

# 基于评论和物品描述的深度学习推荐算法

王美玲 刘晓楠 尹美娟 乔猛 荆丽娜

数学工程与先进计算国家重点实验室(信息工程大学) 郑州 450002

(751882186@qq.com)

**摘要** 评论文本中蕴含着丰富的用户和物品信息,将其应用于推荐算法有助于缓解数据稀疏问题,提高推荐准确度。然而,现有的基于评论的推荐模型对评论文本的挖掘不够充分和有效,并且大多忽视了用户兴趣随时间的迁移和蕴含物品属性的物品描述文档,使得推荐结果不够准确。基于此,文中提出了一种基于深度语义挖掘的推荐模型(Deep Semantic Mining based Recommendation,DSMR),通过深度挖掘评论文本和物品描述文档的语义信息,更精确地提取用户特征和物品属性特征,从而实现更准确地推荐。首先,所提模型利用BERT预训练模型来处理评论文本和物品描述文档,深度挖掘用户特征和物品属性,有效缓解了数据稀疏和物品冷启动问题;然后,利用前向LSTM来关注用户偏好随时间产生的变化,得到了更精确的推荐;最后,在模型训练阶段,将实验数据按1~5分1:1:1:1:1等量随机抽取,保证每个分值的数据量相等,使结果更加准确,模型鲁棒性更强。在4个常用的亚马逊公开数据集上进行实验,结果表明,以均方根误差为评价指标,DSMR推荐结果的误差比2个仅基于评分数据的经典推荐模型至少平均降低了11.95%,同时优于基于评论文本的3个最新推荐模型,且比其中最优的模型平均降低了5.1%。

**关键词:**推荐算法;深度学习;评论文本;物品描述;数据稀疏性;冷启动

**中图分类号** TP391

## Deep Learning Recommendation Algorithm Based on Reviews and Item Descriptions

WANG Mei-ling, LIU Xiao-nan, YIN Mei-juan, QIAO Meng and JING Li-na

State Key Laboratory of Mathematical Engineering and Advanced Computing (Information Engineering University), Zhengzhou 450002, China

**Abstract** Reviews contain rich user and item information, which helps to alleviate the problem of data sparsity. However, the existing recommendation model based on reviews is not sufficient and effective enough to mine the review texts, and most of them ignore the migration of user interest over time and the item description documents containing the item attribute, which makes the recommendation result not accurate enough. In this paper, a deep semantic mining based recommendation model (DSMR) is proposed. By mining the semantic information of review texts and item description documents in depth, user characteristics and item attributes can be extracted more accurately, so as to realize more accurate recommendation. Firstly, the BERT pre-training model is used to process the comment text and item description document, and the user characteristics and item attributes are excavated deeply, which effectively alleviated the problems of data sparse and item cold start. Then, the forward LSTM is used to pay attention to the change of user preferences over time, and more accurate recommendations are obtained. Finally, in the model training stage, the experimental data are randomly selected from 1 to 5 points at 1:1:1:1:1 to ensure the same amount of data for each score value, so as to make the results more accurate and the model more robust. Experiments on four commonly used Amazon open datasets show that the root mean square error (RMSE) of DSMR is at least 11.95% lower than the two classical recommendation models based only on rating data, and it is better than the three new recommendation models based only on review text, and 5.1% lower than the optimal model.

**Keywords** Recommendation algorithm, Deep learning, Review, Item description, Data sparsity, Cold start

### 1 引言

推荐系统自诞生以来受到了极大的关注,研究者提出了很多优秀的算法来提高推荐的效率和准确性。深度学习是在传统推荐算法的基础上应用深度学习模型来挖掘深层次的用户偏好特征,使得推荐准确性进一步提高。早期的算法主要

利用评分数据进行推荐,随着用户数量和物品数量的剧增,数据稀疏性和冷启动等问题越来越突出,成为限制推荐精度进一步提高的主要原因。电子商务在带来大量商品信息的同时也产生了大量的评论信息,评论中包含了用户对商品的功能、质量等方面是否满意的信息<sup>[1]</sup>。充分利用评论信息可以准确地获得用户的喜好以及商品全面的属性,有效缓解数据

