

## ■ 数据挖掘专题(主持人:谢娟英)

引用格式:朱瑞,张俊三,朱杰,等.一种基于 KCNN 和 MKR 的两阶段深度学习多任务推荐模型[J].陕西师范大学学报(自然科学版),2020,48(6):82-89.[ZHU R,ZHANG J S,ZHU J,et al. A two-stage deep learning multi-task recommendation model based on KCNN and MKR[J]. Journal of Shaanxi Normal University (Natural Science Edition), 2020, 48(6): 82-89.] DOI: 10.15983/j.cnki.jsnu.2020.01.020

# 一种基于 KCNN 和 MKR 的两阶段深度学习多任务推荐模型

朱 瑞<sup>1</sup>,张俊三<sup>1\*</sup>,朱 杰<sup>2</sup>,张世栋<sup>3</sup>

(1 中国石油大学(华东) 计算机科学与技术学院 山东 青岛 266580;

2 中央司法警官学院 信息管理系 河北 保定 071000;

3 国网山东电科院, 山东 济南 250003)

**摘 要:**深度学习和知识图谱的结合在推荐领域受到广泛关注,然而部分模型输入向量较为稀疏,不仅增加了整个模型的训练难度,还容易导致模型陷入局部最优;此外,大多推荐模型并未充分挖掘用户和物品的特征交互,使得用户和物品的向量表示不准确,影响最终的推荐模型性能。基于此,本文提出一种基于 KCNN 和 MKR 的两阶段深度学习多任务推荐模型 TMR。首先,利用文本卷积网络,提取物品名称的特征,将其转化为稠密向量,再结合物品自身属性,作为物品特征向量的初始化表示;其次,采用交替训练的方式,获取知识图谱中的辅助信息,再以 DeepFM 为特征提取层,挖掘用户(user)和物品(item)的特征交互。实验结果表明:与当前主流推荐方法相比,TMR 模型在准确度等评价指标上有很好的表现,提高了推荐系统的性能。

**关键词:**深度学习;推荐系统;知识图谱;多任务

**中图分类号:** TP391 **文献标志码:** A

**文章编号:** 1672-4291(2020)06-0082-08

**开放科学(资源服务)标识码(OSID):**

## A two-stage deep learning multi-task recommendation model based on KCNN and MKR

ZHU Rui<sup>1</sup>, ZHANG Junsan<sup>1\*</sup>, ZHU Jie<sup>2</sup>, ZHANG Shidong<sup>3</sup>

(1 College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, Shandong, China;

2 Department of Information Management, The National Police University for Criminal Justice, Baoding 071000, Hebei, China;

3 State Grid Shandong Electric Power Research Institute, Jinan 250003, Shandong, China)

**Abstract:** The combination of deep learning and knowledge graphs has attracted the attention of researchers in the field of recommender system. However, some models use sparse feature vectors as input which will not only increase the training difficulty of the model, but also cause the model to fall into a local optimum. In addition, most models can't fully describe the relationship between the items in the recommender system and the entities in the knowledge graphs. The vector representation is not accurate and it will affect the performance of the

**收稿日期:** 2020-06-01

**基金项目:**国家自然科学基金(61873280);河北省自然科学基金青年基金项目(F2018511002);中央司法警官学院校级科研项目(XYZ201602);河北省高等学校科学技术研究项目(Z2019037)

\* **通信作者:**张俊三,男,副教授,博士,主要从事信息检索、机器学习、数据挖掘等领域的研究。E-mail: zhangjunsan@upc.edu.cn

recommender system. Based on these issues, a multi-task recommendation model based on two-stage deep learning (TMR) is proposed. First, using KCNN to extract the features of the items' name, convert it into a dense vector, and then combine the property of the item as the initial representation of the item feature vector. Second, alternate training is used to obtain extra information in the knowledge graph, and then DeepFM is used as a feature extraction layer to mine the feature interaction between user and item. The experiments demonstrate that our proposed TMR model achieves better results in several evaluation indicators and improves the performance of the recommender system, comparing with other state-of-the-art methods.

