

融合评分矩阵和评论文本的深度学习推荐模型

王 艳, 彭 治, 杜永萍
(北京工业大学 信息学部, 北京 100124)

摘要:用户对项目评分数据的稀疏性是影响推荐质量的主要因素之一,提出了融合评分数据和评论文本的深度学习模型,通过引入辅助信息缓解评分数据稀疏性的影响。利用评论文本可以获取用户的偏好信息和项目特征,而评分数据中又包含了用户和项目之间的潜在关联。现有的融合模型对评分数据的处理大多数都是采用矩阵分解方法,为了更好地利用评分数据中的有效信息,文中利用卷积神经网络处理评论文本,并引入注意力机制提取评论信息中具有代表性的评论,从而更好地表征用户偏好和项目特征。利用深度神经网络处理评分数据提取其中的深度特征,将特征进行融合来预测用户对项目的评分。文中在 Amazon 数据集上进行验证,以均方误差 MSE 作为评价指标,结果表明所提出的模型优于多个优秀的基线模型。

关键词:推荐系统;评论文本;评分矩阵;注意力机制;深度学习

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2021)08-0013-06

doi:10.3969/j.issn.1673-629X.2021.08.003

A Deep Learning Recommendation Model Combining Rating Matrix and Review Text

WANG Yan, PENG Zhi, DU Yong-ping

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: The sparsity of user rating data is one of the main factors affecting recommendation quality. A deep learning model combining rating data and comment text is proposed to mitigate the impact of rating data sparsity by introducing auxiliary information. The review text can be used to obtain the user's preference information and item characteristics, and the rating data contains the potential association between the user and the item. Most of the existing fusion models deal with rating data by using matrix factorization methods. In order to make better use of the effective information in the rating data, the convolutional neural network is used to process the comment text, and the attention mechanism is introduced to extract the representative comments from the comment information, so as to better characterize the user preferences and project characteristics. The deep neural network is used to process the score data to extract the deep features which are fused to predict the user's score on the item. It is verified on the Amazon data set with the mean square error (MSE) as the evaluation index. The results show that the proposed model is better than many excellent baseline models.

Key words: recommendation system; review text; rating matrix; attention mechanism; deep learning

0 引言

随着网络的飞速发展,给人们的生活带来了巨大的影响,在提供便利的同时也产生了很多问题。由于数据量成指数增长,人们想要从海量的数据中获取有用信息变得越来越困难^[1],因此推荐系统应运而生。推荐系统在很多领域都得到了广泛的应用,例如:医疗健康^[2]、交通出行^[3]、学习教育^[4]、电子商务^[5]等。近年来,推荐系统在解决信息过载方面被证明是有效的,可以帮助人们快速地在海量数据中找到自己需要的信息;因此得到了学术界和工业界的广泛关注。然而如

何更高效地提高推荐系统的准确性是现在推荐领域的研究热点。

传统的推荐算法主要有协同过滤^[6-7]和基于内容的推荐算法^[8]。这两种方法都是基于用户的历史行为来进行推荐的。基于内容的推荐算法^[8]利用用户对项目的评分信息来确定这些物品之间的共同特征,若是新物品也有这个特征,就会将该物品推荐给用户;协同过滤^[6]是利用历史数据找到与该用户喜好相同的用户,然后将找到的这些相似用户过去喜欢的物品推荐给该用户。或者是找到和某一物品具有相同特征的物

收稿日期:2020-10-12

修回日期:2021-02-22

基金项目:国家重点研发计划(2019YFC1906002)

作者简介:王 艳(1993-),女,硕士研究生,研究方向为推荐系统;杜永萍,硕导,博士,研究方向为自然语言处理。

品,若用户喜欢这一物品则将其其他具有相同特征的物品推荐给用户。

基于近邻的推荐算法广泛用于协同过滤中。该算法利用评分信息计算用户或物品的相似度,根据和某一用户相似的用户对目标商品的评分来预测该用户对目标商品的评分,或者根据用户对和目标商品相似的物品的评分来预测该用户对目标商品的评分。基于近邻的方法易于实现,并且对于预测结果有一个直观的解释。