稀疏和冷启动问题,使推荐更加准确。

起初,研究者尝试将评论文本用于主题建模<sup>[2-9]</sup>,实现了比仅利用评分数据的模型更高的预测准确度。但这种方式只关注主题线索,忽略了语义内容,而且通常将评论表示为词袋,忽略了上下文信息<sup>[10]</sup>,因此限制了预测准确度的进一步提高。近年来,很多研究开始将深度学习与评论文本相结合,提出了许多优秀的算法,并得到了比基于主题建模方法精确度更高的推荐结果。文献[11-14]将多条评论连接成一个长文档,利用卷积神经网络(Convolutional Neural Network, CNN)从评论文本中学习有用的特征。然而,基于文档的建模将所有评论无差别地连接到同一个文档中,没有区分不同评论的不同重要程度,不利于提取有效特征<sup>[15]</sup>。于是,研究者开始使用基于评论的建模方式,即对每一条评论单独进行建模,最后将每条评论的特征聚合为总特征。文献[15-17]都是基于评论建模的,并且都利用了注意力机制来区分不同评论的重要程度,得到了比基于文档建模的模型更高的推荐准确度。

综上,我们注意到目前许多工作存在的局限性:1)很多模型依然使用 CNN 来提取评论中的用户和物品特征,只能捕捉局部特征,对于长序列文本的特征无法有效提取,在一定程度上限制了推荐准确度的提升。2)在基于评论的模型中,很多工作都没有考虑用户的兴趣和偏好会随时间变化<sup>[11-16]</sup>,而是同等看待过去的偏好与近期的偏好。3)上述提到的利用评论文本来提高推荐准确度的优秀模型都没有在利用评论文本的同时也重视对物品描述文档的使用,物品描述文档中包含着对物品属性比较全面的介绍,对于缓解物品冷启动有着十分重要的作用。4)对于训练数据,现有方法没有考虑不同分值的数量差别很大,4分和5分的评分占很大比重,训练结果对于低分值数据是不公平的,容易造成过拟合,模型鲁棒性差。为了解决这些问题,我们提出了基于评论文本和物品描述的深度学习推荐模型。

本文的工作可以总结为以下3点:

(1)使用谷歌提供的预训练好的 BERT<sup>[18]</sup> 模型(bert\_base\_uncase)代替 CNN 来处理评论文本,克服了 CNN 只能提取局部特征的弱点,而且能较为准确地捕捉词在不同语境中的语义,度量不同评论对用户特征的贡献程度,并结合前向长短期记忆模型(Long Short-Term Memory, LSTM)来学习用户随时间变化的兴趣迁移,提高了推荐准确度。很多模型选择双向循环神经网络(Recurrent Neural Network, RNN)来处理数据,但对于我们的模型,语义信息已经由 BERT 学习得到,我们只期望 LSTM 学习用户兴趣随时间的变化。由于只有已有的评论可以影响未来的评论,未来的评论并不能对已有的评论产生影响,并且后向 LSTM 在学习兴趣迁移方面效果不佳,只会增加模型的复杂度,因此我们不采用。

(2)将物品描述文档与评论一起引入模型中,帮助我们更好地描述物品特征,提高预测准确度,并且在新物品缺少评论时,物品描述文档可以很好地缓解物品冷启动问题。

(3)对于实验数据,我们将 1~5 分这 5 个分值的评论数据按 1:1:1:1:1 进行随机抽样,确保每个分值的数据量相等,

以减轻过拟合并提高模型的鲁棒性。

在 4 组公开数据集上进行对比实验,结果表明,我们的基于深度语义挖掘的推荐模型 DSMR 的预测评分准确度高于目前最优秀的基于评论文本的模型,如 DeepCoNN<sup>[10]</sup>, NARRE<sup>[15]</sup>, DER<sup>[17]</sup> 等。

## 2 相关工作

近年来,深度学习在自然语言处理、计算机视觉等领域的成功使得推荐领域开始注意到这个强大的工具,学者们开始探索利用深度学习方法来改善目前推荐系统的一些难以克服的弱点,如数据稀疏、冷启动、可解释性差等问题<sup>[19-20]</sup>。特别是 CNN 和 RNN 的出现,在很多自然语言处理(Natural Language Processing, NLP)任务中取得了巨大成功。于是大家开始尝试用深度学习方法,如 DeepCoNN, D-Attn<sup>[12]</sup> 等,挖掘评论文本中用户的偏好、商品的特征,从而直接应用于预测评分。DeepCoNN 由两个以 CNN 为基本模型的并行神经网络组成,分别学习用户和物品的隐表示,在网络顶部将两部分连接起来学习交互,证明了评论文本对于缓解稀疏问题的有效性。

