

文章编号: 1006-2475(2020)10-0007-05

# 一种基于文本卷积神经网络的推荐算法

杨 辉,王月海,豆震泽

(北方工业大学信息学院,北京 100144)

**摘要:** 传统的矩阵因子分解模型不能有效提取用户和物品特征,而基于深度学习模型可以很好地提取特征信息。当前,主流的基于深度学习推荐算法只是单一地将神经网络的输出或物品特征与用户特征乘积的形式来做推荐预测,不能充分挖掘用户和物品之间的关系。基于此,本文提出一种基于文本卷积神经网络与带偏置项的奇异值分解(BiasSVD)结合的推荐算法,利用文本卷积神经网络(TextCNN)来充分提取用户和物品的特征信息,然后用奇异值分解方法来做推荐,深层次理解文档上下文信息,进一步提高推荐的准确性。将该算法在 MovieLens 的 2 个真实数据集上做广泛的评估分析,推荐的准确度要明显优于 ConvMF 算法及主流深度学习推荐算法。

**关键词:** 矩阵分解; 奇异值分解; 深度学习; 文本卷积神经网络

**中图分类号:** TP301.6      **文献标志码:** A      **DOI:** 10.3969/j.issn.1006-2475.2020.10.002

## A Recommendation Algorithm Based on Text Convolutional Neural Network

YANG Hui, WANG Yue-hai, DOU Zhen-ze

(School of Information, North China University of Technology, Beijing 100144, China)

**Abstract:** The traditional matrix factorization model can not effectively extract the features of users and items, while the deep learning model can extract the feature information well. At present, the mainstream recommendation algorithm based on deep learning only uses the output of neural network or the product of item features and user features to make recommendation prediction, which can not fully mine the relationship between users and items. Based on this, this paper proposes a recommendation algorithm based on the combination of text convolutional neural network and bias singular value decomposition (BiasSVD). Text convolutional neural network (TextCNN) is used to fully extract the feature information of users and items, and then singular value decomposition method is used to make recommendations, which can deeply understand the document context information and further improve the accuracy of recommendation. After extensive evaluation and analysis on two real datasets of MovieLens, the recommendation accuracy of this algorithm is obviously better than that of ConvMF algorithm and mainstream deep learning recommendation algorithm.

**Key words:** matrix decomposition; singular value decomposition; deep learning; text convolutional neural network

## 0 引 言

随着云时代的到来,电子商务、在线新闻和社交媒体数据呈现爆炸式增长。据 IDC 预测,到 2020 年全球将拥有 35 ZB 的数据量<sup>[1]</sup>。这些海量的数据给人类社会带来了变革性的发展,但同时带来了“信息过载”和“长尾效应”问题,如何从这些纷繁复杂的数据中获取有价值的信息,成为当今大数据处理的难题。

推荐系统作为解决这些问题的方法之一发挥着关键的作用,但同时也面临一些问题,比如用户对物品评级稀疏性的不断增加,以及如何去处理这些海量数据,这些都是影响推荐系统质量的主要因素。

协同过滤是推荐算法中最成功的算法之一,特别是基于模型的协同过滤算法因为其可扩展性强、精确、可以比较好地处理稀疏问题而得到大量的研究<sup>[2]</sup>。

收稿日期: 2020-03-18; 修回日期: 2020-03-24

基金项目: 国家自然科学基金资助项目(61573019)

作者简介: 杨辉(1993—),男,安徽合肥人,硕士研究生,研究方向: 人工智能、大数据挖掘、推荐系统, E-mail: 215835380@qq.com; 王月海(1975—),男,山东莒南人,教授,CCF 容错专委会委员,博士,研究方向: 人工智能、大数据挖掘、智能机器人, E-mail: wangyuehai@ncut.edu.cn; 豆震泽(1995—),男,北京人,硕士研究生,研究方向: 人工智能, E-mail: 379113096@qq.com。