**Keywords:** deep learning; recommender system; knowledge graph; multi-task

随着信息技术和互联网的发展,人们面临的问题不再是信息匮乏,而是如何筛选信息。推荐系统正是以海量数据为基础,面向用户建立的智能化平台,以服务用户为目标,为用户实现信息的个性化推荐<sup>[1]</sup>。所谓个性化,即根据用户的不同偏好推荐不同的商品信息,以帮助用户发现他们想要的服务。这不仅会提高用户体验,也会给社会带来巨大的效益<sup>[2]</sup>。

目前,很多研究人员将深度学习与传统的推荐算法相结合,并取得了不错的成果<sup>[3-5]</sup>。而基于深度学习的推荐模型在研究用户、商品或两者交互行为间的高阶特征<sup>[6]</sup>有很大优势,却难以捕获低阶特征交互,且如果输入向量较为稀疏,模型便容易陷入局部最优<sup>[7]</sup>。此外,只从用户行为交互记录难以挖掘到物品间的深层联系。因此,我们可以借助知识图谱(knowledge graph,KG)作为辅助信息来源,并在一定程度上解决冷启动问题<sup>[8]</sup>。

知识图谱以三元组的形式,构建了实体之间的关系,其包含的语义信息可以为推荐系统提供很大帮助,现已在多个推荐场景得到了广泛应用。而知识图谱特征学习(knowledge graph embedding, KGE)<sup>[9]</sup>旨在挖掘实体的潜在信息以及实体间的关联,最终将知识图谱中的实体和关系表示为低维向量,从而更好地应用到推荐领域。当前,有研究者在这一方面做过深入探究,并提出了 3 种学习方式,即依次学习、联合学习和交替学习。

对于依次学习(如 DKN 模型<sup>[10]</sup>)的方式,其主要问题是将推荐模块和知识图谱特征学习模块分离,无法挖掘物品(item)和实体(entity)之间的联系,使知识图谱没有发挥应有的作用;联合学习(如 RippleNet 模型<sup>[11]</sup>)与此正好相反,其主要思想是将推荐模块和知识图谱特征学习模块的目标融合并同步训练,但两个模块间的权重分配需要通过实验来

确定。因此,联合学习训练的难度会比较大。而交替学习(如 MKR 模型<sup>[12]</sup>)是将推荐系统和知识图谱特征学习看成两个分离但又相关的任务,并采用多任务学习的框架,使得两模块的信息互补。但在 MKR 中,模型的向量是随机初始化的,这会导致模型收敛较慢。此外,MKR 模型只关注高阶特征间的交互,较少研究低阶特征,使得用户和物品的特征向量表示不够准确。

针对以上问题,本文提出了一种推荐模型 TMR,该模型的主要创新表现在 3 个方面:(1)基于 MKR 模型,提出一种两阶段深度学习的多任务推荐模型 TMR,并用交替学习的方法进行模型的训练和优化,使整个推荐模型达到更优;(2)优化向量初始化方法,以文本卷积网络<sup>[13]</sup>提取物品名称特征,并结合物品自身的属性,得到稠密的物品向量,作为后续模型的初始化向量,从而加速模型的收敛;(3)通过 DeepFM 层<sup>[14]</sup>建模低阶、高阶特征交互对推荐性能的影响,充分挖掘用户和物品间的特征交互,也为推荐模块和知识图谱特征学习模块提供更多的共享特征。

## 1 相关工作

推荐系统就是预测用户对物品的评分,从而向用户推荐其喜爱的商品。如以今日头条为代表的新闻 APP 均在首页根据用户偏好推送不同内容的定制化新闻<sup>[15]</sup>。据数据科学中心统计,推荐系统为 Amazon 等的商务平台提高了很大一部分收入<sup>[16]</sup>。个性化推荐算法已成为新闻、视频、音频、电商等相关平台中不可缺少的一部分。

推荐系统经历了从基于内容的推荐到协同过滤的推荐,从基本的基于用户<sup>[17]</sup>和物品<sup>[18]</sup>的协同过滤到基于 model 的协同过滤<sup>[19]</sup>。深度学习的兴起使得推荐系统得到进一步的发展。在 2016 年,

YouTube 团队尝试将深度学习技术应用到推荐系统领域,其核心思想是将线性模型和 DNN 模型相融合,兼顾二者的优势,使得模型同时具备记忆能力和泛化能力,从而在特征提取层面达到更好的效果,最终提出了 Wide & Deep 模型<sup>[20]</sup>。然而,该模型的输入通过 One-Hot 编码得到,输入向量稀疏,不适合神经网络训练,增加了模型训练的成本。

此外,研究人员还提出了类似的推荐模型,其中应用最广泛的就是 DeepFM 模型。该模型主要关注如何挖掘用户交互行为下隐藏的特征组合,充分学习用户、物品特征间的交互,从而使得推荐系统的性能更优。