注意力机制<sup>[21]</sup>的关键是去学习一个权重来标识重要程度,其自被提出后就被广泛应用于自然语言处理,在机器翻译<sup>[22-23]</sup>、阅读理解<sup>[24-25]</sup>、语音识别<sup>[26]</sup>等领域都取得了最先进的结果<sup>[27]</sup>。于是注意力机制受到了推荐领域的关注,开始被用于基于评论的推荐算法<sup>[12, 15-16, 28]</sup>。NARRE<sup>[15]</sup>使用注意力机制来学习不同评论的有用性,更好地建模用户和物品,预测物品评分和生成解释。不同于 D-Attn 词级别的注意力机制, NARRE 采用的是评论级别的注意力机制。MPCN<sup>[16]</sup>受 Transformer<sup>[29]</sup>的启发,不使用 RNN 和 CNN,完全依赖 Attention 机制,提出了一个新的基于指针的学习方案,使用户和物品之间可以进行深层次的文字交互,并取得了很好的效果。

NLP 的发展对评论文本在推荐领域的应用有很大的促进作用。预训练语言模型<sup>[14]</sup>自被提出后快速发展,产生了很多优秀的方法,如基于特征的 ELMo<sup>[30]</sup>和基于微调的 OpenAI GPT<sup>[31]</sup>。但这些语言模型在本质上都是单向的,限制了预训练的表达能力。于是文献[18]提出了一个双向的预训练模型 BERT,该模型利用 Transformer 的 Encoder 一次性读取整个文本,使模型能够基于单词的两侧进行学习,从而较为准确地把握单词在句子中所表达的含义。因此, BERT 就有了天然的双向性,泛化能力较强,为下游任务提供了很好的基础。

## 3 DSMR 模型

### 3.1 模型框架

每个用户都会购买很多商品,并对许多商品进行评论,因此我们可以将评论作为用户偏好的一种表示。但对于用户而言,物品的描述同样重要,因为用户只有受到物品描述的吸引才会选择浏览此物品并查看此物品收到的评论;此外,对于一个新物品,还没有或者很少被购买和评价,而物品描述中提供了

丰富的物品属性信息,有助于解决物品冷启动问题。很多模型在利用文本进行建模时都只利用了评论文本,而没有重视物品描述文档,我们认为这会损失一部分重要信息,因此我们将物品描述也输入模型中,以得到更准确的预测结果。

DSMR 利用 BERT 预训练模型来处理文本数据并区分不同评论的重要性,从而帮助我们更准确地预测一个用户对一件商品的评分。DSMR 模型的结构如图 1 所示。模型分为并行的两个部分,一个是用户模型,一个是物品模型。在用户模型,输入为该用户评论过的所有物品的描述文档和每个物品收到的所有评论;在物品模型,输入为此物品收到的所有评论和对此物品的描述。最后将两个模块所得到的结果进行点积,得到该用户对此物品的预测打分。由于用户模型和物品模块的结构类似,因此本文以用户模块为例来详细介绍我们的模型。