在协同过滤中,矩阵分解是最成功应用最广泛的一种方法。Koren 等人^[9]最早提出了基于矩阵分解(matrix factorization, MF)的推荐算法,该算法在评分矩阵的基础上,加入了隐向量,并将评分矩阵分解为用户矩阵和项目矩阵乘积的形式,增强了模型对稀疏矩阵的处理能力,从而取得了很好的推荐效果。Salakhutdinov 等人^[10]提出了概率矩阵因子分解(probabilistic matrix factorization, PMF)模型,该模型通过添加用户、项目隐式特征的高斯概率分布来改进算法。实验研究表明,基于矩阵分解的算法在评分预测方面要好于基于近邻的推荐算法。

虽然基于矩阵分解的推荐算法在推荐领域取得了不错的推荐效果,但是数据稀疏性对推荐系统性能的影响仍是当下急需解决的问题。一些研究表明^[11]引入评论文本信息能有效缓解数据稀疏性对推荐效果的影响,在一定程度上提高算法推荐的精确度。在推荐系统中,用户除了可以直接进行评分外还可以写评论信息,一般评论信息中会包含用户给出这个评分的理由,相较于评分而言,评论信息包含的内容更加丰富。一方面可以得到用户给出这个评分的原因,另一方面也可以体现用户的偏好信息或商品的一些特征。

最初使用评论文本进行建模的推荐算法大都是基于主题模型,Devid 等人^[12]提出了 LDA(latent Dirichlet allocation)模型,该模型用于推测文本中的主题分布,然后利用推测的主题分布获取用户、项目的的相关信息从而提高推荐的准确度。LDA 是主题模型研究领域最具有影响力的模型。Ganu 等人^[13]提出从评论文本主题中推导出基于文本的评分方法。然后,使用聚类技术将文本中的主题和情感相似的用户分为一组。McAuley 等人^[14]提出将评论文本和评分数据相结合的 HFT 模型,将用户评论集或商品评论集作为输入,将评论文本主题和评分数据中的隐因子结合。HFT 是将评论文本和评分数据相结合的最经典的模型。基于主题模型在推荐效果上取得了很好的结果,但是该方法不能很好地保留文本信息中的词序信息,从而忽略了上下文之间的逻辑对应关系^[15]。

近年来,深度学习在很多领域被广泛应用,特别是

在自然语言处理^[16]方面更取得了突破性的进展。利用自然语言在文本内容挖掘方面的优势,将其运用到推荐系统中对评论文本进行处理,这也为推荐系统提供了新的研究方向^[11];目前在推荐系统中使用的深度学习技术主要有 CNN^[17]、RNN^[18]、注意力机制^[19-20]、RBM^[21]、自编码器^[22]等。

深度学习中的 CNN、RNN 网络在处理文本信息时能够很好地保留词序信息,近年来逐渐取代了主题模型。Kim 等人^[23]提出的 ConvMF(convolution matrix factorization)模型将卷积神经网络和概率矩阵分解相结合,利用卷积神经网络处理评论文本信息,从而有效地保留词序信息。但是该模型仅仅利用了项目的评论文本信息,而忽略了用户的评论信息。随后 Zheng 等人^[24]提出了 DeepCoNN(deep cooperative neural network)模型,该模型利用两个并行的 CNN 来分别处理项目的评论信息和用户的评论信息,在最后一层利用因子分解机来进行评分预测。Chen 等人^[25]提出的 NARRE(neural attentional regression model with review level explanations)模型,将评分矩阵和评论文本作为输入,同时融合评分数据中的隐向量,极大地提高了推荐系统的预测性能。

上述模型中在与评论文本特征融合时仅用到了评分的潜在特征,没有对评分数据进行深度建模,学习评分数据的深度特征。因此文中提出了融合评分矩阵和评论文本的深度学习推荐模型,利用深度学习技术提取评分中的深度特征,并和评论信息中提取的用户偏好、项目特征进行融合。通过更好地对用户项目特征进行表示,从而提升推荐性能。用 MSE 作为评价指标,在亚马逊 6 个不同领域的公开数据集上进行对比实验,结果表明所提出的算法优于目前多个优秀的公开模型。

