

赞同 9

分享

百度2024：EmerG-解锁冷启动内容潜力——基于特征交互图的CTR优化框架



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

9 人赞同了该文章

Introduction

冷启动问题在[推荐系统](#)⁺中是一个显著的挑战，特别是在新内容从没有用户交互（称为冷启动阶段）过渡到积累初始点击（称为暖启动阶段）的背景下。深度学习模型以其捕捉复杂特征交互的能力而闻名，然而，这些模型通常需要大量数据集以实现最佳性能，这一要求在冷启动阶段成为了一种限制。由于这些模型的参数规模庞大，它们在有限的交互记录阶段适应性能不佳，从而加剧了在冷启动和暖启动阶段准确预测CTR和更新模型而无需承担高额成本的挑战。

近期的研究着重于通过改进内容ID嵌入的初始化来缓解推荐系统中的内容冷启动问题，这使得在暖启动阶段随着可用交互记录的增加，可以通过梯度下降进行后续更新。然而，他们忽略了一个关键方面：不同用户和内容之间特征[交互模式](#)⁺的独特性。这种忽视限制了这些模型完全捕捉用户内容交互的微妙动态的能力，可能在需要个性化推荐的场景中影响CTR预测的准确性和有效性。

例如，将高价奢侈品与低价日常必需品进行比较。对于高价奢侈品，内容价格与用户收入水平之间的交互是决定用户购买意愿的关键因素。具体而言，价格与收入的二阶特征交互可能是决定因素。而对于低价日常必需品，用户收入水平对购买决策的影响较小。在这种情况下，其他特征交互，如用户年龄与内容类别之间的交互，变得相对更为重要。这种差异强调了建模特定内容特征交互的必要性。现有工作都学习了一个全局的用户与内容特征交互模式，这使得具有大量交互记录的旧内容掩盖了具有有限交互记录的新内容。

因此，我们引入了EmerG，通过学习特定内容[特征图](#)⁺来解决具有递增交互数据（从无交互记录到少量记录再到大量记录）的新出现内容的CTR预测问题。如下：

- 我们提出了一种独特的方法，着重于特定内容特征间的交互，通过减少具有大量数据的老内容对新内容[点击率](#)⁺预测的覆盖效应，解决了新内容CTR预测的挑战。利用超网络，我们构建了特定于内容的特征图，其中节点代表特征，边代表它们之间的交互，捕捉到了每个内容特有的复杂交互模式。
- 我们使用了一种[图神经网络](#)⁺（GNN），其中包含了定制的消息传递过程，该过程可以被证明能够捕获任意阶次特征交互，这些交互可以聚合成精细且准确的预测。
- 鉴于数据有限，为了减少[过拟合](#)⁺，我们采用了一种元学习策略，该策略在不同商品点击率预测任务中优化了超网络与GNN的参数，同时调整每个任务内部的商品特定参数，调整量较小。通过这种方式学习的超网络与GNN，预期能够轻松地对每个任务进行泛化。

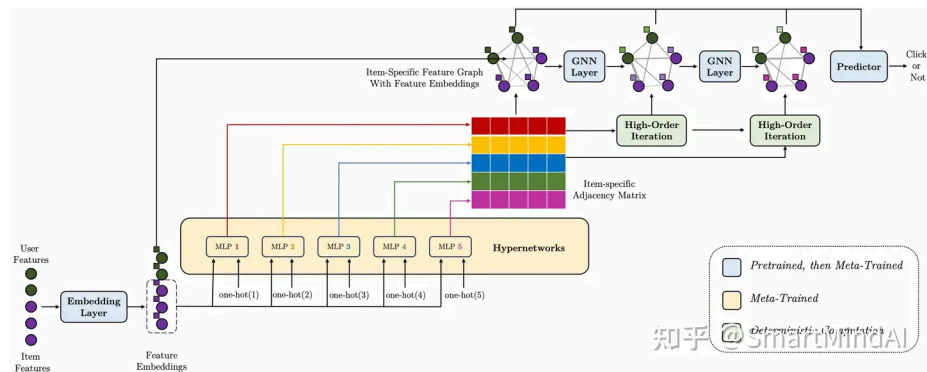
Problem Formulation

设 $\mathcal{V} = \{v_i\}$ 表示一组内容集合，其中每个内容 v_i 关联有 N_v 个内容特征，如内容ID、类型和价格。同样，设 $\mathcal{U} = \{u_j\}$ 表示一组用户集合，其中每个用户 u_j 关联有 N_u 个用户特征，如用户ID、年龄和家乡。

