识。通过soft-label可以将模型学到的知识传递给新版本, 有助于减少模型的不可再现性和不稳定性。

4.3 Ad-hos业务目标

- 一些目标的cost是不可微函数, 特别是来自临时业务需求。例如在搜索结果中显示更多的新item, 来帮助新的业务者和市场长期增长。在顶级搜索结果中对高质量item进行rank, 来提高市场的品牌。
- 这种目标本身就是模糊的, 不能从过去的学习中学习优化, 它们通常与用户行为模型观察到的方向相反。例如用户倾向购买完善的item。模型可以很容易从数据中学到并对完善的item进行排序, 但是新item很难。尽管可以用item和item的协同过滤帮助新item, 但是和目标高度相关的特征(例如评论数量、评论等级、ctr、cvr等)仍然会使得学习偏向成熟的item, 这些特征在新item缺失。传统的多目标优化很难包括这种不可微的目标。
- 一种解决方案是通过平均 相似的成熟item的特征值 作为新item的特征值, 但无法保证学习效率。没有定量的方法这种做法对新item的影响。而且盲目假设新item是一个平均水平 可能对过去表现不佳的item不公平, 甚至损害用户体验和其他商业目标, 例如cvr。
- 另一种解决方案是直接修改label。hard-label是购买是1, 非购买是0, 新item给一个小的提升。最终label是购买>新item>非购买。这没有任何意义, 因为将新item放到其他所有item之上, 然后购买。但是这种提升可以用于soft-label。例如软标签给6个item打分, 其中新item的分数提升一些。或许有人争辩直接将boost添加到rank排序中, 但这将对业务指标伤害更大, 如5.5节中的实验, 将到时模型性能下降。将boost添加到soft-label中, 让模型和其他目标平衡可能产生更好的结果。

4.4 训练开销和不可再现性

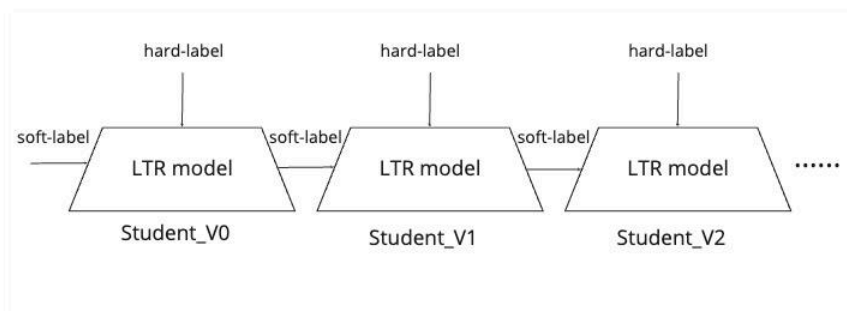


Figure 2: Transfer knowledge among models by soft-label

soft-label在模型间传递知识

- 冷启训练模型, 会忘记之前学到的东西, 并导致不可再现性和不稳定性。这在多任务学习中尤为重要, 因为预期多目标之间的平衡是持续稳定的。
- soft-label是从预训练模型中训练出来的, 会存在训练开销的问题。同时这些预训练模型也要定期重新训练, 也具有不可再现性。收到Born-Again神经网络的启发[7,17], 我们在每个模型再训练过程中, 通过自蒸馏的方式传递soft-label到新版本模型中。
- 自蒸馏
 - v0的soft-label来自多个预训练模型
 - v1的soft-label来自v0
 - v2的soft-label来自v1
- 通过这种设计, 训练复杂度显著降低: 不需要更新和维护预训练模型。而是将知识提炼为soft-label验证模型训练的路径传递。当然, 如果需要, 这些预训练模型可以随时插入。在5.3和5.4

5 实验

5.1 实验设置

- 搜索rank是多目标，主要目标是cvr，次要目标是用户取消率、host取消率、host拒绝率、平台长期增长等。
- 提出的模型MO-LTR-MD⁺使用最近几个月收集的360万个训练样本，仅包含cvr样本。每个目标的预训练模型使用相同的训练集，同一日期范围内收集的500万个训练样本，每个样本都有多个标签：order、点击、取消等。这些预训练模型是多任务学习系统[21]联合训练的。基线模型是相同的多任务学习系统。
- 提出的MO-LTR-MD模型 只有一个蒸馏权重，我们进行了网格搜索，发现0.2是最合适的。soft-label的权重计算（14）来自我们的生产设置，这是一个模型融合系统（在此无法共享这些权重的绝对值以包含我们的核心业务数据）。当我们移动到自蒸馏阶段，soft-label来自学生模型的最后一个版本的分数，（14）中的权重就不再需要了。