1 模型结构

文中提出了融合评分矩阵和评论文本的深度学习推荐模型 RMRT(rating matrix and review text based recommendation model),模型结构如图 1 所示。该模型包括两个部分,左侧部分对用户进行建模,提取用户偏好信息;右侧部分对项目进行建模,提取项目特征。将提取的用户特征和项目特征融合后送入预测层,得到预测评分。

在用户、项目建模过程中又分别包含了两个部分:(1)从评论文本中提取特征,输出用户、项目的深层特征表示为 T_u 、 T_i ;(2)从评分矩阵中提取用户项目特征,输出为 X_u 、 X_i ;将得到的用户特征和项目特征分别送入交互层得到用户、项目的最终特征表示,最终在预测层得出用户对项目的预测评分。在对用户、项目

建模时所用网络结构相同,只有输入不同。因此文中 重点介绍项目建模部分,用户建模过程与之相似。

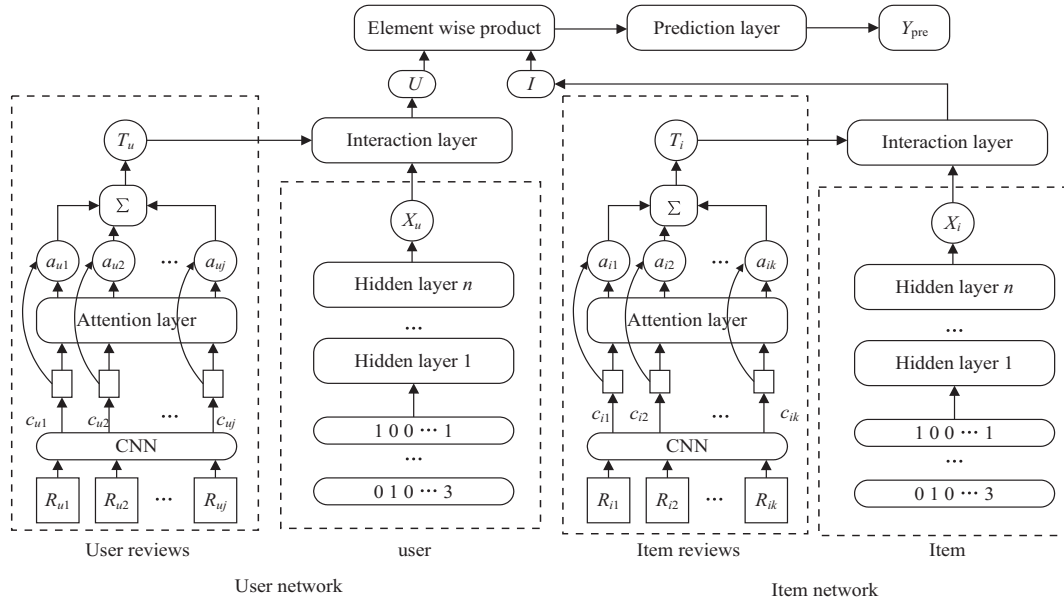


图 1 融合评分矩阵和评论文本的深度学习推荐模型

1.1 文本特征提取模块

用户对项目的评论信息能够反映出项目的特征信息。给定某项目 i 的评论集 $\{R_{i1}, R_{i2}, \dots, R_{ik}\}$, k 表示模型允许输入的最大评论数,经过嵌入层后,每条评论信息被映射为 d 维的向量,得到固定长度的词嵌入矩阵。嵌入层之后为卷积神经网络层,该层以嵌入层得到的词向量矩阵作为输入,得到特征向量。设卷积层由 m 个神经元组成,使用大小为 s 的滑动窗口上的卷积滤波器来提取上下文特征,具体如公式(1):

$$k_j = \text{Relu}(M_i * f_j + b_j) \quad (1)$$

式中, $*$ 表示卷积操作, b_j 表示偏差。

进入卷积模块中的最大池化层, max-pooling 能够捕获重要的具有高价值的特征,并且将卷积层的输出压缩成一个固定大小的向量,即:

