用于点击率预测的深度行为路径匹配网络

董健[1], 于益松[2,3], 张亚鹏[1], 吕一民[2,3], 王树立[1], 金北虹[2，3*]、王永康[1*]、王星星[1], 王栋[1]美团[2]中国科学院软件研究所[3]中国科学院大学，北京，中国 董健03@meituan.com，Beihong@iscas.ac.cn，Wangyongkang03@meituan.com

arXiv:2302.00302v1 [cs.AI] 1 Feb 2023

摘要

电商应用中的用户行为不仅包含对物品的不同类型反馈，有时还蕴含着用户决策的认知线索。为了理解用户决策背后的心理过程，我们提出了行为路径，并通过将用户当前的行为路径与历史行为路径进行匹配，来预测用户在应用中的用户行为。进一步地，我们设计了一个用于行为路径匹配的深度神经网络，解决行为路径建模中的三个难题：稀疏性、噪声干扰以及行为路径的精准匹配。具体而言，我们利用对比学习增强用户行为路径，引入行为路径自激活以减轻噪声的影响，并采用两级匹配机制识别最合适的候选项。我们的模型在两个真实世界数据集上展现出卓越性能，优于最先进的点击率预测模型。此外，我们的模型已部署于美团外卖平台，使点击率提升了1.6%，广告收入提升了1.8%。

CCS 概念

• **信息系统** →**推荐系统**。

关键词

点击率预测，用户行为建模

美国计算机协会参考格式：

董健，于益松，张亚鹏[1],吕一民[2,3],王树立[1],金北虹[2，3*]，王永康[1*]，王星星[1], 王栋[1]。2023年。《深度行为路径匹配网络：点击率预测》。载于《（会议简称“*XX*”）论文集》。美国计算机协会，纽约州，纽约，美国，5页。https://doi.org/XXXXXXX.XXXXXXX

1 引言

美团外卖APP是一款面向餐饮与零售的应用。用户可以通过该应用浏览并选择兴趣点（POI，例如餐馆、食品店和咖啡馆），并下单购买由配送员快速送达的食品。该应用需要理解用户决策背后的心理学，并向用户推送相关候选项，从而提高点击率（CTR），进一步提升交易量和广告收入。

我们注意到，用户在应用中的行为是用户决策心理的重要体现。然而，尽管一些现有的点击率预测模型从长序列或多种行为的角度分析了用户行为，但它们采用的是候选项与历史行为序列中单个行为之间的点到点激活，没有考虑包含用户决策轨迹的序列行为所产生的影响。因此，对于点击目标兴趣点这一行为，我们将用户此前的序列行为（包括浏览兴趣点、下单等）视为一条行为路径。通过观察美团外卖APP上的历史数据，我们发现行为路径与点击行为之间存在密切关联。

上述观察促使我们开发一种模型，通过行为路径匹配来预测用户下一次点击。其核心思想是从用户行为路径中学习与决策心理学相关的潜在因子，以生成其嵌入向量。获得行为路径的嵌入向量后，模型将对历史行为路径与当前行为路径进行匹配，并估算候选项的点击率。

然而，用户行为路径的建模颇具挑战，因为存在三个难点：行为路径的稀疏性、行为路径中的噪声干扰，以及行为路径之间的精确匹配。首先，对于单个用户而言，用户与应用之间的交互并不多，因此难以捕捉用户的所有行为模式。为了解决行为路径的稀疏性问题，我们利用对比学习来扩充用户行为路径的正样本，并优化用户行为路径的学习。其次，用户行为路径中存在大量噪声。例如，用户可能因为某个兴趣点的封面吸引人而点击它，但一旦觉得不喜欢，便会立即返回。这种行为实际上会成为路径中的噪声。为减少噪声的影响，我们构建了一个动态激活网络，以关注路径中的若干主要行为。与对路径中的所有行为一视同仁相比，动态激活更有效且高效，因为某些行为确实对后续行为有更明显的影响。最后，我们提出了一种两级匹配机制。在第一级，对于当前路径，我们计算每条历史行为路径的激活权重，然后选出最相似的前 k 条历史路径。在第二级，给定候选项和