传统的矩阵模型是协同过滤算法中的一种比较常见的建模方法。它是将原来的评分矩阵近似地分解成2个小矩阵的乘积,将用户和物品的特征分别映射到各自的潜在空间中。在实际推荐中不再使用原来的大矩阵分析,而是用这2个带有用户特征和物品特征的小矩阵进行推荐。因为矩阵分解模型的简单和可拓展性,其得到了广泛的研究。随之涌现出了大量相关的算法,如比较经典的 LFM 算法,因为它只要变成2个矩阵,从而避开了需要填充稀疏性矩阵的问题。虽然此方法比较简洁,但效果一般,原因是考虑的信息太少,无法学习到用户与项目间的深层次特征。之后 Koren 等人考虑到用户评分偏好问题,比如一些用户习惯性打低分,这些错误的偏好会对推荐的准确度有很大影响。基于此, Koren 等人<sup>[3,4]</sup>在原来的矩阵上增加了用户和项目的偏置项,即 BiasSVD。但针对 LFM 算法出现的问题仍没有得到有效的解决,数据的高稀疏性和海量数据处理仍是推荐系统需迫切解决的难题。

近年来,神经网络被大量运用于自然语言处理<sup>[5-6]</sup>、图像处理<sup>[7]</sup>、语音识别<sup>[8]</sup>中,并取得了卓著的效果。深度学习不仅具有自学习的能力,而且具有强大的特征提取和组合能力,因而可以解决传统矩阵分解的一些问题,比如稀疏性、维度风险、计算量、不能有效提取特征信息、可解释性等。基于此,深度学习被广泛运用到推荐系统中,成为近些年研究的热点。

为了解决数据高稀疏问题,一些推荐算法考虑加入辅助信息来提高评级精确度。例如文献[9]提出一种神经协同过滤模型(NCF),加入用户和物品ID信息。文献[10]提出一种深度协同推荐,把项目内容和用户行为信息结合起来,通过使用深度学习技术深层次表达用户特征,以此解决文本稀疏的问题。文献[11]提出卷积矩阵分解推荐模型,利用卷积网络来提取文本中的信息。文献[12]将循环神经网络(RNN)应用到推荐系统中,文献[13]用长短期记忆网络来记忆以前特征,达到精准推荐。

卷积神经网络广泛应用于图像分类、目标识别、语音识别等领域<sup>[14]</sup>,并取得了很好的效果。同样地,在推荐领域,也需要一种复杂度不高、能够高效地处理海量数据的网络,所以本文采用文本卷积神经网络来改进 Netflix 推荐大赛冠军 Koren 等人<sup>[3,4]</sup>的 BiasSVD 模型。本文将针对此算法做出改进,结合深度学习具有强大的从样本中学习数据集本质特征的能力,以及能够从多源异构数据中进行自动学习的能力,采用词嵌入结合前馈神经网络的方法和文本卷积神经网络组合来提取用户和物品特征,最后采用神经

网络并联方式输出预测评分项。本文算法经实验表明在面临极其稀疏的数据集,表现也比其它协同过滤算法要好,而且由于利用了用户对物品的交互信息,从而有了可解释性。本文算法在 MovieLens-1M 和 MovieLens-10M 数据集上进行的广泛对比评估表明,在 RMSE 和 MAE 指标上有了很大的提升。

本文工作主要如下:

1) 提出词嵌入结合前馈神经网络方式和文本卷积神经网络组合来提取用户和物品的描述信息和交互信息,可提高提取特征的能力。

2) 对 BiasSVD 算法进行改进,成功地将该算法对用户评分偏好的处理方式和本文算法结合起来,可提高评级预测的精度。

3) 在 MovieLens-1M 和 MovieLens-10M 这2个真实的公开数据集上做算法对比实验,验证算法的可靠性。

## 1 方法描述

### 1.1 BiasSVD 算法

矩阵分解是通过物品评分矩阵来推断表征用户和物品特征的,因为用户和物品特征之间具有高度一致性,所以能够进行推荐。这些矩阵分解的方法能够迅速流行起来源于它们有良好的可扩展性和比较精确的特性,而且也很多建模的实际情况提供了很大的灵活性。矩阵分解模型将用户和项目特征映射到维度为  $\omega$  的联合潜在因子模型,设对项目  $i$  相关联的向量  $q_i \in \mathbf{R}^\omega$ ,每个用户  $u$  相关联的向量  $p_u \in \mathbf{R}^\omega$ ,  $q_i$  表示项目拥有的特性,消极的或积极的,而  $p_u$  表示用户对这些项目的感兴趣程度,  $r_{ui}$  表示用户对项目的评分,则其预测值为:

$$\hat{r}_{ui} = q_i^T p_u \quad (1)$$

这个模型源于 SVD,当在处理稀疏的评分矩阵时,传统矩阵分解带来不确定性,即会出现很大偏差,所以在计算时往往需要对评分矩阵进行填充。然而遇到大量数据时这种方法既费时又不容易实现,而且处理较少项目时还容易过拟合,所以出现了改进版本:

$$\min_{p, q} \sum_{(u, i) \in I} (r_{ui} - q_i^T p_u)^2 + \lambda (\|q_i\|^2 + \|p_u\|^2) \quad (2)$$

其中  $I$  表示数据集,  $\lambda$  为控制正则化程度的参数,采用随机梯度下降优化参数:

误差项:

$$e_{ui} = r_{ui} - q_i^T p_u \quad (3)$$

按梯度下降对  $q_i$  和  $p_u$  进行优化:

$$q_i \leftarrow q_i + \gamma (e_{ui} p_u - \lambda q_i) \quad (4)$$

$$p_u \leftarrow p_u + \gamma (e_{ui} q_i - \lambda p_u) \quad (5)$$

其中  $\gamma$  为学习速率。因为存在用户偏好问题,这些因素的存在都严重影响了最后结果的精确性。例如,

假设所有电影的平均评分为 3.2 分,即式(6)中的  $\mu$ ,小明是个挑剔的人,他倾向于对电影的评分要比电影平均评分低 0.5 分,即式(6)中的  $b_u$ ,而《泰坦尼克号》是一部很好的电影,它的评分要比平均评分高 1.3 分,那么  $b_{ui} = 3.2 - 0.5 + 1.3 = 4$ ,因此估计小明对《泰坦尼克号》的评分为 4 分。所以就有了显式隔离项:

$$b_{ui} = \mu + b_i + b_u \quad (6)$$

其中  $\mu$  表示全局偏移量或平均评分,  $b_i$  表示与项目有关的偏移项,  $b_u$  表示与用户有关的偏移项。所以式(2)可以改写为:

$$\min_{p, q, b} \sum_{(u, i) \in I} (r_{ui} - \mu - b_i - b_u - q_i^T p_u)^2 + \lambda (\|q_i\|^2 + \|p_u\|^2 + \|b_u\|^2 + \|b_i\|^2) \quad (7)$$