传统的推荐系统只是将用户的历史行为信息作为输入,在一定程度上会导致信息的稀疏性和系统冷启动问题<sup>[21]</sup>,而解决此类问题的常用方法就是在推荐算法中额外引入一些辅助信息,将其作为原信息的补充。此类辅助信息不仅可以更充分地描述物品信息,对于用户画像的建立也有很大意义。在各种辅助信息中,知识图谱作为一种新兴技术逐渐得到了广泛应用,在推荐领域大放异彩<sup>[22]</sup>。其中,基于交替学习的 MKR 模型的核心在于交叉压缩单元的设计,其主要利用了知识图谱中实体和推荐系统中物品的同一性,通过建立共享单元来连接推荐模块和知识图谱特征学习模块,从而获取更多的额外辅助信息,解决信息的稀疏性问题。其中,交叉操作包含物品  $v$  和与之相关的实体  $e$  每个可能的特征交互:

$$C_l = \begin{bmatrix} v_l^{(1)} e_l^{(1)} & \cdots & v_l^{(1)} e_l^{(d)} \\ \vdots & \cdots & \vdots \\ v_l^{(d)} e_l^{(1)} & \cdots & v_l^{(d)} e_l^{(d)} \end{bmatrix}. \quad (1)$$

其中,  $v_l$  和  $e_l$  分别表示第  $l$  层的物品和实体。压缩操作将交叉特征矩阵从  $\mathbf{R}^{d \times d}$  降维到  $\mathbf{R}^d$ ,并在下一层输出物品和实体的潜在特征表示:

$$v_{l+1} = v_l e_l^T w_l^{VV} + e_l v_l^T w_l^{EV} + b_l^V, \quad (2)$$

$$e_{l+1} = v_l e_l^T w_l^{VE} + e_l v_l^T w_l^{EE} + b_l^E, \quad (3)$$

其中,  $w_l \in \mathbf{R}^d$  和  $b_l \in \mathbf{R}^d$  分别是网络的权重和偏置。MKR 模型通过交叉压缩单元,实现了推荐系统中物品和知识图谱中实体的信息共享。但在训练用户向量时,对用户进行了随机初始化,并直接用全连接网络做特征提取。这不仅会忽略低阶特征间的交互,还可能导致网络收敛速度较慢,到达局部最优,

在一定程度上影响模型的性能。

综上,虽然目前将深度学习技术应用于推荐系统领域已取得了较大的成就,但仍可以发现部分模型的输入极大地影响模型整体的学习效果,且大多数推荐模型对用户和物品的特征提取还存在不足之处。因此,本文提出了 TMR 模型,将文本卷积网络和词频统计相结合,对各向量进行初始化;通过 DeepFM 层建模用户、物品的特征交互,并采用交替学习实现模型的全局优化。

## 2 TMR 模型

### 2.1 TMR 模型整体架构

TMR 模型在 MKR 模型的基础上,优化了模型的初始化,并以 DeepFM 作为特征提取层。TMR 模型通过交替训练推荐系统任务和知识图谱特征学习任务,以多任务框架进行模型的优化,其整体架构如图 1 所示。

以 MovieLens-1M 数据集为例,介绍 TMR 推荐模型的工作原理首先,模块①在读取电影数据集后,统计所有电影样本类型的频率,即每个电影类型对应一个词频;用每个类型相对应的词频来代替该类型,得到样本对应的向量(用词频统计得到的向量,可以在一定程度上体现电影的热门程度,即类型词频越高的电影种类越受用户欢迎)。在此基础上,对其做标准化处理,保证数据分布一致。此外,再用文本卷积网络处理电影名称,通过卷积层提取名称特征,通过池化层统一各样本向量维度,从而得到物品名称向量表示。将上述两部分向量进行连接,最终得到初始的物品特征向量并作为模块①的输出。根据用户物品交互记录,按评分对物品向量进行加权,得到对应的用户初始特征向量。模块②在获取了初始化的用户特征向量和物品特征向量后,采用多任务学习的框架,与模块③进行交替训练。用交叉压缩单元实现推荐系统中物品和知识图谱中实体之间的信息共享,补充稀疏的物品信息,并输出物品的潜在特征表示,将其作为最终的电影特征向量。同时,在模块④中用 DeepFM 层提取初始用户特征向量间的低阶、高阶特征交互,将最终的用户特征向量表示输出到模块②中,并与最终的电影特征向量计算内积,从向量相似度的角度对用户行为进行预测,并向用户推荐其偏好的电影,实现推荐的个性化。