$$c_j = \max(k_1, k_2, \dots, k_{d-s+1}) \quad (2)$$

卷积层最终输出的特征向量可表示为: $c_{i1}, c_{i2}, \dots, c_{ik}$ 。

对同一项目来说,不同用户会有不同的评论信息,有些用户的评论准确描述了项目的相关特性,而有些用户写的评论可能和项目没有太大的关联性。不同评论信息所作的贡献是不同的。为了选择更能代表项目特征的评论信息,引入了注意力机制^[20],对相应的评论文本使用一个注意力层,得到注意力得分:

$$g_{ik} = \text{Relu}(w_i \times c_{ik} + b) \quad (3)$$

式中, Relu 表示激活函数, w_i, b 分别为网络自动学习的权重和偏移量。

使用 softmax 函数对注意力得分 g_{ik} 进行归一化处理,得到最终的评论权重 a_{ik} 如下式所示:

$$a_{ik} = \frac{\exp(g_{ik})}{\sum_{l=1}^k \exp(g_{il})} \quad (4)$$

利用评论权重来突出当前评论集中第 n 条评论对项目 i 的特征的影响,将评论的注意力分数 a_{ik} 一一作用在对应的卷积层输出的特征向量上,得到项目 i 特征向量的加权和:

$$C_i = \sum_{l=1}^k a_{il} c_{il} \quad (5)$$

然后将其送入到全连接层,得到项目 i 基于文本信息的最终特征表示:

$$T_i = W_i C_i + b_i \quad (6)$$

式中, W_i, b_i 分别是全连接层的权重和偏差量。同理对用户建模可以得到用户 u 基于评论信息的特征表示 T_u 。

1.2 评分数据处理模块

用户对项目的评分数据是一种显性反馈,可以直接反映出用户对项目的喜爱程度,不同的评分说明用户对项目的喜好程度不同,因此利用神经网络根据评分数据获取用户的偏好和项目的特征表示。

以项目 i 的评分数据作为输入 $p_i = \{y_{i1}, y_{i2}, \dots, y_{im}\}$,其中网络的模型定义为:

$$\begin{cases} X_i^1 = f(w_i^1 p_i + b_i^1) \\ X_i^2 = f(w_i^2 X_i^1 + b_i^2) \\ \dots \\ X_i = f(w_i^n X_i^{n-1} + b_i^n) \end{cases} \quad (7)$$

其中, w_i^x 表示第 x 层的权重矩阵, b_i^x 表示第 x 层的偏置向量, $f(\cdot)$ 表示激活函数,这里使用的是 Relu 函数,最后一层的输出 X_i 表示项目 i 的深度特征表示。

同理对用户建模可以得到用户 u 基于评分数据的特征表示 X_u 。

1.3 特征融合

在本层对文本特征提取模块和评分数据处理模块得到的用户特征和项目特征进行交互,通过线性和高阶交互建模来获取项目的最终特征表示:

$$I = (T_i + X_i) \oplus g \quad (8)$$

特征向量 I 中包含了从评论信息和评分数据中获取的项目特征,其中 \oplus 表示拼接, g 表示高阶交互建模。此处借鉴了因子分解机中的二阶项来获取高阶特征,首先将因子分解机中的二阶项进行转换:

$$\sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle z_i z_j = \text{sum} \left\{ \frac{1}{2} \times \left[\left(\sum_{i=1}^n v_i z_i \otimes \sum_{i=1}^n v_i z_i \right) - \sum_{i=1}^n (v_i z_i \otimes v_i z_i) \right] \right\} \quad (9)$$

其中, \otimes 表示元素积,在执行 $\text{sum}()$ 之前,就已经包含了二阶交互项的所有信息。因此,只要求和前的部分,将评论信息中得到的特征和评分数据中得到的特征进行拼接 $z = T_i \oplus X_i$,利用下面的网络结构来获取高阶特征:

$$g = \frac{1}{2} \times \left[\left(\sum_{i=1}^n v_i z_i \otimes \sum_{i=1}^n v_i z_i \right) - \sum_{i=1}^n (v_i z_i \otimes v_i z_i) \right] \quad (10)$$