允许免费为个人或课堂使用而制作本作品的全部或部分数字或纸质副本，但不得以营利或商业利益为目的制作或分发副本，且副本必须保留本声明及首页上的完整引文。对于本作品中归美国计算机协会以外的其他方所有的组件，必须尊重其版权。允许在注明出处的情况下撰写摘要。若要以其他方式复制、转载、发布到服务器或分发至列表，则需事先获得具体许可和/或支付费用。请通过 permissions@acm.org 申请许可。会议简称“*XX*”，*2023*年*6*月*3*日至*5*日，纽约州伍德斯托克© 2023年 美国计算机协会。美国计算机协会 ISBN 978-1-4503-XXXX-X/18/06...$15.00 https://doi.org/XXXXXXX.XXXXXXX

所选路径，我们计算所选路径之后发生的点击行为的激活权重，用于点击率预测。

我们的主要贡献总结如下。

- 我们首次将用户行为路径匹配引入工业界的点击率预测。我们指出了行为路径建模面临的挑战，即行为路径的稀疏性、噪声和匹配问题。
- 我们提出了一种用于点击率预测的深度行为路径匹配网络（DBPMaN），该网络对行为路径进行增强，提供行为路径自激活，并执行两级匹配（首先在行为路径层面匹配，然后在点击行为层面匹配）。
- 我们在两个规模不同的真实世界数据集上进行了离线实验，并在美团广告业务中进行了在线 A/B 测试。实验结果表明，DBPMaN 效果显著，达到了最先进的性能。

2 相关工作

点击率预测作为推荐系统的核心部分，一直是业界和学术界关注的热点 [18]。

点击率预测的经典解决方案是学习特征交互，其中 DeepFM [6]、xDeepFM [9]和 ONN [17] 是早期具有代表性的深度神经网络模型，而 CAN [1] 在当前开源点击率模型中取得了最先进的性能。

近年来，序列行为建模成为点击率预测的新驱动力。建模行为的粒度涵盖从单一行为（例如 DIN [21]）到多种行为（例如 FeedRec [15]），从短序列（例如 DIEN [20]、DSIN [5]）到超长序列（例如 MIMN [11]、SIM [12]、ETA [2]）。这些模型旨在捕捉用户兴趣 [5, 19–21] 或意图 [8]，并且通常采用点对点激活方法，其输入仅包含单一类型的行为（如点击），从概率角度估计用户兴趣/意图对候选项的倾向。事实证明，这些模型持续提升了点击率预测的准确性。此外，随着 Transformer [14] 和 BERT [4] 在自然语言处理领域取得巨大成功，它们也被引入推荐系统，以完成不同的召回任务 [7, 13] 或点击率预测任务 [3]。

与上述工作相比，我们的工作着重于蕴含决策信号的行为路径，并将行为路径用作点击率预测的证据基础。

3 我们的模型

3.1 概述

首先，我们给出本文中使用的以下定义。

定义 1.（用户行为序列）设 U 表示一个用户集合。对于用户 $u \in \mathrm{U}$，其行为序列由其用户行为组成，按发生时间排序，记为 $s = [b_1, ..., b_i, ..., b_T]$，其中 b_i是第 i个行为，T 是行为序列的长度。在我们的场景中，该序列包含用户过去一年内的用户行为。每个行为包括交互物品的 ID、行为类型、发生时间与当前时间之间的时间间隔，以及其相对位置等。在我们的场景中，存在三种行为类型：点击、曝光和下单。

定义 2.（用户点击序列）在用户行为序列 s 中，存在大量点击行为。因此，我们可以从 s 构建点击序列 s^c：$s^c = [b_1^c, ..., b_i^c, ..., b_t^c]$，其中 t 表示点击序列的长度。

定义 3.（用户行为路径）对于 s^c 中第 i 个点击行为 b_i^c，令 $b_{m(i)}$ 表示行为序列 s 中对应的行为，其中 $m(i)$ 表示 s 中第 i 个点击行为出现的位置。然后，与点击行为 b_i^c 对应的用户行为路径记为 p_i，它是 s 中的子序列 $[b_{m(i)-l}, ..., b_{m(i)-2}, b_{m(i)-1}]$，其中 l 是预设的行为路径长度。

根据用户行为路径的定义，我们可以明显看出，点击行为与行为路径一一对应。图 1 给出了用户行为路径的示例。在历史用户行为序列中，有三条长度预设为 3 的用户行为路径：一条对应于 h_4，一条对应于 h_7，还有一条对应于 h_{11}。