由于偏差往往能捕捉到大部分观测到的信号,因此它们的精确建模至关重要。因此,其他研究提供了更精细的偏见模型<sup>[3-4]</sup>。

## 1.2 深度学习与 BiasSVD 算法的结合

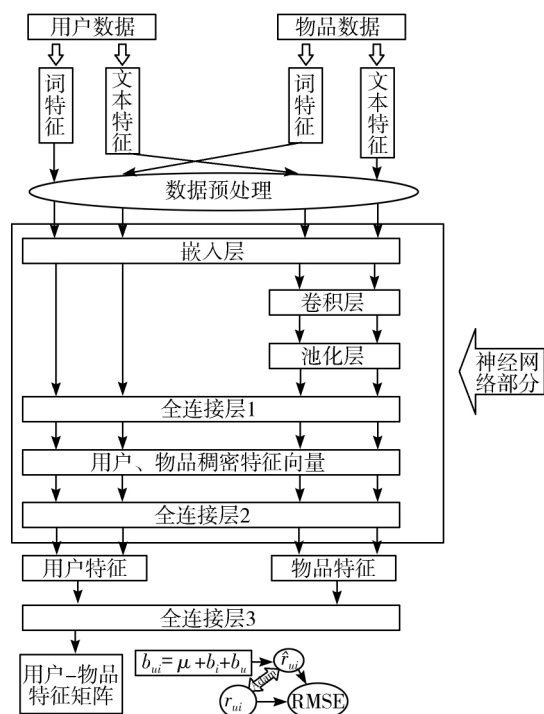


图1 本文算法模型流程图

本文主要针对 BiasSVD 算法中式(1)做出改进,不再用式(1)来表示预测评分项,而是采用深度学习神经网络输出预测评分项。如图1所示,深度学习神经网络部分有5层,分为嵌入层、卷积层、池化层、全连接层第1层和第2层。在嵌入层中输入经编码后的用户和物品信息数据,例如用户年龄、用户职业、用户ID、用户性别、电影名称、电影年代、电影分类、用户评价等信息,这里没有采用 one-hot 编码,因为它会导致词向量维度很大,而且没有上下词语义联系。本

文采用 Keras 神经网络的嵌入层,它是把一个密集向量表示的单词映射到一个连续的向量空间中。Keras 的嵌入层基于标签的信息来更新权重,从而达到较高的监督学习的目的,其本身也会学习到词与词之间的关系。Keras 在效率上虽然不如 Word2vec<sup>[5]</sup>,但它可以作为深度学习的一部分,利于本文算法来实现。卷积层、池化层与全连接第1层,主要用来处理用户和物品的一些文本信息,例如电影名称、用户评价等这些文本信息富有情感且对用户的选择起到很大的影响,需要用文本卷积神经网络<sup>[15]</sup>进行深层次提取,以便起到更好的文本分类作用。特别地在第1层网络中的激活函数采用 ReLU,ReLU 函数有克服梯度消失和加快训练速度的优点。用  $U$  表示经过全连接第2层输出的用户信息,用  $I$  表示经过全连接第2层输出的物品信息,则全连接第3层输出的用户和物品矩阵为  $\hat{r}_{ui}$ ,数学表达式为:

$$\hat{r}_{ui} = \text{conlayer}(U, I) \quad (8)$$

最后经过 Adam 算法优化:

$$\text{Adam}(\min_b \sum_{(u, i) \in I} (r_{ui} - \mu - b_i - b_u - \hat{r}_{ui})^2 + \lambda (\|b_u\|^2 + \|b_i\|^2)) \quad (9)$$

得到电影特征矩阵和用户特征矩阵后用 Top-N<sup>[16]</sup>推荐。

## 2 实验分析

### 2.1 数据分析

本文实验采用 MovieLens 数据集,MovieLens 数据集包含多个用户对多部电影的评级数据,也包括电影元数据信息和用户属性信息。这个数据集经常用来做推荐系统机器学习算法的测试数据集,尤其在推荐系统领域,很多推荐算法文献都是基于这个数据集的<sup>[2-4, 17-20]</sup>。

表1与表2分别给出了 MovieLens 数据集 MovieLens-1M 和 MovieLens-10M 的数据分布。

表1 MovieLens-1M 数据集情况

电影数	用户数	评分段	评分数	稀疏度/%
3706	6040	1~5	993482	4.64

表2 MovieLens-10M 数据集情况

电影数	用户数	评分段	tags	评分数	稀疏度/%
10681	71567	1~5	95580	9945875	1.41

如图2所示,在对 MovieLens 整个数据集进行统计分析时,发现用户评分随着年龄的增长呈现递增趋势,这表明在此数据集上年龄越大对电影的评分标准越低。为了使算法对用户的情感分析更加精准,所以在提取特征时加上了年龄因素这个隐性特征。