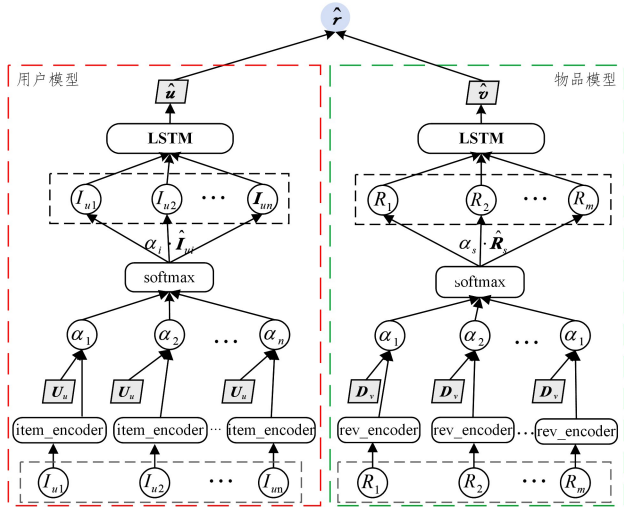


图 1 DSMR 模型整体框架

Fig. 1 DSMR framework

## 3.2 实现细节

### 3.2.1 Encoder

对于某个用户  $u$ , 他评论过的所有物品用  $I_{ui}$  ( $i=1, 2, \dots, n$ ) 来表示。将  $I_{ui}$  传入 item\_encoder 模块中。item\_encoder 的具体结构如图 2 左侧框所示, 其中  $\oplus$  表示相加。在 item\_encoder 模块中, 将物品  $I_{ui}$  的描述文档  $D_i$  和物品  $I_{ui}$  收到的所有评论  $R_{ij}$  ( $j=1, 2, \dots, m$ ) 传入 BERT 中。我们的对比模型 NARRE 利用 CNN 进行评论文本的处理, 只能对输入序列建立短距离依赖, 而 Transformer 中的 Self-attention 可以通过动态生成不同连接的权重来处理变长的信息序列, 并且可以实现并行化, 提高训练速度。

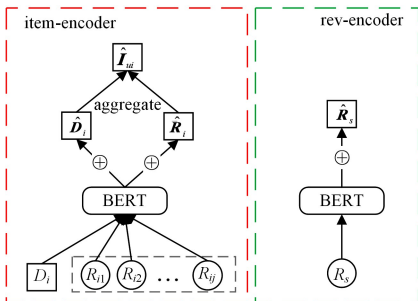


图 2 编码部分

Fig. 2 Encoder part

物品描述文档  $D_i$  经过 BERT 预训练后得到物品描述的词向量表示, 将各词向量相加得到  $\hat{D}_i$ ; 将物品的各条评论经过 BERT 预训练后得到各条评论的隐表示, 将各隐表示相加得到  $\hat{R}_i$ , 合并  $\hat{D}_i$  与  $\hat{R}_i$  后得到物品嵌入向量  $\hat{I}_{ui}$ ,  $\hat{I}_{ui}$  刻画了物品  $i$  的特征。公式如下:

$$\hat{D}_i = \text{BERT}(D_i) \quad (1)$$

$$\hat{R}_i = \text{sum}(\text{BERT}(R_{i1}, R_{i2}, \dots, R_{im})) \quad (2)$$

$$\hat{I}_{ui} = \hat{D}_i \odot \hat{R}_i \quad (3)$$

其中,  $\odot$  表示将两向量进行拼接。

对于物品  $v$ , 它所收到的所有评论用  $R_s$  ( $s=1, 2, \dots, m$ ) 来表示, 评论经过 BERT 模型后得到评论隐表示  $\hat{R}_s$ , 如图 2 右侧 rev\_encoder 部分。

### 3.2.2 LSTM

我们利用词嵌入将用户  $id$  表示为用户嵌入向量  $U_u$  ( $u=1, 2, \dots, d$ ),  $d$  为总的用户数量。将  $U_u$  映射到与物品嵌入向量  $\hat{I}_{ui}$  同一空间中进行点积运算, 得到用户  $u$  与物品  $i$  的特征的关联程度  $\alpha_i$ ,  $\alpha_i$  的值越大, 说明关联程度越高, 用户对物品就越感兴趣。

$$\alpha_i = U_u \cdot \hat{I}_{ui}, i=1, 2, \dots, n \quad (4)$$

将  $\alpha_i$  ( $i=1, 2, \dots, n$ ) 通过 softmax 进行归一化, 并将归一化后的  $\alpha_i$  与  $\hat{I}_{ui}$  相乘, 得到每个物品对用户特征的贡献程度。最后, 将  $\hat{I}_{ui}$  送入 LSTM 学习用户随时间产生的兴趣迁移, 得到用户模型的输出向量  $\hat{u}$ 。

$$\hat{u} = \text{LSTM}(\text{softmax}(\alpha_i) \cdot \hat{I}_{ui}), i=1, 2, \dots, n \quad (5)$$

同理, 我们将物品  $v$  的描述文档表示为  $D_v$ , 并将  $D_v$  与物品评论嵌入向量  $\hat{R}_s$  映射到同一空间进行运算, 可得物品模型的输出向量  $\hat{v}$ 。

### 3.2.3 评分预测

将用户模型的输出向量  $\hat{u}$  与物品模型的输出向量  $\hat{v}$  做点积得到最终的预测评分  $\hat{r}$ 。

$$\hat{r} = \hat{u} \cdot \hat{v} \quad (6)$$

### 3.2.4 模型学习

DSMR 模型的目标实际上就是要提高评分预测的准确度, 这相当于一个回归问题。对于回归问题, 最常用的目标函数是平方损失函数。在训练集样本  $M$  中, 用户  $u$  对物品  $i$  的预测评分为  $\hat{r}_{ui}$ , 真实评分为  $r_{ui}$ , 则目标函数可表示为:

$$L = \sum_{u,i \in M} (\hat{r}_{ui} - r_{ui})^2 \quad (7)$$

我们的任务是最小化目标函数。我们选择 Adam<sup>[32]</sup> 优化算法来优化目标函数, 因为 Adam 使用动量和自适应学习率来加快收敛速度, 适合数据量较大的问题且只需要很少的内存。