的点击率。每个任务 \mathcal{T}_i 对应一个旧内容 v_i ，包含一个训练集 $\mathcal{S}_i = \{(v_i, u_j, y_{i,j})\}_{j=1}^{N_i}$ ，其中包含与内容 v_i 相关的现有交互历史，以及一个测试集 $\mathcal{Q}_i = \{(v_i, u_j, y_{i,j})\}_{j=1}^{N_q}$ ，用于预测用户 u_j 是否点击内容 v_i 。 N_s 和 N_q 分别是 \mathcal{S}_i 和 \mathcal{Q}_i 中的交互数量。

在测试阶段，我们考虑从新内容开始的点击率预测，这些内容在开始时没有任何交互记录。我们处理三个阶段：第一阶段：新内容开始，无交互记录。第二阶段：逐渐收集新内容的一些交互记录。第三阶段：积累足够的交互记录，以进行预测。

The Proposed EmerG



特征交互对于捕捉内容的独特性至关重要，我们提出EmerG（图）来捕获通过关联的特征交互模式。此系统包含两个关键组件：(ii) 不同任务共享的超网络，用于生成用于编码特征图的内容特定邻接矩阵 $\mathbf{A}_i^{(1)}$ ；和 (i) 对于生成的内容特定特征图操作的GNN，其消息传递机制专门设计以证明能够捕捉任意阶的特征交互。

考虑到冷启动和热身阶段，我们进一步设计了一种元学习策略，该策略优化超网络和GNN的参数，适用于各种内容点击率预测任务，同时也在每个任务中仅调整少量内容特定参数。这种策略有效地降低了处理数据有限时过拟合的风险。

Embedding Layer

给定一个样本 (u, v) ，嵌入层将用户特征 u 和内容特征 v 映射到稠密向量中。对于第 m 个特征 f_m ，其中 m 在 $[1, N_v + N_u]$ 范围内，其特征嵌入 e_m 是为：

$$e_m = \begin{cases} W_{e,m} \cdot \text{one-hot}(f_m) & \text{if } f_m \text{ is single-valued} \\ \sum_e W_{e,m} \cdot \text{multi-hot}(f_m) & \text{if } f_m \text{ is multi-valued} \\ W_{e,m} \cdot f_m & \text{if } f_m \text{ is continuous} \end{cases}$$

其中向量矩阵 $W_{e,m}$ 表示第 m 个特征的嵌入矩阵，向量 $\text{one-hot}(f_m)$ 表示单值特征 f_m 的 one-hot 向量，向量 $\text{multi-hot}(f_m)$ 表示多值特征 f_m 的 multi-hot 向量。

Item-Specific Feature Graph Generation

我们采用超网络策略，遵循Ha，生成针对每个商品的特征图。超网络是小型神经网络，用于训练生成主网络的大参数，因为其整合高度依赖于问题的具体情况，它们在教育中面临一个独特的挑战。在EmerG中，超网络被用于产生初始邻接矩阵 $\mathbf{A}_i^{(1)}$ ，用于编码第一个GNN层的针对每个商品的特征图。

我们通过简化流程，允许后续的GNN层从初始的 $\mathbf{A}_i^{(1)}$ 中推导出其邻接矩阵，从而优化存储效率，同时不牺牲模型对每个商品的特定性。对于任务 \mathcal{T}_i 对于商品 v_i ，对于商品的特征 f_1, \dots, f_{N_v} ，商品特征的嵌入分别表示为 $e_{1,i}, \dots, e_{N_v,i}$ ，特征图是一个图，其中每个节点代表特征 f_m ，节点之间的边记录了它们之间的交互。