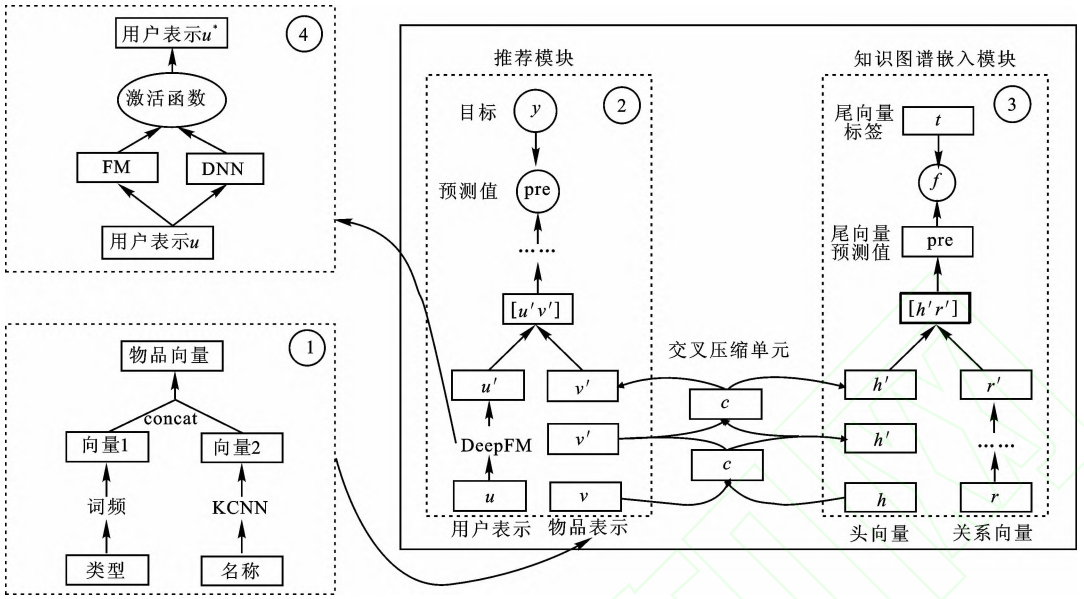


Fig. 1 Framework of the TMR model

图 1 TMR 模型的整体架构

## 2.2 模型预训练

在 TMR 模型中,为加速模型的收敛,防止模型陷入局部最优,要通过模型预训练对数据进行处理,即图 1 中的模块①。

以 MovieLens-1M 为例,对于数据集中的电影类别,统计各类别出现的次数,用类别词频代替原类别。但在本文中,没有直接使用替换后的词频作为输入。首先,由于每个电影所属的类别数不一样,因此需要对类别数进行缺失值补插,使得每个样本的维度保持一致。在填充样本维度时,计算每个维度的平均值,并在此基础上加一个随机值,保证各样本间的差异性,并增加模型的鲁棒性。在保证样本维度一致的情况下,再对样本做标准化处理,即

$$x = (x - \mu) / \sigma. \quad (4)$$

其中,  $\mu$  为当前维度的样本均值,  $\sigma$  为当前维度的样本方差。

其次,本文使用文本卷积网络提取电影名称的特征。在处理数据集时,首先需要将这一属性列中(如 Toy Story (1995))的时间字段去除,只保留电影名称字段(因为电影上映时间对提取电影特征没有帮助);然后,将电影名称输入网络:

$$W = [W_1 \ W_2 \ \cdots \ W_n], \quad (5)$$

其中:  $W$  为电影名称表示,  $W_n$  为电影名称中第  $n$  个单词的向量表示,  $n$  为当前电影名称的维度。对输

入向量进行卷积操作:

$$c_i = f(h \times w_{i:i+l-1} + b). \quad (6)$$

其中:  $h$  是卷积矩阵;  $l$  表示窗口大小,即每次卷积操作覆盖的单词数;  $b$  为偏置。通过卷积层后,将得到的特征图进行最大池化:

$$\tilde{c} = \max\{c_1, c_2, \dots, c_{n-l+1}\}. \quad (7)$$

从而保证电影名称特征向量维度一致。

然后,将电影类别向量和电影名称向量进行连接,喂入全连接层,对其权重参数进行 L2 正则化,防止隐藏层单元的自适应,从而减轻过拟合,得到最终的电影特征向量。

对于用户特征向量,根据用户与电影的交互记录,对相应的电影向量赋予权重。在本数据集中,评分全都分布在 1~5 分,且 50% 以上的评分都是 4 分或 5 分。因此,在对电影赋权重时,评分低于 4 分的赋予负值。经过加权便得到了用户特征向量。