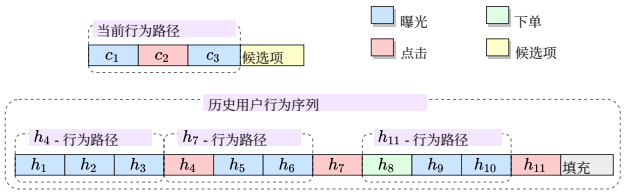

图1：用户行为路径示例，其中 l=3。

定义4.（行为路径序列）对于 s中的所有点击行为，我们可以分别获取其对应的行为路径，从而构成行为路径序列 $P = [p_1, ..., p_i, ..., p_t]$，其中$p_i$是与点击行为 b_i^c相对应的用户行为路径。

其次，我们介绍DBPMaN模型的组成、结构和流程。

DBPMaN由一个嵌入层和三个模块组成，即路径增强模块（PEM）、路径匹配模块（PMM）和路径扩充模块（PAM），其结构如图2所示。

DBPMaN以多种特征作为输入，这些特征包括：（1）物品档案，包括物品ID及其侧信息（例如类别、位置、评分等）；（2）用户档案，包括用户ID及其侧信息（例如年龄、性别、城市等）；（3）用户行为序列中的行为。这些特征被输入嵌入层。对于所有特征，我们都可以通过查询嵌入表获取其嵌入向量。然后，我们对不同类型的特征进行求和池化，计算用户行为序列的嵌入向量$\mathbf{s} = [\mathbf{e}_1, ..., \mathbf{e}_i, ..., \mathbf{e}_T]$、用户嵌入$\mathbf{e}^u$以及候选物品的嵌入向量$\mathbf{e}^{ct}$。

用户行为路径中的各项行为对相应后续点击行为的贡献往往各不相同，也就是说，后续点击行为是指用户行为路径之后发生的点击行为。PEM 旨在挖掘这一信息，从而学习更准确的路径嵌入。简而言之，给定每条行为路径及其相应点击行为的嵌入向量，PEM 首先激活

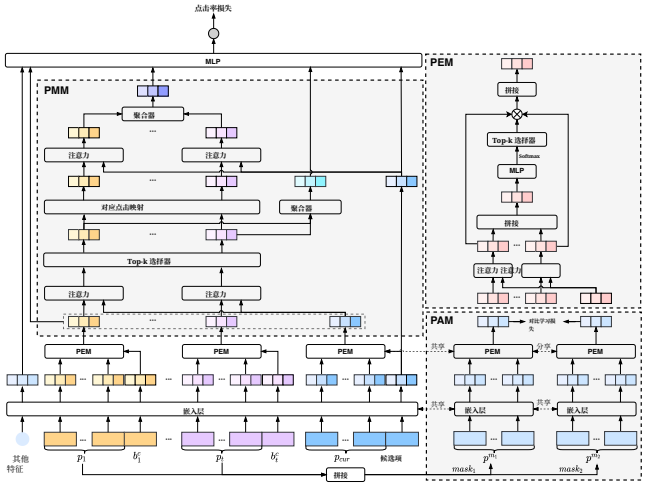

图2：DBPMaN 的结构。(1) PEM 增强路径表征。(2) PMM 匹配路径。(3) PAM 增强路径。

行为路径中的重要行为，然后优化该行为路径的嵌入，以获得更准确的行为路径表示。

然后，PMM 使用当前行为路径和历史行为路径的嵌入向量作为输入，搜索与当前行为路径最相似的 k 历史行为路径，随后激活与候选项对应的 k 点击行为。

此外，PAM 旨在通过对比学习获得更精确、信息更丰富的行为路径嵌入。具体而言，我们对每条历史行为路径进行掩码处理，以获得两条增强路径，并将它们输入嵌入层和 PEM，计算各自的嵌入向量；随后，采用 InfoNCE 损失 [10] 作为对比损失，将来自同一行为路径的嵌入向量拉近。