用于编码在第一个GNN层中使用的特征图：表示第 m 行的 $\bar{\mathbf{A}}_i^{(1)}$ ，其计算方式为：

$$m: = \text{MLP}_{\mathbf{W}_a}([\mathbf{e}_{1,i}, \dots, \mathbf{e}_{N_v,i}, \text{one-hot}(m)]),$$

其中 $\text{MLP}_{\mathbf{W}_a}$ 表示参数为 \mathbf{W}_a 的多层感知器⁺ (MLP)。然后，我们计算 $\bar{\mathbf{A}}_i^{(l)}$ 如下：

$$\bar{\mathbf{A}}_i^{(l)} = \bar{\mathbf{A}}_i^{(l-1)} \cdot \bar{\mathbf{A}}_i^{(1)}$$

在最小最大归一化处理 (min-max normalization) 之后，我们首先对密集的 $\bar{\mathbf{A}}_i^{(l)}$ 进行稀疏化处理，使得只有两个高度相关的特征相连。进一步，由于 \odot 运算的交换律⁺，我们将 $\hat{\mathbf{A}}_i^{(l)}$ 转换为一个对称矩阵⁺。除了上述考虑之外，我们还期望在低阶特征图中断开连接的节点在高阶特征图中也断开连接。例如，如果在第 l 层的GNN层中节点 n_2 的消息没有传播到节点 n_1 ，那么在更高层的GNN层中，节点 n_2 的消息也不会传播到节点 n_1 。因此，我们获得最终的 $\mathbf{A}_i^{(l)}$ 。

Customized Message Passing Process on Item-Specific Feature Graph

在已学习到的内容特定特征图上，我们使用了一个GNN，它具有一个自定义的消息传递过程，这个过程能够证明捕捉任意阶的特征交互，然后这些交互被明确地组合到最终的点击率预测中。在第 l 层GNN层中，特征 f_m 的节点嵌入值 $\mathbf{h}_m^{(l)}$ 被更新为：

$$\mathbf{h}_m^{(l)} = \text{UPD}^{(l)}\left(\mathbf{h}_m^{(l-1)}, \text{AGG}^{(l)}\left(\left\{\mathbf{h}_n^{(l-1)} : f_n \in \mathcal{N}(f_m)\right\}\right)\right),$$

在 EmerG 中，其中节点嵌入 $\mathbf{h}_m^{(l)}$ 更新了 f_m 的节点嵌入，节点嵌入 $\mathbf{h}_m^{(0)} = \mathbf{e}_m$

在经过 N_l 层后，节点嵌入 $\mathbf{h}_m = \mathbf{h}_m^{(N_l)}$ 被作为最终特征表示返回。

$$\mathbf{h}_m^{(l)} = \mathbf{h}_m^{(l-1)} \odot \left[\sum_{n=1}^{N_v+N_u} [\mathbf{A}_i^{(l-1)}]_{mn} \mathbf{W}_g^{(l-1)} \mathbf{h}_n^{(0)} \right],$$

因此，我们通过公式 (eq:mask) 获得最终的专门邻居矩阵为 $\mathbf{A}_i^{(l-1)}$ ，以及可学习的参数 $\mathbf{W}_g^{(l-1)}$ 。与现有的GNNs不同，我们聚合 $\mathbf{h}_m^{(l-1)}$ 与 $\mathbf{h}_n^{(0)}$ ，而不是与 $\mathbf{h}_n^{(l-1)}$ 进行聚合。这样，正如定理所示，第 $(l-1)$ 层GNN层的输出是 l 阶特征交互，这使得我们的方法能够明确地模拟任意阶特征交互。