## 4 实验

### 4.1 数据集

在数据集的选取上, 我们参考文献[15-16]中最先进的

模型所使用的数据集,并从中选取了亚马逊的公开数据集中较常用的4个数据集作为我们的数据:Movies\_and\_TV,Toys\_and\_Games,Kindle\_Store以及Videos\_Games数据集。基本数据统计信息如表1所列。

表1 数据集

Table 1 Statistics of datasets

数据集	用户数	物品数	评论数
Movies_and_TV	123960	50052	1679533
Toys_and_Games	19412	11924	167597
Kindle_Store	68223	61935	982619
Videos_Games	24303	10672	231780

在对数据集进行处理的过程中,我们考虑到虽然评分有1~5分5种分值,但5分和4分仍然占评分的大多数,几乎所有已提出的模型都没有考虑这种情况。我们认为这对于1分或2分的数据是不公平的,会使训练结果过拟合。我们将1~5分这5种分值的数据按1:1:1:1:1进行随机抽取,这样数据集中各分值的数据是等量的,结果更加客观,模型鲁棒性更强。

#### 4.2 对比模型

为了验证模型的有效性,我们选择了2个早期的仅利用评分矩阵的经典模型和3个最近提出的利用评论文本的先进模型作为对比模型。

MF<sup>[33]</sup>:矩阵分解是非常流行的基于协同过滤的推荐方法。它只使用评级矩阵作为输入,用用户和物品低秩矩阵的内积来表示评分,采用交替最小二乘(ALS)技术对其目标函数进行最小化。

PMF<sup>[34]</sup>:概率矩阵分解是一种传统的矩阵分解方法,只使用评分数据进行协同过滤,引入高斯分布对用户和物品的潜在因素进行建模。

DeepCoNN:以CNN为基本模型,由两个并行神经网络组成,其中一个并行网络利用用户评论集学习用户行为,另一个并行网络利用物品评论集学习物品属性。位于两个神经网络顶部的额外共享层将两个并行网络连接起来,使学习到的用户和项目潜在因素可以交互预测评级。此模型证明了利用评论文本可以有效缓解稀疏性问题。

NARRE:在DeepCoNN的基础上利用注意力机制来判断一条评论的贡献程度,通过选择更有用的评论进行建模来提高模型的准确度和可解释性。

DER:与前两个模型类似,DER也是通过CNN来提取物品属性。此外,DER认为传统的GRU没有考虑在较大的时间间隔后用户的兴趣会发生改变,因此提出通过添加一个时间门来改进GRU,以此来更准确地预测用户当前的喜好。

此外,我们还设置了一个对比模型review-DSMR,基于本文提出的DSMR但只利用评论文本、不添加物品描述文档的推荐模型,以此来验证物品描述文档对推荐效果的促进作用。

#### 4.3 评价指标

我们采用广泛应用于算法性能评价的均方根误差(Root Mean Square Error, RMSE)作为评价指标,公式如下:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{R}_{u,i} - R_{u,i})^2} \quad (8)$$

其中, $N$ 为测试集中样本的数量, $\hat{R}_{u,i}$ 为用户 $u$ 对物品 $i$ 的

预测评分, $R_{u,i}$ 为用户 $u$ 对物品 $i$ 实际评分。RMSE的值越小,说明模型的性能越好。

#### 4.4 实验过程与参数设置

数据随机排序后,70%作为训练集,20%作为验证集,10%作为测试集。我们使用的BERT预训练模型为谷歌训练好的bert\_base\_uncase,review-DSMR和DSMR模型的初始学习率为0.01,之后使用NoamOpt优化器来进行动态调整。损失率设置为[0.05,0.1,0.3,0.5],批量大小设置为[3,5,8,16,32],潜在因子数量设置为[32,64,128,256]。