DBPMaN 使用负对数似然函数作为主要损失，该函数广泛应用于大多数点击率模型。最后，DBPMaN 将主要损失与对比损失相结合，以此作为优化目标进行训练。

由于篇幅有限，我们仅详细介绍 PEM 和 PMM，省略模型的其他部分。

3.2 路径增强模块（PEM）

对于一条行为路径 $[b_{m(i)-l}, ..., b_{m(i)-2}, b_{m(i)-1}]$ 以及后续点击行为 b_i^c，嵌入层将生成它们的嵌入向量，分别记为 $\mathbf{e}_{m(i)-l}, ..., \mathbf{e}_{m(i)-2}, \mathbf{e}_{m(i)-1}$ 和 $\mathbf{e}_i^c$。此外，将嵌入向量序列 $[\mathbf{e}_{m(i)-l}, ..., \mathbf{e}_{m(i)-2}, \mathbf{e}_{m(i)-1}]$ 记为 $\mathbf{s}_i$。

我们首先在用户行为路径上应用局部激活单元，通过加权拼接池化自适应地计算行为路径的嵌入表示，如公式 1 所示。

$$\begin{aligned} \mathbf{e}^{te}_{m(i)-j} &= a\left(\mathbf{e}_{m(i)-j}, \mathbf{e}_i^c\right) \cdot \mathbf{e}_{m(i)-j}, \quad 1 \le j \le l \\ \mathbf{s}_i^{te} &= [\mathbf{e}^{te}_{m(i)-l}, ..., \mathbf{e}^{te}_{m(i)-2}, \mathbf{e}^{te}_{m(i)-1}] \\ \mathbf{p}_i^{te} &= \text{concat}(\mathbf{s}_i^{te}) \end{aligned} \tag{1}$$

其中，$a(\cdot)$ 是一个多层感知机，其输出用作第一级激活分数。

然后，我们将 $\mathbf{p}_i^{te}$ 输入另一个多层感知机，并通过 softmax 激活函数学习行为路径中每个行为的第二级激活分数，如公式 2 所示。

$$\mathbf{score}_i = \text{softmax}(\text{MLP}(\mathbf{p}_i^{te})) \tag{2}$$

$\mathbf{score}_i$ 是一个 l 维向量，其各项表示行为路径中各行为的第二级激活分数。然后，仅选取分数最高的 top-k 项。根据选取的分数，将它们与 $\mathbf{s}_i^{te}$ 中对应的行为嵌入相乘。将按分数缩放后的嵌入向量拼接起来，即可得到增强的路径嵌入 $\mathbf{p}_i^e$。

这样，我们就可以得到增强路径嵌入序列 $\mathbf{P}_e = [\mathbf{p}_1^e, \mathbf{p}_2^e, ... \mathbf{p}_t^e]$。

3.3 路径匹配模块（PMM）

对于某个用户，其用户行为序列中可能存在大量行为路径。然而，只有其中少数与当前行为路径相似，因而能够反映用户的当前兴趣。PMM 旨在搜索与当前路径最相似的前 k 个行为路径，然后获取相应的 k 个点击行为，我们认为这些行为对用户当前兴趣有重要影响。

具体而言，给定增强后的历史路径嵌入序列 $\mathbf{P}_e = [\mathbf{p}_1^e, \mathbf{p}_2^e, ..., \mathbf{p}_t^e]$ 以及当前行为路径的增强嵌入 $\mathbf{p}_{cur}^e$，我们将每个 $\mathbf{p}_i^e \in \mathbf{P}_e$ 和 $\mathbf{p}_{cur}^e$ 输入评分门控，并得到相似度分数 g_i^p，该分数反映相应历史行为路径的重要性。评分门控的计算如公式 3 所示。

$$g_i^p = \text{MLP}(\text{concat}(\mathbf{p}_{cur}^e, \mathbf{p}_i^e, \mathbf{p}_{cur}^e \otimes \mathbf{p}_i^e)) \tag{3}$$

其中，$\otimes$ 表示哈达玛积，MLP( )则通过前馈神经网络实现。

由此，我们可以得到一组相似度分数 $g^p = [g_1^p, g_2^p, ..., g_t^p]$。我们对所有分数进行排序，并选取得分最高的前 k 个。根据这前 k 个分数，我们可以获得相应的历史路径，以及这些历史路径对应的点击行为。对于选出的点击行为，其点击行为嵌入序列表示为 $\mathbf{s}_c = [\mathbf{e}_{c1}, \mathbf{e}_{c2}, ..., \mathbf{e}_{ck}]$。此外，我们将每个选出的路径嵌入乘以对应的分数，得到调整后的所选路径嵌入，如公式 4 所示。

$$\mathbf{E}^p = \text{concat}(Filter(g^p, [g_i^p \cdot \mathbf{p}_i^e, 1 \le i \le t], k)) \tag{4}$$