根据特征交互的不同阶数，我们通过多头注意力明确地将每个节点 f_m 的所有节点嵌入组合到更新后的节点嵌入 $\hat{\mathbf{H}}_m$ 中：

$$\begin{aligned} \text{attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \text{softmax}(\mathbf{Q}\mathbf{K}^\top / \sqrt{N^d})\mathbf{V}, \\ \text{head}_h &= \text{attention}(\mathbf{W}_{g,h}\mathbf{H}_m, \mathbf{W}_{k,h}\mathbf{H}_m, \mathbf{W}_{v,h}\mathbf{H}_m), \\ \hat{\mathbf{H}}_m &= [\text{head}_1; \dots; \text{head}_{N_h}], \end{aligned}$$

在其中 $\mathbf{H}_m = [\mathbf{h}_m^{(0)}; \dots; \mathbf{h}_m^{(N_l)}]$ 包含了 N_l 个维度为 N_d 的行向量，注意力头的数量为 N_h 。然后，我们为每个特征计算贡献度因子，其中特征数量为 $N_v + N_u$ 。

$$[c_1, \dots, c_{N_v+N_u}] = \text{sigmoid}\left(\text{MLP}_{\mathbf{W}_{c,1}}([\hat{\mathbf{H}}_1, \dots, \hat{\mathbf{H}}_{N_v+N_u}])\right)$$

由 $\mathbf{W}_{c,1}$ 参数化⁺，我们预测用户 u 和内容 v 的互动状态为：

$$\hat{y} = \sum_{m=1}^{N_v+N_u} c_m \cdot \text{MLP}_{\mathbf{W}_{c,2}}(\hat{\mathbf{H}}_m), \text{ 其中 } \mathbf{W}_{c,2} \text{ 是一个可训练的参数。}$$

Learning and Inference

在处理有限数据时降低过拟合风险，我们引入了一种元学习策略，该策略在各种内容点击率预测任务中跨任务优化了超网络和GNN的参数，同时每个任务中仅调整了一组最小的内容特定参数。

我们将超网络表示为 $\text{hyper}_{\theta_{\text{hyper}}}$ ，其中 $\theta_{\text{hyper}} = \mathbf{W}_a$ 是共享的可训练参数。

参数预测器的参数为 $\mathbf{W}_{c,1}$, $\mathbf{W}_{c,2}$ 以及 ϕ_i 代表内容特定参数:

$$\phi_i = \{\text{hyper}_{\theta_{\text{hyper}}}(v_i), \mathbf{e}_{\text{ID},i}\},$$

由超网络生成的物品特定邻接矩阵 $\mathbf{A}^{(1)}$ 和物品 v_i 的随机化物品 ID 嵌入 $\mathbf{e}_{\text{ID},i}$ 包含在内。我们旨在学习 θ_{GNN}^* 和 θ_{hyper}^* 通过仅生成 ϕ_i 并使用梯度下降对 ϕ_i 进行预热, 可以在新物品 v_i 上实现良好的冷启动和热启动性能。

我们针对 \mathcal{T}^{old} 中的 N_t 个任务, 对 EmerG 进行优化: Optimization Objective: 仅生成 ϕ_i 并使用梯度下降对 ϕ_i 进行预热, 以在新物品 v_i 上实现良好的冷启动和热启动性能。通过针对 \mathcal{T}^{old} 中的 N_t 个任务的优化, 我们旨在提升 EmerG 的性能, 特别是在新物品 v_i 的冷启动和热启动方面。

$$\sum_i^{N_t} \gamma \mathcal{L}_{\mathcal{S}_i}(\theta_{\text{GNN}}, \phi_i) + (1 - \gamma) \mathcal{L}_{\mathcal{Q}_i}(\theta_{\text{GNN}}, \phi'_i)$$

其中 γ 是一个超参数⁺, 用于平衡两个损失项的贡献。特别是, 第一项可以表示模型在未接触 \mathcal{S}_i 中的标签时的冷启动阶段性能。我们计算 $\mathcal{L}_{\mathcal{S}_i}(\theta_{\text{GNN}}, \phi_i)$