5.2 整体模型性能

- 离线评估：使用NDCG，其中booking的相关性分数是1，其他搜索曝光的相关性分数是0。和多任务学习基线相比[21]，NDCG+1.1%的显著改进。
- 在线AB测试：cvr+0.37%。其他次要目标的变化是中性的。soft-label对来自每个目标的预训练模型的知识进行编码，更重要的是充当正则化器来减轻数据不平衡问题，该模型可以更好的平衡主要目标和次要目标。
- 模型训练成本和在线infer耗时：使用更少的数据进行训练，尽管加载预训练模型生成soft-label，但是总训练时间和基线模型相同。在线infer，由于只有学生模型，而基线系统中有多个rank模型，所以infer时间显著降低，-1.6%。

5.3 自蒸馏测试

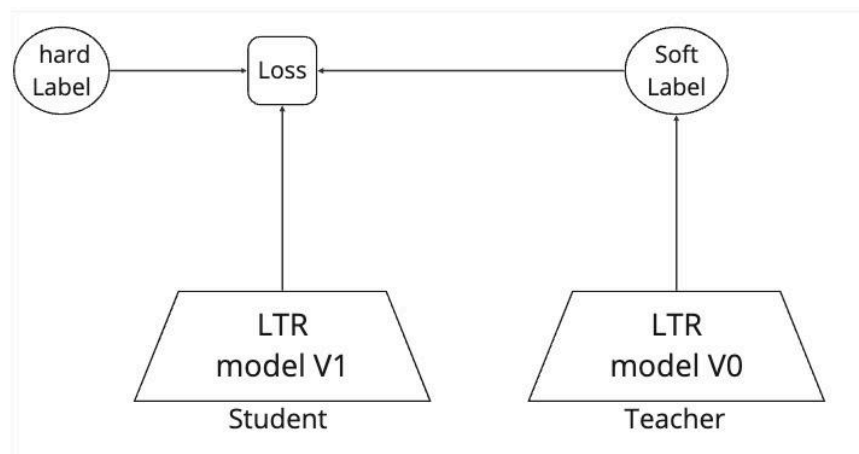


Figure 3: Self distillation

自蒸馏示意图

- 4.4节中所叙述，为了保证模型的再训练是稳定的，并减少模型蒸馏引起的操作。我们提出了自蒸馏。
- 自蒸馏的两个问题：1在给定预训练模型的情况下，自蒸馏是否损坏业务指标。2如果将新的训练数据添加到模型训练中，自蒸馏过程中，预训练模型不存在，无法更新，这将损坏模型和业务指标。
- 测试步骤
- 结果
 - 离线NDCG几乎相同。表明，即使没有预训练模型的情况下，仅使用soft-label将知识转移到更新模型中，也不会伤害，添加新训练数据不会稀释soft-label的作用。

5.4 模型非再生性测试

Model	Change rate	PD
Baseline	77%	0.407
MO-LTR-MD	36%	0.363
Improvement	53%	11%

Table 2: Model Irreproducibility

模型的非再生性

5.5 Ad-hoc业务目标测试

Methods	NDCG
directly boost score	-0.5%
Inject boost to soft-label	-0.1%

Table 3: Boosting Impact To NDCG

Boosting的影响


- 为了保护我们的核心业务数据，我们无法披露系统中的真实临时业务目标。相反，模拟了一些从未在我们系统中应用过的临时目标。例如在顶级搜索结果中提高更高评论评级的item。
- 如果直接提升rank score，会有-0.5%的NDCG损失。在训练时增加soft-label只需要牺牲-0.1%的NDCG损失。为什么注入soft-label会伤害更少的指标。深度虚席模型训练通常是一个非凸优化问题，最终可能会获得不同的局部最小值。soft-label可以被认为是正则化器，通过一些提升注入到正则化器中，我们将模型推向和临时业务目标更对齐的局部最小值。

6 未来的工作

- MO-LTR-MOD可以找到更好的优化解决方案，并提高模型的不可逆性，并且可以处理不可微目标。
- 在自蒸馏测试中，soft-label可以通过通过自蒸馏的方式稳定的转移到新模型。尽管在线测试只测试了两轮自蒸馏模型，我们将更仔细的跟进在线测试，确保知识转移不会长期衰减。特别是大多数次要目标只能在在线测试中检查。
- 我们从现有的模型融合设置中 引导我们的蒸馏系统，因此不需要调整 (14) 中的soft-label权重。图4所示，我们将通过导入MoE层[9] 将权重学习折叠到模型训练中，这项工作将在未来完成。

编辑于 2024-10-30 14:15 · IP 属地北京

内容所属专栏

 单篇论文

订阅专栏

推荐算法 知识蒸馏