g 能进一步增强潜在特征的交互,同理可以得到用户的最终特征表示 U 。

1.4 评分预测

首先将上述得到的用户最终特征 U ,项目最终特征 I 进行元素积操作,得到一个向量 $H = \text{multiply}(U, I)$,将其传递到预测层得到预测评分 $\hat{R}_{u,i}$:

$$\hat{R}_{u,i} = WH + b_u + b_i + b_g \quad (11)$$

式中, W 表示预测层的权重参数, b_u 、 b_i 分别表示用户的偏差量、项目的偏差量, b_g 表示全局偏差量。

为了使预测评分 $\hat{R}_{u,i}$ 和数据集中的真实评分 $R_{u,i}$ 之间的差值最小,下面给出损失函数 loss 的表达式:

$$\text{loss} = \sum_{u,i \in O} (\hat{R}_{u,i} - R_{u,i})^2 + \lambda \| \omega \|^2 \quad (12)$$

式中, O 表示训练集中的样本数, $\hat{R}_{u,i}$ 、 $R_{u,i}$ 分别表示预测评分和真实评分。为了防止过拟合,损失函数中使用了 L2 范数正则项约束,其中 λ 用来控制正则化的程度,使用 Adam 优化函数最小化 loss 来得到模型中的相关参数。

2 实验结果与分析

2.1 数据集与实验设置

在实验中,使用了 6 个不同领域的公开数据集来

评估文中提出的模型,这些数据集来自 Amazon 5-core,分别为:Toys_and_Games、Office_Product、Musical_ents、Instant_Video、Digital_Music、Kindle_Store,下面简称(TG,OP,MI,IV,DM,KS)。这些数据集中主要包含了以下内容:用户对项目的评分信息(1-5分)、用户对项目的评论信息、用户 ID 和项目 ID。具体统计信息如表 1 所示。

表 1 数据集基本信息

数据集	用户数	项目数	评论 & 评分	稀疏度/%
TG	19,412	11,924	167,597	92.75
OP	4,798	2,419	52,673	99.55
MI	1,397	900	10,216	99.19
IV	4,902	1,683	36,486	99.56
DM	5,426	3,568	64,475	99.67
KS	68,223	61,934	982,619	97.67

在实验中将数据集划分为训练集、测试集和验证集,所占比重分别为 80%、10%、10%。其中,验证集主要是用来调整参数,在测试集上对模型的性能进行评估。

在实验过程中,使用 Google 新闻上训练的 300 维的 word2vec 向量来初始化词嵌入并在训练的过程中进行微调。使用 Adam 函数优化目标函数,默认初始化学率设置为 0.005,卷积核大小设置为 3。为了防止过拟合,在实验中采用了 dropout 策略,将值设置为 0.5。

2.2 评价指标

在实验中,模型的性能评价指标选用了均方误差 MSE,该值越小表示模型的性能越好,计算如下:

$$\text{MSE} = \frac{1}{k} \sum_{u,i} (\hat{R}_{u,i} - R_{u,i})^2 \quad (13)$$

式中, k 表示样本的个数, $\hat{R}_{u,i}$ 表示预测评分, $R_{u,i}$ 表示真实评分。

2.3 对比模型

为了验证文中提出模型的有效性,选择了基于评分和评论数据相关的推荐模型进行对比,具体如下所示:

- MF^[9]:矩阵分解模型,这个模型中只用到了评分数据,是协同过滤中最常使用的方法。

- PMF^[10]:概率矩阵分解模型,该模型在推荐建模时用到了高斯分布,来获取用户和项目的隐特征。

- HFT^[14]:该模型结合评分数据和评论文本信息,在建模时有两种组合方式,用户评论文本和评分 HFT(user),项目评论文本信息和评分 HFT(item)。文献中显示 HFT(item)的效果要优于 HFT(user),文中在对比时 HFT 指的是 HFT(item)。

- ConvMF^[23]:该模型结合卷积神经网络和概率矩阵分解,利用 CNN 来处理评论文本信息,保留了文本中的词序,但是仅仅只用到了项目的评论信息。

- DeepCoNN^[24]:深度协同神经网络模型,是首个同时利用了用户和项目的评论文本信息的深度学习模型,使用两个并行的卷积神经网络来处理文本数据,并在融合层利用因子分解机进行评分预测,取得了很好的结果。