其中，函数 $Filter(score, embedding, k)$ 根据得分对历史行为路径排序，并选取得分最高的 top-k 条路径。

所有选出的点击行为对用户当前兴趣的贡献应该各不相同。因此，我们采用与式 3 相同的方法，计算候选项与选出的点击行为之间的相似度得分，从而考虑它们之间的相关性，自适应地计算用户兴趣的表示向量，如式 5 所示。

$$\begin{aligned} g_i^c &= \text{MLP}(\text{concat}(\mathbf{e}^{ct}, \mathbf{e}_{ci}, \mathbf{e}^{ct} \otimes \mathbf{e}_{ci})) \\ \mathbf{E}^c &= \text{concat}([g_i^c \cdot \mathbf{e}_{ci}, 1 \le i \le k]) \end{aligned} \tag{5}$$

最后，将 $\mathbf{P}_e$、$\mathbf{E}^p$、$\mathbf{E}^c$、$\mathbf{e}^u$ 和 $\mathbf{e}^{ct}$ 拼接起来，然后输入多层感知机层，输出预测的点击率。

4 实验评估

4.1 实验设置

数据集。 我们采用以下两个数据集进行实验。

• 淘宝：一个包含10天交互数据的公开数据集。我们按照 CAN 在 [1] 中采用的相同方式对数据进行预处理。
• 美团：由美团外卖APP收集的一个工业数据集，包含1亿用户14天的交互数据。

表1列出了处理后数据集的统计信息。

对比模型。 我们选择以下专注于特征交互建模的点击率模型作为对比模型。

• DeepFM [6]。它将因子分解机与深度学习相结合，以建模低阶和高阶特征交互。• xDeepFM [9]。它使用提出的压缩交互网络（CIN）生成特征交互，并进一步将 CIN 与基础深度神经网络整合为一个统一模型。

• DIN [21]。它设计了一个局部激活单元，根据候选项从历史行为中学习用户兴趣的表征。
• DIEN [20]。它设计了兴趣提取层和兴趣演化层，以从行为序列中捕捉用户兴趣。• ONN [17]。它为不同操作学习不同的表征。

• CAN [1]。它通过协同行为单元将表征学习与特征交互建模解耦。

指标。 我们使用 AUC 和 RelaImpr [16] 作为离线实验的指标，使用点击率和千次展示成本作为在线实验的指标。

实现细节。 我们通过 Tensor- flow 实现 DBPMaN [1]。对于所有模型，我们使用 Adam 作为优化器，学习率设为 0.001。模型参数初始化为高斯分布（均值为 0，标准差为 0.01）。物品嵌入维度设为 18。

4.2 性能比较

我们进行了对比实验，将我们的模型与上述竞品进行比较。不同模型在两个数据集上的性能结果如表 2 所示。

从结果中可以看出，所有模型在两个数据集上的性能排名相同，且我们的 DBPMaN 超越了所有竞品。我们认为，DBPMaN 中针对路径到路径激活的注意力机制有助于其胜过其他模型，包括基于注意力的模型（即 DIN、DIEN 和 ONN）；后者采用点到点激活，将候选项与历史行为序列中的单个行为进行关联。

一个有趣的发现是，与淘宝数据集相比，我们的 DBPMaN 模型在美团数据集上取得了更显著的提升。这可能源于外卖场景的特点。用户周边的食品店相对较少，加上用户对食物的兴趣特点，导致用户行为历史中反复出现与同一家食品店的交互。这使得激活与当前路径相关的历史路径变得容易。

[1]代码可在 https://github.com/Ethan-Yys/DBPMaN 获取。

表1：数据集统计信息。

数据集	#用户数	#物品数	#类别数	#交互数
淘宝	987991	4161138	9437	100095182
美团	100000000	15755909	184	5648932411