为:

$$\begin{aligned} \mathcal{L}_{\mathcal{S}_i}(\theta_{\text{GNN}}, \phi_i) \\ \equiv 1/|\mathcal{S}_i| \cdot \sum_{(v_i, u_j, y_{ij}) \in \mathcal{S}_i} \text{BCE}(y_{ij}, \text{GNN}_{\theta_{\text{GNN}}, \phi_i}(v_i, u_j)) \end{aligned}$$

其中 $\text{BCE}(y, \hat{y}) = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y})$ 是二元交叉熵⁺。

我们计算 $\mathcal{L}_{\mathcal{Q}_i}(\theta_{\text{GNN}}, \phi'_i)$ 为:

$$\mathcal{L}_{\mathcal{Q}_i}(\theta_{\text{GNN}}, \phi'_i) = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y})$$

在更新 ϕ_i 后, 使用在 \mathcal{S}_i 中提供的少量新内容实例。

$$\begin{aligned} \mathcal{L}_{\mathcal{Q}_i}(\theta_{\text{GNN}}, \phi'_i) \\ \equiv 1/|\mathcal{Q}_i| \cdot \sum_{(v_i, u_i, u_{ii}) \in \mathcal{Q}_i} \text{BCE}(y_{ij}, \text{GNN}_{\theta_{\text{GNN}}, \phi'_i}(v_i, u_j)) \end{aligned}$$

通过执行梯度下降步骤, 得到 ϕ'_i :

$$\phi'_i = \phi_i - \alpha_1 \nabla_{\phi_i} \mathcal{L}_{\mathcal{S}_i}(\theta_{\text{GNN}}, \phi_i),$$

在这种情况下 α_1 表示学习速率。

Experiments

Experimental Settings

数据集 我们使用两个基准数据集: (i) 电影Lens, (ii) 淘宝 ### 评价指标 遵循现有工作的方法, 我们通过以下指标来评价性能: (i) AUC (ii) F1分数, 是精确率和召回率的调和平均数⁺。

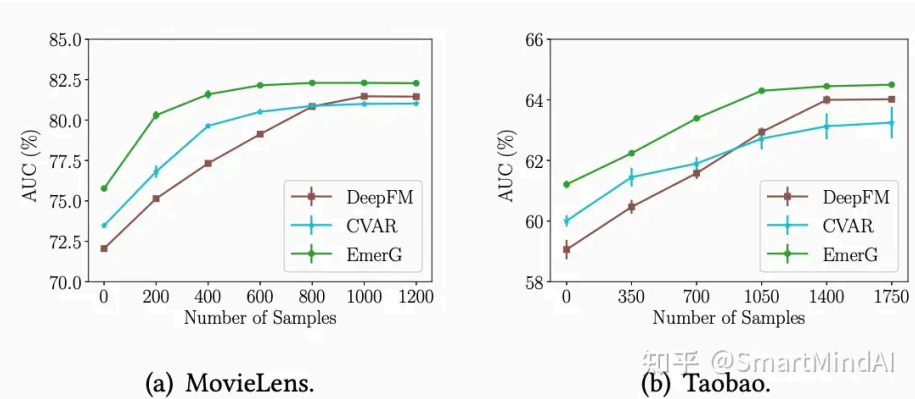
Performance for Cold-Start & Warm-Up Phases