对于MF和PMF,我们分别根据文献[33]和文献[34]的设置策略,使用网格搜索从[25,50,100,150,200]中寻找潜在因子的最佳值,从[0.001,0.01,0.1,1,10]中寻找正则化参数的最佳值。对于DeepCoNN和NARRE,我们分别按照文献[10,15]的设置进行复现,学习率为[0.005,0.01,0.02,0.05],批量大小为[50,100,150],损失率为[0.1,0.3,0.5,0.7,0.9],潜在因子数量为[8,16,32,64]<sup>[15]</sup>;对于CNN文本处理器,卷积层的神经元数量为100,窗口大小为3。对于对比模型DER,设置的学习率为[0.001,0.01,0.1,1],批量大小为[50,100,150],用户/物品的嵌入尺寸在[8,16,32,64,128]中进行调整。

为了验证对训练数据进行1:1:1:1的等量控制能够提高算法的准确度,我们将所有模型分别在不做等量控制和做等量控制两种情况下进行实验,观察实验结果。

#### 4.5 实验结果与分析

经过多次实验,DSMR模型在损失率为0.1、批量大小为5、潜在因子数量为128时效果最好。

各模型的实验结果如表2、表3所列。

表2 无数据等量控制性能对比(RMSE)

Table 2 Performance comparison without data equal

	control(RMSE)			
	Movies_and_TV	Toys_and_Games	Kindle_Store	VideosEOS_Games
MF	1.522	1.379	1.286	1.503
PMF	1.276	1.158	1.102	1.311
DeepCoNN	1.193	1.044	1.025	1.231
NARRE	1.147	1.008	0.976	1.192
DER	1.106	0.983	0.942	1.145
review-DSMR	<b>1.098</b>	<b>0.977</b>	<b>0.913</b>	<b>1.115</b>
DSMR	<b>1.073</b>	<b>0.935</b>	<b>0.884</b>	<b>1.097</b>

表3 数据等量控制性能对比(RMSE)

Table 3 Performance comparison with data equalization

	control(RMSE)			
VideosEOS_Games	Movies_and_TV	Toys_and_Games	Kindle_Store	
MF	1.357	1.239	1.208	1.415
PMF	1.122	1.026	0.974	1.206
DeepCoNN	1.107	0.993	0.955	1.154
NARRE	1.075	0.974	0.937	1.141
DER	1.049	0.954	0.902	1.109
review-DSMR	<b>1.035</b>	<b>0.921</b>	<b>0.874</b>	<b>1.083</b>
DSMR	<b>1.017</b>	<b>0.897</b>	<b>0.839</b>	<b>1.058</b>

为了使实验结果更直观,我们将表2和表3制作成了图3(以Toys\_and\_Games数据集为例),来展示有无数据等量控制的性能差异,并将表3以柱状图的形式展示出来(见图4),方便后文做各方面的分析(以Movies\_and\_TV为例)。

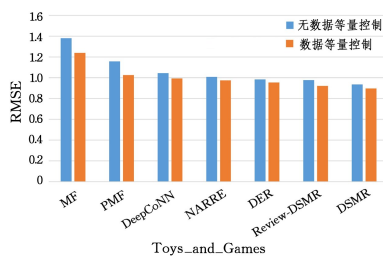


图3 有无数据等量控制效果对比

Fig. 3 Effect comparison with/without data equal control

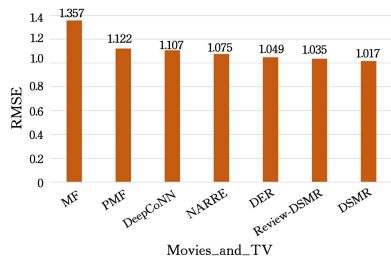


图4 数据等量控制性能对比

Fig. 4 Performance comparison with data equal control

从图3可以看出,对所有模型的1~5分这5种分值的训练数据按1:1:1:1:1等量抽取后,所有模型的RMSE相比不做数据等量控制时均有所降低,证明对训练数据进行等量控制有助于提升推荐效果。因为不作处理的数据低分值的评论很少,4分和5分的数据占大多数,模型容易过拟合,所以等量处理后,模型鲁棒性更强。

从表3可以看出,在对所有模型都做了数据等量控制的情况下,DSMR模型依然比之前最先进的模型的效果更好, RMSE比MF,PMF,DeepCoNN,NARRE和DER模型在4个数据集上分别平均降低了26.98%,11.95%,9.46%,7.66%和5.1%。