- NARRE^[25]:该模型在建模时用到了评分数据和评论文本,并在 DeepCoNN 的基础上引入了 attention 机制,来识别评论的有用性,预测性能有一定的提高。

2.4 实验结果分析

使用亚马逊数据集进行对比实验,在 6 个不同的数据集上对不同模型的性能进行了比较,实验结果如表 2 所示。

表 2 在 Amazon 数据集上不同模型的结果对比(MSE)

模型	TG	OP	MI	IV	DM	KS
MF	1.406	1.787	1.471	1.205	1.169	1.502
PMF	1.144	1.092	1.401	-	1.211	0.983
HFT	1.048	1.617	1.378	0.951	1.049	1.341
ConvFM	0.862	0.960	0.991	1.199	1.084	-
DeepCoNN	0.801	0.739	0.908	0.934	0.809	0.673
NARRE	0.784	0.717	0.903	0.919	0.793	0.631
RMRT(ours)	0.773	0.709	0.869	0.909	0.789	0.626

由上面的对比结果可以看出,基于评分的推荐模型,由于数据稀疏性问题其推荐性能要差于基于评论文本的推荐模型。融合评分矩阵和评论文本的推荐模型性能要优于单独采用评分数据或者评论数据模型的推荐性能。因为评论能够反映用户的偏好,项目的特征,而评分数据仅反映了用户的偏好情况。引入文本信息能够在一定程度上缓解数据稀疏性对性能的影响;一般情况下,基于深度学习的推荐模型在推荐性能上优于常规的推荐模型,因为深度学习可以更好地学习特征表示。从表 2 结果可以看出 RMRT 在不同的数据集上均取得了优于其他模型的效果。

2.5 评分处理模块有效性验证

为了验证模型中利用深度学习技术对评分处理模块的有效性,将评分部分的处理方式替换为其他模型中常用的矩阵分解,其他部分保持不变,将替换后的方法表示为 RMRT-D+MF。在亚马逊数据集上进行对比实验,结果如图 2 所示,可以看出利用深度学习技术处理评分数据得到了更优的 MSE 值。

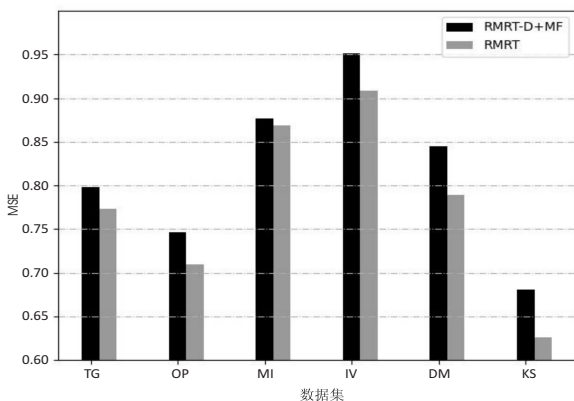


图 2 评分处理模块有效性验证性能比较

3 结束语

提出了融合评分矩阵和评论文本的深度推荐模型 RMRT。该模型可以同时从评分数据和评论文本信息中学习用户和项目的特征。在处理评分矩阵时利用深度学习算法进行深度特征学习,提取用户和项目的特征表示,将其和评论信息中提取的特征进行融合。对比实验结果表明,文中提出的方法在不同的数据集上均取得了优于其他模型的效果,因此提取深度特征能够在一定程度上提高预测的准确性。

文中工作主要关注的是评分数据和评论文本信息的特征提取,没有考虑时间因素及文本情感对推荐性能的影响,在今后的工作中将考虑引入时间变化对用户偏好及商品特征的影响,以及潜在特征的权重和文本中的情感信息对推荐性能的提升。

参考文献:

- [1] 黄立威,江碧涛,吕守业,等. 基于深度学习的推荐系统研究综述[J]. 计算机学报,2018,41(7):1619-1647.
- [2] SUBRAMANIASWAMY V, MANOGARAN G, LOGESH R, et al. An ontology-driven personalized food recommendation in IoT-based healthcare system[J]. The Journal of Supercomputing, 2018, 75(6):3184-3216.
- [3] HUANG Z H, SHAN G X, CHENG J J, et al. TRec: an efficient recommendation system for hunting passengers with deep neural networks[J]. Neural Computing and Applications, 2018, 31:209-222.
- [4] CHIEN T C, CHEN Z H, CHAN T W. Exploring long-term behavior patterns in a book recommendation system for reading[J]. Journal of Educational Technology & Society, 2017,

- 20(2):27–36.
- [5] GRBOVIC M, RADOSAVLJEVIC V, DJURIC N, et al. E-commerce in your inbox: product recommendations at scale [C]//Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining. Sydney, Australia; ACM, 2015:1809–1818.
- [6] 冷亚军, 陆青, 梁昌勇. 协同过滤推荐技术综述[J]. 模式识别与人工智能, 2014, 27(8):720–734.
- [7] 文俊浩, 袁培雷, 曾骏, 等. 基于标签主题的协同过滤推荐算法研究[J]. 计算机工程, 2017, 43(1):247–252.
- [8] 单京晶. 基于内容的个性化推荐系统研究[D]. 长春: 东北师范大学, 2015.
- [9] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8):30–37.
- [10] MNIH A, SALAKHUTDINOV R R. Probabilistic matrix factorization[C]//Proceedings of the advances in neural information processing systems. Vancouver, Canada: [s. n.], 2008:1257–1264.
- [11] LING G, LYU M R, KING I. Ratings meet reviews, a combined approach to recommend[C]//Proceedings of the 8th ACM conference on recommender systems. [s. l.]: ACM, 2014:105–112.
- [12] BLEI D M, NG A Y, JORDAN M I. Latent Dirichlet allocation[J]. J. Mach. Learn. Res., 2003, 3:993–1022.
- [13] GANU G, KAKODKAR Y, MARIAN A. Improving the quality of predictions using textual information in online user reviews[J]. Information Systems, 2013, 38(1):1–15.
- [14] MCAULEY J, LESKOVEC J. Hidden factors and hidden topics: understanding rating dimensions with review text[C]//Proceedings of the 7th ACM conference on recommender systems. Hong Kong, China; ACM, 2013:165–172.
- [15] WANG S, MANNING C D. Baselines and bigrams: simple, good sentiment and topic classification[C]//Proceedings of the 50th annual meeting of the association for computational linguistic. Jeju Island, Korea; ACL, 2012:90–94.
- [16] COLLOBERT R, WESTON J. A unified architecture for natural language processing; deep neural networks with multitask learning[C]//Proceedings of the 25th international conference on machine learning (ICML '08). Helsinki, Finland; ACM, 2008:160–167.
- [17] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6):1229–1251.
- [18] 高茂庭, 徐彬源. 基于循环神经网络的推荐算法[J]. 计算机工程, 2019, 45(8):198–202.
- [19] 王永贵, 尚庚. 融合注意力机制的深度协同过滤推荐算法[J]. 计算机工程与应用, 2019, 55(13):8–14.
- [20] AL-SABAHI K, ZUPING Z, NADHER M. A hierarchical structured self-attentive model for extractive document summarization (HSSAS) [J]. IEEE Access, 2018, 6:24205–24212.
- [21] SALAKHUTDINOV R, MNIH A, HINTON G. Restricted Boltzmann machines for collaborative filtering [C]//Proceedings of the 24th international conference on machine learning. Corvalis, USA: [s. n.], 2007:791–798.
- [22] WU Y, DUBOIS C, ZHENG A X, et al. Collaborative denoising auto-encoders for top-n recommender systems [C]//Proceedings of the ninth ACM international conference on web search and data mining. California, USA; ACM, 2016:153–162.
- [23] 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. Boston, USA; ACM, 2016:233–240.
- [24] 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. Cambridge, UK; ACM, 2017:425–434.
- [25] CHEN C, ZHANG M, LIU Y, et al. Neural attentional rating regression with review-level explanations [C]//Proceedings of the 2018 conference on world wide web. Lyon, France; [s. n.], 2018:1583–1592.