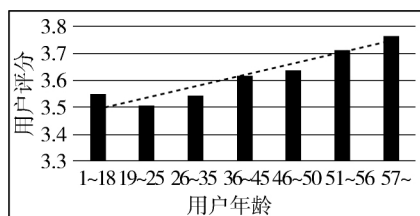


图2 用户年龄—评分情况

如图3所示,在对电影年代和用户评分之间进行统计分析时,发现用户评分随着电影年代的递增呈现递减趋势,这表明在 MovieLens 整个数据集上大量用户可能更青睐于年代久一点的电影。这些都是影响用户评分的因素,所以在对算法进行评估时也加上了这一隐性特征。

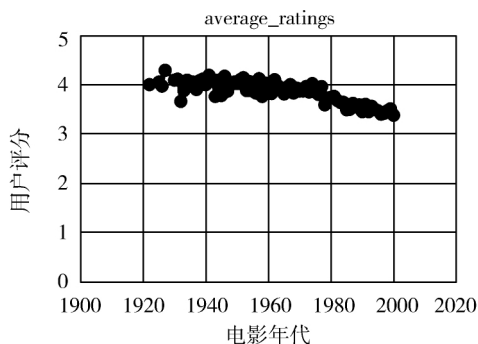


图3 电影年代—评分情况

## 2.2 评价指标

本节给出评价指标:均方根误差(RMSE)和平均绝对误差(MAE)经常被各种推荐系统用来评价系统优良的指标<sup>[2-3,11,17-19,21-23]</sup>,百度、Netflix、阿里巴巴等公司举办的推荐系统大赛也用此指标。

RMSE 测量的是模型预测值与观测值之间的差异,其中  $m$  为测试集规模,  $y_i$  为真实评分,  $\hat{y}_i$  为预测评分。

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (10)$$

MAE 测量的是预测值和实际值之间的差值的绝对值。

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (11)$$

## 2.3 实验结果与分析

将本文算法与7种常用的算法进行比较,这7种算法分别是:

1) SVD (Singular Value Decomposition)。Sarwar 等人<sup>[20]</sup>在2002年发现预测评分矩阵  $\hat{R} = U^T V$  与目标矩阵  $R$  的距离的平方最小化。

2) SVD++。Koren<sup>[3]</sup>在2010年将显式反馈和隐式反馈集成到一个模型中,在当时实现了最先进的性能。

3) CDL (Collaborative Deep Learning)。Wang 等人<sup>[10]</sup>在2015年发现利用辅助信息可解决用户对评分矩阵的稀疏问题,而协作主题回归表现得更好。

4) MLP (Multi-layered Perceptron)。He 等人<sup>[9]</sup>在2017年将基于多层感知机网络加上 ReLU 激活函数,用于推荐。

5) ConvMF (Convolutional MF)。Kim 等人<sup>[11]</sup>在2016年将卷积神经网络(CNN)集成到概率矩阵分解的算法。

6) NeuMF (Neural Matrix Factorization)。He 等人<sup>[9]</sup>在2017年结合多层感知机 MLP 和广义矩阵分解 GMF 得到的算法。

7) LSTM 模型。Yu 等人<sup>[2]</sup>在2019年利用 LSTM (长短期记忆神经网络)增强系统的可靠性。

由表3可以得到,在 MovieLens-1M 数据集上本文算法的 RMSE 较最先进的 LSTM 算法提升约 19.75% 以上,MAE 约提升 17.82% 以上。表中提升比例的计算方法为: (LSTM 算法值 - 本文算法值) / LSTM 算法值。