表 2：离线评估的性能结果。

模型	淘宝		美团	
	AUC	RelaImpr	AUC	RelaImpr
DeepFM	0.8125	-13.19%	0.6673	-1.55%
xDeepFM	0.8366	-10.61%	0.6693	-1.25%
ONN	0.8689	-7.16%	0.6705	-1.08%
DIN	0.9308	-0.54%	0.6753	-0.37%
DIEN	0.9324	-0.37%	0.6761	-0.25%
CAN	0.9359	0.00%	0.6778	0.00%
DBPMaN	**0.9381**	**0.24%**	**0.6812**	**0.50%**

表 3：美团数据集上的消融实验。

模型	AUC	CTR	CPM
DBPMaN	0.6812	-	-
DBPMaN（不含 PEM）	0.6789	-0.6%	-0.5%
DBPMaN（不含 PMM）	0.6769	-1.0%	-1.1%
DBPMaN（不含 PAM）	0.6801	-0.2%	-0.2%

表 4：在线 A/B 测试结果。

模型	CTR	CPM	推理时间（毫秒）
基线	-	-	48
DBPMaN (l=8)	+1.6%	+1.8%	+0.8
DBPMaN (l=16)	+1.3%	+1.5%	+2.1

4.3 消融实验

我们在美团数据集上进行了消融实验，以评估 DBPMaN 各关键模块的贡献。我们将模型与三个变体进行比较，即 DBPMaN（不含 PEM）、DBPMaN（不含 PMM）和 DBPMaN（不含 PAM）。结果如表 3 所示。

从表 3 可以看出，与原始 DBPMaN 相比，三个变体在三项指标上均有所下降。DBPMaN（不含 PMM）的性能下降幅度最大，这表明 PMM 比另外两个模块发挥着更重要的作用。

4.4 在线 A/B 测试

该 A/B 测试在美团外卖平台上进行，持续 14 天，从 2022-08-10 到 2022-08-23。其中，基线模型是我们上一个在线点击率模型，仅使用点对点激活方法。结果如表 4 所示，其中 l 表示行为路径的长度。目前，DBPMaN（l=8）已部署上线，为用户的主要流量提供服务。

5 结论

本文提出了 DBPMaN，这是首次将用户行为路径建模用于点击率预测的方法。除展现出优异性能外，DBPMaN 还通过对行为路径进行建模，展示了探索用户决策心理的可能性。