预处理模块流程如算法 1 所示。

### 算法 1.

输入: 用户集合  $U = \{u_1, u_2, \dots, u_n\}$ 、电影集合  $M = \{m_1, m_2, \dots, m_n\}$ 、用户电影交互评分记录  $R = \{r_1, r_2, \dots, r_n\}$

输出: 用户特征向量、电影特征向量

for  $m \in M$  do

    取出 type 列,以词频代替类型得到 type 向量

根据式(4)对 type 矩阵标准化  
取出 title 列,根据式(5)–(7)得到 title 向量  
end for  
连接 type 向量和 title 向量得到 movie 特征向量  
for  $r \in R$  do  
    if  $\text{score} < 4$      $\text{score} = -\text{score}$   
end for  
for  $u \in U$  do  
     $u = \frac{1}{n} \sum_{(u,m) \in r} \text{movie} * \text{score}$   
end for

2.3 特征提取

TMR 模型在预训练阶段对物品提取特征,求出物品特征向量;但并未对用户向量提取特征。因此,本文选用 DeepFM 网络对用户预训练得到的用户特征向量提取特征,如图 1 中模块 ④ 所示。

DeepFM 模型由 FM 和 DNN 两部分构成,分别负责提取低阶、高阶特征。而在 TMR 模型中,这两部分的输入都是预训练阶段得到的特征向量,该向量不具备稀疏性,因此 DeepFM 就无需再做稀疏向量的嵌入。

DeepFM 模块的输出也由两部分构成,即  
 $\hat{y} = y_{\text{FM}} + y_{\text{DNN}}$ 。(8)

$y_{\text{FM}}$  负责提取一阶和二阶特征,即

$$y_{\text{FM}} = \langle w, x \rangle + \sum_{j_1=1}^d \sum_{j_2=j_1+1}^d \langle v_i, v_j \rangle x_{j_1} \cdot x_{j_2} \quad (9)$$

其中,  $w$  为线性模型部分的参数,  $v_i$  为特征向量  $x_i$  的辅助向量。

经过 DeepFM 层提取特征之后,用户向量便可以 and 物品特征向量进行运算,对用户进行电影的个性化推荐。

2.4 评分预测

推荐系统在获取了用户特征向量和物品特征向量后,便可以对用户进行行为预测。在相关推荐场景下,可以将问题抽象成回归问题,即将用户的行为看作是对电影的评分预测。本文选择的方法是直接计算用户向量和物品向量间的内积

$$\hat{y}_{uv} = \sigma(f_{\text{RS}}(u, v)), \quad (10)$$

即通过向量的相似度来进行用户的个性化推荐。其中,  $\sigma$  表示激活函数;  $f_{\text{RS}}$  表示推荐模型计算预测结果的函数,此处具体为内积;  $u$  表示用户向量;  $v$  表示

物品向量。

2.5 模型训练与参数学习

TMR 模型采用多任务框架交替学习的方法进行特征建模和推荐预测,损失函数由三部分组成,即推荐系统部分、知识图谱特征学习部分和正则化项,具体表示为

$$L = L_{\text{RS}} + L_{\text{KG}} + L_{\text{REG}} \quad (11)$$

推荐系统部分损失为

$$L_{\text{RS}} = \sum_{u \in U, v \in V} f(\hat{y}_{uv}, y_{uv}) \quad (12)$$

其中:  $u, v$  分别表示用户向量和物品向量,  $\hat{y}_{uv}$  是  $u, v$  经内积操作后,再通过激活函数的结果;  $f$  为交叉熵函数,此函数是收敛的。

知识图谱特征学习部分损失为

$$L_{\text{KG}} = \lambda_1 \left( \sum s(h, r, t) - \sum s(h', r, t') \right) \quad (13)$$

其中:  $h, r, t$  分别为知识图谱中的头结点表示、关系表示和尾节点表示,而  $h', t'$  是对知识图谱的负取样,  $s$  为知识图谱特征学习模块的得分函数,选取的计算方式是向量内积,该函数是收敛的。

正则化项为

$$L_{\text{REG}} = \lambda_2 \|w\|_2^2 \quad (14)$$

需要注意的是,  $\lambda_1, \lambda_2$  为超参数,其中  $\lambda_1$  的值为负,从而保证知识图谱部分的特征学习效果。此外, TMR 模型通过设置 Dropout<sup>[23]</sup> 参数的值,随机丢弃部分神经网络单元,在一定程度上防止了过拟合现象的出现。