表3 MovieLens-1M 数据集上各算法的 RMSE 与 MAE

算法	RMSE	MAE
SVD	0.8730	0.686
SVD++	0.8620	0.673
CDL	0.8879	0.691
MLP	0.8773	0.687
ConvMF	0.8549	0.676
NeuMF	0.8631	0.674
LSTM	0.8480	0.668
本文算法	0.6805	0.5489
本文算法性能最小提升比例/%	19.75	17.82

在表4中本文对用户与物品的交互信息 tags 做了处理,结果本文算法较 LSTM 算法的 RMSE 约提升 35.6% 以上,MAE 约提升 47.4% 以上,可以得出用户与物品的交互信息在给用户推荐时起到非常重要的作用。

表4 MovieLens-10M 数据集上各算法的 RMSE 和 MAE

算法	RMSE	MAE
SVD	0.8135	0.631
SVD++	0.8108	0.629
CDL	0.8186	0.637
MLP	0.8143	0.632
ConvMF	0.7930	0.606
NeuMF	0.8107	0.674
LSTM	0.7882	0.601
本文算法	0.5073	0.316
本文算法性能最小提升比例/%	35.6	47.4

由表3与表4可以看出 ConvMF 算法表现较好,且与本文算法较为类似,也比较易于实现,故在图4中将本文的改进算法与 ConvMF 算法在 MovieLens-1M 数据集上进行了对比实验。其中 Epochs 在深度

学习中表示全数据集迭代次数。

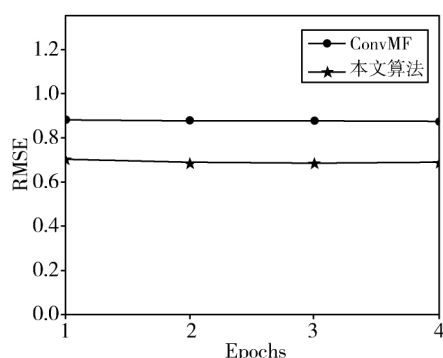


图4 本文算法与 ConvMF 算法在 MovieLens-1M 数据集上不同迭代次数时对评分精度 RMSE 的比较

由图4可以明显地看出,在预测观众喜好的准确性上,本文对 BiasSVD 算法改进的算法要明显优于 ConvMF 算法。

为了使实验结果更具说服力,本文分别对神经网络的迭代次数与学习速率进行了实验,如图5与图6所示。

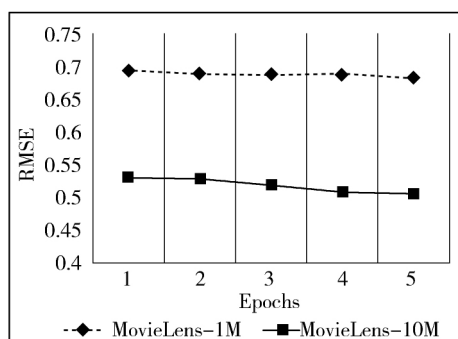


图5 本文算法在不同数据集上不同迭代次数时对评分精度 RMSE 的比较

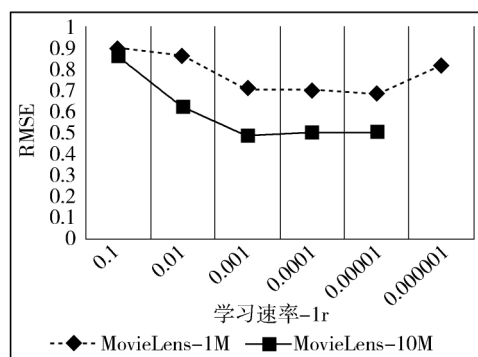


图6 本文算法在不同数据集上不同学习速率时对评分精度 RMSE 的比较

在图5与图6中可以看到学习速率对算法的影响明显,MovieLens-1M 和 MovieLens-10M 数据集分别在学习速率为 0.00001 和 0.001 时达到最佳。