Wide&Deep	69.44 _(0.29)	59.53 _(0.31)	74.82 _(0.29)	64.44 _(0.21)	77.58 _(0.25)	66.82 _(0.24)	79.09 _(0.22)	67.67 _(0.27)
AutoInt	68.64 _(0.24)	59.63 _(0.14)	75.60 _(0.31)	64.93 _(0.36)	77.65 _(0.33)	66.84 _(0.42)	79.20 _(0.34)	67.77 _(0.36)
LorentzFM	68.91 _(0.15)	56.22 _(0.27)	75.35 _(0.17)	62.77 _(0.21)	78.46 _(0.08)	66.23 _(0.25)	79.85 _(0.02)	67.93 _(0.08)
AFN	71.23 _(0.42)	61.76 _(0.37)	74.26 _(0.08)	64.39 _(0.09)	76.19 _(0.24)	65.84 _(0.16)	77.36 _(0.35)	66.71 _(0.28)
Fi-GNN	71.37 _(0.05)	61.46 _(0.08)	74.62 _(0.03)	63.83 _(0.12)	76.83 _(0.06)	65.71 _(0.05)	78.49 _(0.05)	66.74 _(0.06)
FinalMLP	69.51 _(0.06)	60.59 _(0.17)	78.48 _(0.12)	67.34 _(0.10)	78.47 _(0.16)	67.27 _(0.07)	79.07 _(0.17)	68.00 _(0.10)
FINAL	71.64 _(0.15)	61.72 _(0.17)	77.87 _(0.13)	66.99 _(0.10)	77.94 _(0.10)	66.93 _(0.14)	78.29 _(0.09)	67.42 _(0.10)
DropoutNet	72.94 _(0.17)	62.43 _(0.18)	-	-	-	-	-	-
ALDI	65.53 _(0.13)	57.47 _(0.23)	-	-	-	-	-	-
MeLU	-	-	77.54 _(0.06)	66.71 _(0.11)	79.43 _(0.10)	68.51 _(0.05)	80.26 _(0.03)	68.13 _(0.06)
MAMO	-	-	77.69 _(0.10)	66.92 _(0.13)	79.61 _(0.07)	68.72 _(0.04)	80.37 _(0.05)	68.49 _(0.04)
TaNP	-	-	79.15 _(0.10)	68.39 _(0.14)	80.49 _(0.17)	69.43 _(0.15)	80.71 _(0.09)	69.63 _(0.09)
ColdNAS	-	-	77.45 _(0.03)	67.01 _(0.03)	77.88 _(0.12)	67.25 _(0.21)	78.06 _(0.09)	67.31 _(0.11)
MetaE	71.82 _(0.70)	61.76 _(0.30)	79.53 _(0.25)	67.96 _(0.15)	80.27 _(0.09)	68.31 _(0.12)	80.47 _(0.04)	68.46 _(0.12)
CVAR	73.58 _(0.21)	63.15 _(0.12)	78.23 _(0.10)	67.03 _(0.26)	80.28 _(0.06)	68.76 _(0.12)	81.06 _(0.04)	69.33 _(0.14)
GME	71.54 _(0.13)	64.31 _(0.10)	75.81 _(0.20)	67.50 _(0.26)	78.10 _(0.18)	69.26 _(0.20)	79.15 _(0.12)	69.95 _(0.16)
MWUF	73.19 _(0.66)	62.61 _(0.74)	78.88 _(0.11)	67.34 _(0.22)	80.26 _(0.08)	68.40 _(0.13)	80.57 _(0.05)	68.66 _(0.10)
EmerG	75.44 _(0.05)	64.76 _(0.15)	79.92 _(0.27)	68.61 _(0.24)	81.28 _(0.21)	69.71 _(0.14)	81.82 _(0.16)	70.26 _(0.14)