参考文献

[1] Weijie Bian、Kailun Wu、Lejian Ren、Qi Pi、Yujing Zhang、Can Xiao、Xiang-Rong Sheng、Yong-Nan Zhu、Zhangming Chan、Na Mou 等。2022。CAN：用于点击率预测的特征协同作用网络。载于第十五届 ACM国际网络搜索与数据挖掘会议论文集。57–65。[2] Qiwei Chen、Changhua Pei、Shanshan Lv、Chao Li、Junfeng Ge 和 Wenwu Ou。2021。点击率预测模型中的端到端用户行为检索。arXiv 预印本 arXiv:2108.04468（2021）。[3] Qiwei Chen、Huan Zhao、Wei Li、Pipei Huang 和 Wenwu Ou。2019。阿里巴巴电子商务推荐中的行为序列转换器。载于首届高维稀疏数据深度学习实践国际研讨会论文集。1–4。[4] Jacob Devlin、Ming-Wei Chang、Kenton Lee 和 Kristina Toutanova。2019。BERT：用于语言理解的深度双向变换器预训练。载于2019年计算语言学协会北美分会会议：人类语言技术论文集，NAACL-HLT 2019，美国明尼苏达州明尼阿波利斯，2019年6月2日至7日，第1卷（长篇与短篇论文）。4171–4186。[5] Yufei Feng、Fuyu Lv、Weichen Shen、Menghan Wang、Fei Sun、Yu Zhu 和 Keping Yang。2019。用于点击率预测的深度会话兴趣网络。载于第二十八届国际人工智能联合会议论文集，IJCAI 2019，中国澳门，2019年8月10日至16日。2301–2307。[6] Huifeng Guo、Ruiming Tang、Yunming Ye、Zhenguo Li 和 Xiuqiang He。2017。DeepFM：基于因子分解机的神经网络，用于点击率预测。载于第二十六届国际人工智能联合会议论文集，IJCAI 2017，澳大利亚墨尔本，2017年8月19日至25日，Carles Sierra（编）。1725–1731。[7] Wang-Cheng Kang 和 Julian McAuley。2018。自注意力序列推荐。载于2018年IEEE 国际数据挖掘会议（ICDM）。IEEE，197–206。[8] Feng Li、Zhenrui Chen、Pengjie Wang、Yi Ren、Di Zhang 和 Xiaoyu Zhu。2019。赞助搜索中的点击率预测图意图网络。载于第42届国际 ACM SIGIR 信息检索研究与开发会议论文集。961–964。[9] Jianxun Lian、Xiaohuan Zhou、Fuzheng Zhang、Zhongxia Chen、Xing Xie 和 Guangzhong Sun。2018。xDeepFM：结合显式和隐式特征交互的推荐系统。载于第24届 ACM SIGKDD 国际知识发现与数据挖掘大会论文集。1754–1763。[10] Aaron van den Oord、Yazhe Li 和 Oriol Vinyals。2018。基于对比预测编码的表征学习。arXiv 预印本 arXiv:1807.03748（2018）。[11] Qi Pi、Weijie Bian、Guorui Zhou、Xiaoqiang Zhu 和 Kun Gai。2019。点击率预测中的长序列用户行为建模实践。载于第25届 ACM SIGKDD 国际知识发现与数据挖掘大会论文集。2671–2679。[12] Qi Pi、Guorui Zhou、Yujing Zhang、Zhe Wang、Lejian Ren、Ying Fan、Xiaoqiang Zhu 和 Kun Gai。2020。利用终身序列行为数据进行基于搜索的用户兴趣建模以实现点击率预测。载于第29届ACM信息与知识管理国际会议论文集。2685–2692。[13] Fei Sun、Jun Liu、Jian Wu、Changhua Pei、Xiao Lin、Wenwu Ou 和 Peng Jiang。2019。BERT4Rec：基于 Transformer 双向编码器表征的序列推荐。载于第28届 ACM 国际信息与知识管理会议论文集。1441–1450。[14] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Łukasz Kaiser 和 Illia Polosukhin。2017。注意力就是你所需要的一切。神经信息处理系统进展 30（2017）。[15] Chuhan Wu、Fangzhao Wu、Tao Qi、Qi Liu、Xuan Tian、Jie Li、Wei He、Yongfeng Huang 和 Xing Xie。2022。FeedRec：融合多种用户反馈的新闻信息流推荐。载于 ACM网络会议 2022论文集。2088–2097。[16] Ling Yan、Wu-Jun Li、Gui-Rong Xue 和 Dingyi Han。2014。展示广告中的网络规模点击率预测耦合组套索。载于国际机器学习会议。PMLR，802–810。[17] Yi Yang、Baile Xu、Shaofeng Shen、Furao Shen 和 Jian Zhao。2020。用于用户响应预测的操作感知神经网络。《神经网络》121（2020），161–168。[18] Weinan Zhang、Jiarui Qin、Wei Guo、Ruiming Tang 和 Xiuqiang He。2021。点击率估计的深度学习。载于第三十届国际人工智能联合会议论文集，IJCAI-21。4695–4703。[19] Yu Zheng、Chen Gao、Jianxin Chang、Yanan Niu、Yang Song、Depeng Jin 和 Yong Li。2022。解耦长期与短期兴趣以实现推荐。载于 ACM网络会议 2022论文集。2256–2267。[20] Guorui Zhou、Na Mou、Ying Fan、Qi Pi、Weijie Bian、Chang Zhou、Xiaoqiang Zhu 和 Kun Gai。2019。用于点击率预测的深度兴趣演化网络。载于AAAI人工智能会议论文集，第33卷。5941–5948。[21] Guorui Zhou、Xiaoqiang Zhu、Chenru Song、Ying Fan、Han Zhu、Xiao Ma、Yanghui Yan、Junqi Jin、Han Li 和 Kun Gai。2018。用于点击率预测的深度兴趣网络点击率预测。在第*24*届 *ACM SIGKDD* 国际知识发现与数据挖掘大会论文集。1059–1068。[22] Han Zhu、Xiang Li、Pengye Zhang、Guozheng Li、Jie He、Han Li 和 Kun Gai。2018。用于推荐系统的基于树的深度模型学习。在第*24*届 *ACM SIGKDD* 国际知识发现与数据挖掘大会论文集。1079–1088。