通过大量实验验证,可以看到本文的算法在 MovieLens-1M 数据集上相较效果最好的 LSTM 算法的 RMSE 提升约 19.75% 以上,MAE 提升约 17.82%

以上,在 MovieLens-10M 数据集上 RMSE 提升约 35.6% 以上,MAE 提升约 47.4% 以上。综上,可以得到,本文的算法在对用户情感、喜好的理解程度比效果最好的 ConvMF 算法与 LSTM 算法要深,更能满足用户的需求。

### 3 结束语

传统的矩阵分解算法因为只是简单地将物品与用户矩阵作内积,不能有效地提取用户与物品特征,卷积矩阵分解 ConvMF 与长短期记忆神经网络 LSTM 先后对其改进,但对用户与物品文档上下文理解不够,导致评级预测精度较差,本文提出用词嵌入结合前馈神经网络和文本卷积神经网络组合来提取特征的方法对 BiasSVD 算法进行改进。在 MovieLens 的 2 个真实数据集上进行广泛的实验对比,发现本文算法评级预测更加准确,验证了用户对物品的交互信息更加能反映用户的喜好,所以基于深度学习的用户和物品信息的提取仍然值得深入研究。

参考文献:

- [1] 陶雪娇,胡晓峰,刘洋. 大数据研究综述[J]. 系统仿真学报, 2013, 25(S1): 142-146.
- [2] YU S, YANG M, QU Q, et al. Contextual-boosted deep neural collaborative filtering model for interpretable recommendation[J]. Expert Systems with Applications, 2019, 136: 365-375.
- [3] KOREN Y. Collaborative filtering with temporal dynamics[J]. Communications of the ACM, 2010, 53(4): 89-97.
- [4] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [5] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[C]// Proceedings of the 1st International Conference on Learning Representations. 2013.
- [6] PENNINGTON J, SOCHER R, MANNING C D. GloVe: Global vectors for word representation[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014: 1532-1543.
- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [8] SAINATH T N, KINGSBURY B, MOHAMED A-R, et al. Improvements to deep convolutional neural networks for LVCSR[C]// Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding. 2013: 315-320.
- [9] HE X N, LIAO L Z, ZHANG H W, et al. Neural collaborative filtering[C]// Proceedings of the 26th International Conference on World Wide Web. 2017: 173-182.

(下转第 16 页)

- [9] 苗夺谦,王珏. 基于粗糙集的多变量决策树构造方法[J]. 软件学报,1997 8(6):425-431.
- [10] 吴海洋,缪巍巍,郭波,等. 基于遗传算法的 BP 神经网络蓄电池寿命预测研究[J]. 计算机与数字工程,2019 47(5):1275-1278.
- [11] KRIZHEVSKY A, HINTON G. Learning Multiple Layers of Features from Tiny Images[R]. University of Toronto, 2009.
- [12] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019: 4171-4186.
- [13] SUN Y Q, WU Z X, WANG X, et al. Exploiting objects with LSTMs for video categorization[C]// Proceedings of the 24th ACM International Conference on Multimedia. 2016: 142-146.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 6000-6010.
- [15] CHEN T L, DING S J, XIE J Y, et al. ABD-Net: Attentive but diverse person re-identification[J]. arXiv preprint arXiv:1908.01114, 2019.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997 9(8):1735-1780.
- [17] LYU Q, ZHU J. Revisit long short-term memory: An optimization perspective[C]// Proceedings of the 2014 Workshop on Deep Learning and Representation Learning. 2014.
- [18] LI Z Y, GAVVES E, JAIN M, et al. VideoLSTM convolves, attends and flows for action recognition[J]. arXiv preprint arXiv:1607.01794, 2016.
- [19] WUH Y, LU X, MIAO W W, et al. Dynamic routing programming for power communication networks by recurrent neural networks based reliability prediction and particle swarm optimization[C]// Proceedings of the 5th IEEE International Conference on Advances in Electrical and Electronics, Information, Communication and Bio-Informatics (AEEICB2019). 2019: 1217-1221.
- [20] 章鑫锋,张彩友,冯毅萍,等. 基于多粒度 LSTM 模型的换流站设备分析研究[C]// 第 30 届中国过程控制会议 (CPCC 2019) 论文集. 2019.
- [21] DONAHUE J, HENDRICKS L A, GUADARRAMA S, et al. Long-term recurrent convolutional networks for visual recognition and description[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. 2015: 2625-2634.
- [22] KINGMA D P, WELING M. Auto-encoding variational Bayes[J]. arXiv preprint arXiv:1312.6114, 2013.
- [23] ZHANG Z L, SABUNCU M R. Generalized cross entropy loss for training deep neural networks with noisy labels[C]// Proceedings of the 32nd Conference on Neural Information Processing Systems. 2018: 8778-8788.
- [24] FARAHNAK-GHAZANI F, BAGHSHAH M S. Multi-label classification with feature-aware implicit encoding and generalized cross-entropy loss[C]// Proceedings of the 24th Iranian Conference on Electrical Engineering (ICEE). 2016: 1574-1579.