Taobao	Cold-Start Phase		Warm-Up Phase A		Warm-Up Phase B		Warm-Up Phase C	
	AUC(%)	F1(%)	AUC(%)	F1(%)	AUC(%)	F1(%)	AUC(%)	F1(%)
DeepFM	59.01 _(0.84)	13.47 _(0.42)	60.68 _(0.65)	14.27 _(0.20)	61.51 _(0.64)	14.56 _(0.34)	62.34 _(0.54)	15.00 _(0.24)
Wide&Deep	59.07 _(0.44)	13.65 _(0.06)	60.92 _(0.56)	14.25 _(0.10)	61.69 _(0.51)	14.56 _(0.15)	62.33 _(0.46)	14.75 _(0.13)
AutoInt	55.69 _(1.37)	12.14 _(0.22)	58.65 _(1.26)	13.61 _(0.51)	59.43 _(1.20)	13.84 _(0.36)	60.07 _(1.13)	14.19 _(0.39)
LorentzFM	56.53 _(0.41)	12.72 _(0.04)	60.83 _(0.50)	14.15 _(0.24)	61.26 _(0.47)	14.31 _(0.14)	61.96 _(0.45)	14.60 _(0.14)
AFN	57.94 _(0.99)	13.28 _(0.15)	58.99 _(0.74)	13.73 _(0.20)	59.81 _(0.87)	13.99 _(0.18)	60.18 _(0.69)	14.18 _(0.13)
Fi-GNN	56.92 _(0.08)	12.79 _(0.10)	60.00 _(0.13)	14.06 _(0.18)	62.09 _(0.14)	14.82 _(0.14)	62.46 _(0.21)	14.90 _(0.05)
FinalMLP	60.64 _(0.12)	13.57 _(0.04)	63.44 _(0.06)	14.83 _(0.03)	63.49 _(0.07)	14.80 _(0.03)	64.05 _(0.02)	15.05 _(0.03)
FINAL	60.53 _(0.24)	13.63 _(0.05)	63.30 _(0.12)	14.81 _(0.06)	63.35 _(0.13)	14.74 _(0.04)	63.93 _(0.12)	15.01 _(0.02)
DropoutNet	60.41 _(0.09)	13.53 _(0.02)	-	-	-	-	-	-
ALDI	50.10 _(0.18)	10.93 _(0.05)	-	-	-	-	-	-
MeLU	-	-	61.37 _(0.17)	14.09 _(0.17)	62.48 _(0.04)	14.34 _(0.05)	63.07 _(0.07)	14.64 _(0.11)
MAMO	-	-	61.96 _(0.11)	14.31 _(0.09)	62.52 _(0.05)	14.34 _(0.04)	63.15 _(0.12)	14.78 _(0.13)
TaNP	-	-	55.67 _(0.22)	11.92 _(0.31)	55.85 _(0.16)	12.07 _(0.16)	56.19 _(0.09)	12.08 _(0.11)
ColdNAS	-	-	54.27 _(0.07)	10.89 _(0.05)	54.86 _(0.14)	11.33 _(0.13)	55.01 _(0.09)	11.71 _(0.13)
MetaE	59.75 _(0.37)	13.58 _(0.06)	61.19 _(0.26)	14.01 _(0.09)	62.06 _(0.31)	14.41 _(0.10)	62.87 _(0.32)	14.71 _(0.07)
CVAR	60.56 _(0.46)	13.71 _(0.13)	62.54 _(0.19)	14.38 _(0.06)	63.17 _(0.10)	14.69 _(0.05)	63.95 _(0.18)	15.09 _(0.12)
GME	60.57 _(0.23)	13.32 _(0.33)	62.55 _(0.17)	13.96 _(0.22)	63.29 _(0.05)	14.39 _(0.12)	63.85 _(0.13)	14.52 _(0.08)
MWUF	59.65 _(0.46)	13.44 _(0.15)	62.08 _(0.17)	14.20 _(0.07)	63.03 _(0.13)	14.33 _(0.07)	63.77 _(0.11)	14.95 _(0.16)
EmerG	61.58 _(0.03)	13.99 _(0.05)	63.56 _(0.03)	15.02 _(0.06)	63.76 _(0.02)	15.15 _(0.01)	64.22 _(0.02)	15.21 _(0.02)

如表所示，为冷启动和热身阶段设计的冷启动方法通常表现更好。在所有四个阶段中，《TheName》始终表现最佳，这验证了通过超网络捕获内容特定特征交互的有效性。

Performance Given Sufficient Training Samples



所有方法在给定更多训练样本时都能获得更好的性能。在拥有大量训练样本（常规阶段）时始终表现最佳，并且其性能优于其他方法，这验证了我们的有效性，能够以不同阶次优雅地捕捉内容特定的特征交互。

发布于 2024-08-12 14:11 · IP 属地北京

百度 ctr 冷启动

理性发言，友善互动