3 实验与分析

3.1 数据集

MovieLens 数据集是一组从 20 世纪 90 年代末到 21 世纪初由 MovieLens 用户提供的电影评分数据。该数据主要包括电影名称、类型和年代,用户的年龄、性别、职业和邮编,以及用户对电影的评分。

本实验使用了其中的 MovieLens-1M 数据集,其主要包括 6 040 个用户、3 882 个电影以及 1 000 209 次用户和电影的交互记录评分。在该数据集中,用户对电影的评分大多集中在 4 分或 5 分,占比分别为 35% 和 22%。因此,在对评分数据进行处理时,将 4 分和 5 分对应的样本看为正样本,其余的

数据看成是负样本。

3.2 评价指标

本文采用准确率(accuracy, ACC) 和 ROC 曲线的线下面积(area under curve, AUC) 这两个指标来评价 TMR 模型,旨在测试该模型在公共数据集上的性能表现。

ACC 指标用于描述对于总体而言预测正确的分类比例,即

$$ACC = \frac{n_{correct}}{n_{total}}.$$
 (15)

其中,  $n_{correct}$  表示预测正确的记录个数,  $n_{total}$  表示全部测试数据的个数。

AUC 指标是对 ROC 曲线的量化。在某些情况下,由于阈值取值的问题,ROC 曲线并不一定是光滑的,很难判定模型的性能。因此,选用 AUC 对模型进行评价,而 AUC 的值就是 ROC 曲线与 FPR 轴线形成的面积。

3.3 实验结果与分析

3.3.1 常规数据集下不同推荐算法的性能对比

为了验证 TMR 模型的有效性,本文将 TMR 模型与当前一些具有代表性的、主流的推荐算法(PER<sup>[24]</sup>、RippleNet<sup>[11]</sup>、LibFM<sup>[25]</sup>、MKR<sup>[12]</sup> 等) 进行比较分析,最终结果取 3 次实验的平均值,如表 1 所示。

表 1 常规数据集下不同推荐算法的性能对比

Tab. 1 Performance comparison of different algorithms in common data set

模型	ACC	AUC
PER	0.667	0.712
RippleNet	0.842	0.920
LibFM	0.812	0.892
Wide&Deep	0.820	0.898
MKR	0.843	0.917
TMR	<b>0.852</b>	<b>0.926</b>

本文实验在相同的实验环境(2.8 GHz Intel Core i7-7700 处理器,16 GB 2400MHz DDR4 内存, NVIDIA GTX 1050 Ti 显卡) 中进行测试,共享模型输入,包括用户特征、电影特征、用户 - 电影交互评分记录以及知识图谱三元组。模型中的参数均通过实验训练得到,如推荐模块的学习率、知识图谱特征学习模块的学习率、特征向量的维度、batch 的大小等。

为从多角度评价 TMR 模型的性能,本文进行了 Top-K 场景实验,选取多组 K 值,并以 Precision 和 Recall 两个评价指标对比不同的推荐算法,其实验结果如图 2 所示。