(上接第 11 页)

- [10] WANG H, WANG N Y, YEUNG D-Y. Collaborative deep learning for recommender systems[C]// Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015: 1235-1244.
- [11] KIM D H, PARK C, OH J, et al. Convolutional matrix factorization for document context-aware recommendation[C]// Proceedings of the 10th ACM Conference on Recommender Systems. 2016: 233-240.
- [12] WU C-Y, AHMED A, BEUTEL A, et al. Joint training of ratings and reviews with recurrent recommender networks[C]// Proceedings of the 5th International Conference on Learning Representations. 2017.
- [13] 曾安,聂文俊. 基于深度双向 LSTM 的股票推荐系统[J]. 计算机科学,2019 46(10):84-89.
- [14] 李南星,盛益强,倪宏. 用于个性化推荐的条件卷积隐因子模型[J/OL]. 计算机工程,(2019-05-28) [2020-02-16]. <https://doi.org/10.19678/j.issn.1000-3428.0054209>.
- [15] KIM Y. Convolutional neural networks for sentence classification[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014: 1746-1751.
- [16] WU Y, DUBOIS C, ZHENG A X, et al. Collaborative denoising auto-encoders for top-n recommender systems[C]// Proceedings of the 9th ACM International Conference on Web Search and Data Mining. 2016: 153-162.
- [17] STRUB F, MARY J. Collaborative filtering with stacked denoising autoencoders and sparse inputs[C]// NIPS Workshop on Machine Learning for eCommerce. 2015.
- [18] SEDHAIN S, MENON A K, SANNER S, et al. AutoRec: Autoencoders meet collaborative filtering[C]// Proceedings of the 24th International Conference on World Wide Web. 2015: 111-112.
- [19] SHIBATA H, TAKAMA Y. Behavior analysis of RBM for estimating latent factor vectors from rating matrix[C]// Proceedings of the 2017 6th International Conference on Informatics, Electronics and Vision, 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMT). 2017.
- [20] SARWAR B, KARYPIS G, KONSTAN J, et al. Incremental singular value decomposition algorithms for highly scalable recommender systems[C]// Proceedings of the 2002 International Conference on Computer and Information Science. 2002: 27-28.
- [21] ZHAO L L, LU Z Q, PAN S J L, et al. Matrix factorization + for movie recommendation[C]// Proceedings of the 25th International Joint Conference on Artificial Intelligence. 2016: 3945-3951.
- [22] HE X N, CHUA T-S. Neural factorization machines for sparse predictive analytics[C]// Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017: 355-364.
- [23] SALAKHUTDINOV R, MNIH A, HINTON G. Restricted Boltzmann machines for collaborative filtering[C]// Proceedings of the 24th International Conference on Machine Learning. 2007: 791-798.