下面从各模型所基于的信息与提取的特征对实验结果的影响进行分析。首先,利用评论文本的模型比传统的仅利用评分数据的模型效果都更好,从图4可以看出,DeepCoNN,NARRE,DER和DSMR的RMSE均比MF和PMF低,这证明了评论数据对模型学习到更准确的用户特征和物品属性是有益的,对推荐准确性的提高的确有促进作用。

其次,对于同样都考虑了评论文本的模型,使用注意力机制的模型比没有使用注意力机制的模型效果更好,如NARRE,DER和DSMR的RMSE均低于DeepCoNN,因为注意力机制可以学习到各个评论对于用户特征和物品属性的贡献程度,所以以不同的权重来利用不同的评论比将所有评论不加区分地利用效果更好。

再者,使用BERT预训练方法的模型比使用静态词向量的模型效果更好,如DSMR的RMSE低于DeepCoNN和NARRE,因为BERT能够学习到单词在不同上下文信息中的不同含义,而静态词向量则不能,所以这就使得两者特征提取的效果相差很大。再加上DeepCoNN,NARRE和DER都利用CNN来提取特征,而CNN只能学习局部特征,对于长序列的信息损失较大。此外,DSMR模型在利用评论文本的同时还引入了物品描述文档,而其他模型都忽视了这一点。物品描述文档不仅丰富了物品属性

信息,还缓解了物品冷启动问题。

最后,使用LSTM方法来探索用户偏好随时间变化的模型比没有关注用户偏好随时间变化的模型表现更好,如DER(使用GRU)和DSMR的RMSE均低于DeepCoNN和NARRE。因为用户偏好并不是一成不变的,过去不喜欢战争片的用户现在可能也会喜欢战争片,过去不受欢迎的电影可能成为现在人们眼中的经典影片,所以时间变化是一个很重要的因素,注重时间变化带来的影响有助于提高预测准确度。

从图5 review-DSMR和DSMR实验结果的比较中可以看出,添加了物品描述文档的DSMR模型的RMSE明显比仅利用评论文本的review-DSMR模型更低,证明物品描述文档的确有利于丰富物品属性,从而提高推荐准确度。

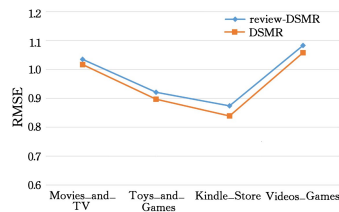


图5 是否利用物品描述信息效果对比

Fig. 5 Effect comparison with/without item description

**结束语** 本文基于评论文本和物品描述,提出了能更准确预测评分的深度学习挖掘推荐模型DSMR,它利用BERT预训练模型来学习单词在上下文信息中较为准确的语义并评估评论的重要程度,还利用LSTM来学习评论间的内在联系,探索用户偏好随时间的变化,同时引入物品描述文档来缓解物品冷启动问题,并在实验数据处理中采用各分值等量抽取来提高模型的鲁棒性。实验结果表明,DSMR模型在预测评分准确度上比目前最新的基于评论文本的推荐模型高5.1%。

对于未来的研究,我们将注重推荐的可解释性问题。推荐的可解释性也是提高推荐效果的一个很重要的方面,有说服力的、贴切的推荐理由会提高用户的信任度。如何利用评论文本生成推荐理由是未来很值得期待的一个研究方向。

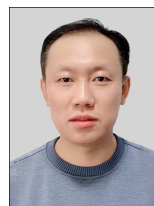
## 参考文献

- [1] KIM D,PARK C,OH J,et al. Convolutional matrix factorization for document context-aware recommendation[C]// Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016:233-240.
- [2] WANG C,BLEI D M. Collaborative Topic Modeling for Recommending Scientific Articles[C]// Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,2011:21-24.
- [3] MCAULEY J,LESKOVEC J. Hidden factors and hidden topics: understanding rating dimensions with review text[C]// Proceedings of the ACM Conference on Recommender Systems. ACM, 2013:165-172.
- [4] BAO Y,FANG H,ZHANG J. Topicmf: Simultaneously exploiting ratings and reviews for recommendation[C]// Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. AAAI Press,2014:2-8.
- [5] TAN Y,ZHANG M,LIU Y,et al. Rating-boosted latent topics:

- Understanding users and items with ratings and reviews[C]// Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. 2016;2640-2646.
- [6] LING G, LYU M R, KING I. Ratings meet reviews, a combined approach to recommend[C]// Proceedings of the ACM Conference on Recommender Systems (RecSys). ACM, 2014; 105-112.
- [7] CATHERINE R, COHEN W. Transnets; Learning to transform for recommendation[C]// Proceedings of the 11th ACM Conference on Recommender Systems. ACM, 2017; 288-296.
- [8] BLEI D M, NG A, JORDAN M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3(4/5): 993-1022.
- [9] LEE D D, SEUNG H S. Algorithms for Non-negative Matrix Factorization[C]// International Conference on Neural Information Processing Systems. MIT Press, 2000; 556-562.
- [10] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation[C]// Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. ACM, 2017; 425-434.
- [11] KIM D, PARK C, OH J, et al. Convolutional Matrix Factorization for Document Context-Aware Recommendation[C]// ACM Conference. ACM, 2016; 233-240.
- [12] SEO S, HUANG J, YANG H, et al. Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction[C]// The Eleventh ACM Conference. ACM, 2017; 297-305.
- [13] WU L, QUAN C, LI C, et al. A context-aware user-item representation learning for item recommendation[J]. ACM Transactions on Information Systems (TOIS), 2019, 37(2): 1-29.
- [14] DAI A M, LE Q V. Semi-supervised Sequence Learning[J]. MIT Press, 2015.
- [15] CHEN C, ZHANG M, LIU Y, et al. Neural attentional rating regression with review-level explanations[C]// Proceedings of the 2018 World Wide Web Conference. 2018; 1583-1592.
- [16] TAY Y, LUU A T, HUI S C. Multi-pointer co-attention networks for recommendation[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018; 2309-2318.
- [17] CHEN X, ZHANG Y, QIN Z. Dynamic Explainable Recommendation Based on Neural Attentive Models[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 53-60.
- [18] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [J]. arXiv:1810.04805, 2018.
- [19] CAO S, YANG N, LIU Z. Online news recommender based on stacked auto-encoder[C]// ACIS 16th International Conference on Computer and Information Science (ICIS). IEEE, 2017; 721-726.
- [20] WANG H, WANG N, YEUNG D Y. Collaborative Deep Learning for Recommender Systems[C]// KDD 2015. ACM, 2015: 1235-1244.
- [21] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. arXiv:1409.0473, 2014.
- [22] GEHRING J, AULI M, GRANGIER D, et al. Convolutional sequence to sequence learning[J]. arXiv:1705.03122, 2017.
- [23] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. Computer Ence, 2014.
- [24] HERMANN K M, KOCISKY T, GREFFENSTETTE E, et al. Teaching machines to read and comprehend[C]// Advances in Neural Information Processing Systems. MIT Press, 2015; 1693-1701.
- [25] SEO M, KEMBHAVI A, FARHADI A, et al. Bidirectional attention flow for machine comprehension[J]. arXiv:1611.01603, 2018.
- [26] AMODEI D, ANANTHANARAYANAN S, ANUBHAI R, et al. Deep Speech 2: End-to-End Speech Recognition in English and Mandarin[C]// ICML. 2015.
- [27] LU Y, DONG R, SMYTH B. Coevolutionary recommendation model; Mutual learning between ratings and reviews[C]// Proceedings of the 2018 World Wide Web Conference. 2018; 773-782.
- [28] CHEN J, ZHANG H, HE X, et al. Attentive Collaborative Filtering; Multimedia Recommendation with Item-and Component-Level Attention [C] // International ACM Sigir Conference. ACM, 2017; 335-344.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv:1706.03762, 2017.
- [30] PETERS M, NEUMANN M, IYYER M, et al. Deep Contextualized Word Representations[C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018.
- [31] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding with unsupervised learning [R]. Technical report, OpenAI, 2018.
- [32] KINGMA D, BA J. Adam: A Method for Stochastic Optimization[J]. arXiv:1412.6980, 2014.
- [33] KOREN Y, BELL R, VOLINSKY C. Matrix Factorization Techniques for Recommender Systems [J]. Computer, 2009, 42(8): 30-37.
- [34] SALAKHUTDINOV R, MNIH A. Probabilistic matrix factorization[C]// Proceedings of the 20th International Conference on Neural Information Processing Systems. 2007; 1257-1264.



**WANG Mei-ling**, born in 1995, post-graduate. Her main research interests include deep learning and recommendation algorithm.



**LIU Xiao-nan**, born in 1977, Ph.D, associate professor, master's supervisor, is a member of China Computer Federation. His main research interests include quantum algorithm and high-performance parallel computation.