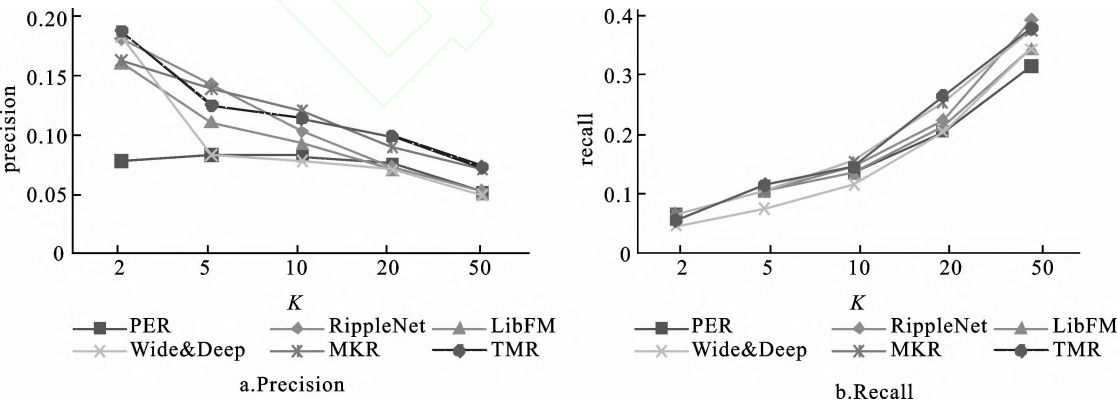


图 2 Top-K 推荐实验结果  
Fig. 2 The results of Top-K recommendation

从图 2 可以看出,TMR 模型在 Top-K 场景下也有不错的表现,总体上优于其他推荐算法。值得一提的是,RippleNet 模型和 MKR 模型在 Top-K 推荐场景下,也有不错的效果。这是因为 MovieLens-1M 数据集信息相对比较稠密,RippleNet 模

型在用户和物品交互记录密集的情况下,可以较为精确地捕获用户的兴趣,获得不错的性能表现;而 MKR 模型通过交叉单元,也可以补充物品信息,达到类似的效果。TMR 模型在共享用户和实体信息的基础上,又用 DeepFM 来建模用户和物品的特征

交互,使得推荐性能更高效。

3.3.2 稀疏数据集下不同推荐算法的性能对比

为了评估 TMR 模型在稀疏数据场景下的性能,本文对 MovieLens-1M 数据集进行了处理,通过设置

训练集比例  $r$  来模拟不同程度的稀疏数据场景,对比不同推荐算法在稀疏场景下的性能,对比试验结果如表 2 所示。

表 2 稀疏数据集下不同推荐算法的性能对比

Tab. 2 Performance comparison of different aalgorithms in sparse data set

模型	$r$									
	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%
PER	0.710	0.697	0.668	0.675	0.662	0.647	0.638	0.621	0.607	0.598(−11.2%)
RippleNet	0.920	0.912	0.901	0.890	0.878	0.870	0.862	0.859	0.851	0.843(−7.7%)
LibFM	0.892	0.886	0.875	0.864	0.850	0.837	0.829	0.816	0.810	0.801(−9.1%)
Wide&Deep	0.898	0.884	0.876	0.858	0.840	0.821	0.815	0.809	0.802	0.788(−11%)
MKR	0.917	0.913	0.908	0.903	0.897	0.889	0.882	0.881	0.874	0.868(−4.9%)
TMR	<b>0.926</b>	<b>0.916</b>	<b>0.913</b>	<b>0.910</b>	<b>0.908</b>	<b>0.905</b>	<b>0.897</b>	<b>0.884</b>	<b>0.881</b>	<b>0.879(−4.7%)</b>

在研究稀疏数据集的实验场景中,本文将训练集比例  $r$  从 100%更改到 10%,并保证验证集和测试集不变。从表 2 看出,随着数据集越来越稀疏,各种推荐算法的性能均有所下降。相对于  $r=100%$ ,当  $r=10%$ 时,各基线的 AUC 指标下降分别达到 11.2%、7.7%、9.1%、11%和 4.9%,而 TMR 模型降低了 4.7%,少于其他基线模型。这表明,在用户和物品交互稀疏的情况下,TMR 模型仍可以保持良好的性能。此外,RippleNet、MKR 和 TMR 推荐模型都引入了知识图谱作为辅助信息,使得数据稀疏场景下,模型性能仍较为稳定。

4 结语

本文提出一种基于 KCNN 和 MKR 的两阶段深度学习多任务推荐模型 TMR,主要解决推荐系统中向量初始化和特征提取的问题。首先,利用 KCNN 对物品进行特征提取,生成初始的物品特征向量,并作为推荐模块的输入。通过这种方法,使得后续 DNN 模型的输入不再是稀疏向量,有助于模型更快地收敛并达到最优。其次,利用 DeepFM 对用户进行特征提取,对其低阶、高阶特征交互建模。最后,对训练好的物品向量和用户向量计算内积,以向量相似度表示用户对物品的喜好程度。实验结果表明,在通用数据集和稀疏数据集上,本文方法较其他主流推荐算法都有明显优势。

由于推荐系统情境较为复杂,因此,在实际应用中仍有许多问题需要深入探讨。例如,在表示用户

向量时,可以挖掘用户社交网络中的信息,从而更准确地建模用户画像,这对于解决用户冷启动也有很大作用;其次,对于物品的特征提取,用户的评论也会有很大帮助,这将涉及 NLP 的相关内容。

参考文献:

[1] 曾义夫,牟其林,周乐,等. 基于图表示学习的会话感知推荐模型[J]. 计算机研究与发展, 2020, 57(3): 590-603.

[2] PAUL COVINGTON, JAY ADAMS, EMRE SARGIN. Deep neural networks for youtube recommendations[C]//Proceedings of the 10th ACM Conference on Recommender Systems. Boston: ACM, 2016: 191-198.

[3] LEI H, ZHANG Y. Learning tree structure in multi-task learning[C] //Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney: ACM, 2015: 397-406.

[4] JAMALI M, ESTER M. A matrix factorization technique with trust propagation for recommendation in social networks[C] //Proceedings of the 4th ACM conference on Recommender systems. New York: ACM, 2010: 135-142.

[5] WANG H X, LIANG G H, ZHANG X M. Feature regularization and deep learning for human resource recommendation[J]. IEEE Access, 2018, 6: 39415-39421.

[6] WU H, ZHANG Z X, YUE K, et al. Dual-regularized matrix factorization with deep neural networks for recommender systems [J]. Knowledge-Based Systems, 2018, 145: 46-58.

[7] 王瑞琴,吴宗大,蒋云良,等. 一种基于两阶段深度学习的集成推荐模型[J]. 计算机研究与发展, 2019, 56(8): 1661-1669.



- [8] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [9] PEROZZI B, ALRFOU R, SKIENA S. DeepWalk: online learning of social representations [C] //Proceedings of the 20th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2014: 701-710.
- [10] WANG H W, ZHANG F Z, XIE X, et al. DKN: deep knowledge-aware network for news recommendation [C] //Proceedings of the 27th World Wide Web Conference on World Wide Web. Lyon: ACM, 2018: 1835-1844.
- [11] WANG H W, ZHANG F Z, WANG J L, et al. RippleNet: propagating user preferences on the knowledge graph for recommender systems [C] //Proceedings of the 27th ACM International Conference on Information and Knowledge Management. Torino: ACM, 2018: 417-426.
- [12] WANG H W, ZHANG F Z, ZHAO M, et al. Multifask feature learning for knowledge graph enhanced recommendation [C] //Proceedings of the 28th World Wide Web Conference on World Wide Web. San Francisco: ACM, 2019: 2000-2010.
- [13] KIM Y. Convolutional neural networks for sentence classification [C] //Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha, 2014: 1746-1751.
- [14] GUO H F, TANG R M, YE Y M, et al. DeepFM: a factorization-machine based neural network for CTR prediction [C] //Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: ACM, 2017. arXiv:1703.04247v1.
- [15] 黄震华, 张佳雯, 田春岐, 等. 基于排序学习的推荐算法研究综述[J]. 软件学报, 2016, 27(3): 691-713.
- [16] MIYAHARA K, PAZZANI M J. Collaborative filtering with the simple Bayesian classifier[J]. PRICAI 2000 Topics in Artificial Intelligence, 2000. DOI: 10.1007/3-540-44533-1\_68.
- [17] 张宜浩, 朱小飞, 徐传运, 等. 基于用户评论的深度情感分析和多视图协同融合的混合推荐方法[J]. 计算机学报, 2019, 42(6): 1316-1333.
- [18] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [19] SOCHER R, CHEN D Q, MANNING C D, et al. Reasoning with neural tensor networks for knowledge base completion [C] //Proceedings of the 26th Int Conf on Neural Information Processing Systems (NIPS' 13). New York: Curran Associates, 2013: 926-934.
- [20] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems [C] //Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. New York: ACM, 2016. DOI: 10.1145/2988450.2988454.
- [21] 高玉凯, 王新华, 郭磊, 等. 一种基于协同矩阵分解的用户冷启动推荐算法[J]. 计算机研究与发展, 2017, 54(8): 1813-1823.
- [22] JAMEEL S, BOURAOUI Z, SCHOCKAERT S. MEMBER: max-margin based embeddings for entity retrieval [C] //Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2017: 783-792.
- [23] XIAO Y, XIANG R, et al. Personalized entity recommendation: a heterogeneous information network approach [C] //Proceedings of the 7th ACM International Conference on Web Search and Data Mining. New York: ACM, 2014: 283-292.
- [24] XIAO Y, XIANG R, SUN Y Z, et al. Personalized entity recommendation: a heterogeneous information network approach [C] //Proceedings of the 7th ACM International Conference on Web Search and Data Mining. 2014: 283-292.
- [25] RENDLE S. Factorization machines with libFM[J]. ACM Transactions on Intelligent Systems and Technology, 2012, 3(3): 1-22.

[责任编辑 宋轶文]