



DOI:10.13364/j.issn.1672-6510.20210184

基于 BERT-GCN-ResNet 的文献摘要自动分类方法

郭羽, 林丽媛, 刘玉良

(天津科技大学电子信息与自动化学院, 天津 300222)

摘要: 为了提升文献摘要自动分类性能,提高分类准确率,有效学习词与文档、文档与文档之间的关联性,本文提出一种基于 BERT-GCN-ResNet 的文献摘要分类方法.该方法采用转换器的双向编码器表示模型(bidirectional encoder representation from transformers, BERT)得到待分类短文本的词向量初始特征,进而构建边和节点特征;将其输入图卷积神经网络(graph convolutional networks, GCN),并在图卷积层之间加入残差网络(residual network, ResNet)模块;最后将利用图卷积层和残差网络层得到的短文本表示输出至 softmax,得到最终的分类结果.将该方法在 4 种不同的公开文本分类数据集上进行实验,模型准确率最高达 97.01%,优于基准模型.同时,在本文构建的基于短文本文献摘要的 Abstext 数据集上进行验证,分类准确率为 96.85%,表明该模型泛化能力较好,能够提高文献自动分类的准确率.

关键词: 图卷积神经网络; 转换器的双向编码器表示模型; 残差网络; 文献自动分类; 预训练模型

中图分类号: TP391

文献标志码: A

文章编号: 1672-6510(2022)02-0051-06

Automatic Abstract Classification Method Based on BERT-GCN-ResNet

GUO Yu, LIN Liyuan, LIU Yuliang

(College of Electronic Information and Automation, Tianjin University of Science & Technology,
Tianjin 300222, China)

Abstract: In order to improve the automatic classification performance of literature abstract, ensure the classification accuracy and effectively investigate the association between words and documents, documents and documents, in this article we propose a literature abstract classification method based on BERT-GCN-ResNet. In this method, using bidirectional encoder representation from transformers (BERT), we first obtain initial word vector features of the short text to be classified, thus constructing the edge and node features. Then, we further input these features into the graph convolutional networks (GCN), and add the residual network (Resnet) module between the graph convolution layers. Finally, the short text representation obtained by graph convolution layer and residual network (ResNet) layer is output to softmax to help get the final classification result. The accuracy of our proposed method was verified on four different public text classification datasets, and the model accuracy reached 97.01%, which is better than the baseline model. Meanwhile, the accuracy of the proposed method was verified on the Abstext dataset based on short text literature abstracts constructed in this article, and the classification accuracy was 96.85%, indicating that the model has good generalization ability and can improve the accuracy of automatic literature classification.

Key words: graph convolutional networks (GCN); bidirectional encoder representation from transformers (BERT); residual network (ResNet); automatic literature classification; pre-training model

文献分类是图书情报学科的重要研究方向. 当前的文献分类主要以《中国图书馆分类法》为依据^[1], 使用题名、摘要、关键词、刊名、作者和机构等

信息进行分类^[2]. 从科技文献检索的角度来说, 摘要中存在文献内容的显性特征, 包含着文献的观点和价值, 因此, 相比使用题目、关键词、刊名等文献信息,

收稿日期: 2021-09-21; 修回日期: 2021-12-21

基金项目: 天津市教委科研计划项目(2019KJ211)

作者简介: 郭羽(1996—), 女, 河北秦皇岛人, 硕士研究生; 通信作者: 林丽媛, 讲师, linly@tust.edu.cn

使用摘要信息进行文献分类能提高文献分类的准确率,对图书情报学科的发展具有重要意义。

文献摘要分类方法分为传统的人工方法以及机器学习和深度学习^[3-4]的方法。机器学习模型的不足之处:一是贝叶斯分类器^[5]在属性个数比较多或者属性关联性较大时分类效果不好;二是支持向量机算法(support vector machine, SVM)^[6]对大规模训练样本难以实施,不适于解决多分类的问题。除此之外,决策树系列模型和随机森林^[7-8]等方法需要足够的词语共现信息完成模型训练,而这些信息恰好是文献摘要文本所缺少的。对原始文本数据进行特征处理后,为了让计算机更好地理解自然语言,需要将处理后的数据转换为文本表示。常用的文本表征方式包括词袋模型^[9]、N元模型(N-gram)、词频逆文档频率(term frequency inverse document frequency, TF-IDF)^[10]、Fasttext等^[11]。文本表示之后,利用网络模型将获取的有效信息拟合并计算模型参数。在过去的几十年中,短文本分类技术迅猛发展,分类性能明显提升。论文摘要是一类极具研究价值的短文本数据,所以利用摘要对论文进行分类与推荐具有重要意义。但是,目前可检索到的摘要自动分类的研究却很少。

随着深度学习技术^[12-15]的广泛应用,为了解决当前文献分类领域存在的问题,将短文本分类方法迁移到摘要自动分类中,希望能够提升摘要自动分类性能。以文献数据库中已广泛存在的大规模摘要数据集为训练语料,实现文献自动分类效果的提升。其中,基于预训练语言模型的方法在这些文本分类任务中取得较大突破,比如 Word2vec^[16]和 One-hot^[17]编码等。但是,这些预训练模型对短文本中词与词的关联性把握不准,不能获得关联紧密的边和节点的特征信息。转换器的双向编码器表示模型(bidirectional encoder representation from transformers, BERT)^[18]用预训练中的双向表示和自注意力机制模块进行训练,得到上下文语义特征,在短文本分类任务中获得了显著提升。图卷积神经网络(graph convolutional networks, GCN)可实现节点特征和边特征的端到端学习。但是,用于短文本分类任务的多层 GCN 会导致节点特征过度平滑,使局部特征收敛到一个近似值,导致分类性能下降。Yao等^[19]提出了文本图卷积网络模型(Text GCN),可根据词共现和文档词关系建立一个文本图的语料库,可学习单词嵌入和文档嵌入。但 Text GCN 不能快速生成嵌入向量,对无标签文档的预测效果不好。

为了改善模型的分类型效果,本文提出一种基于 BERT-GCN-ResNet 的文献摘要自动分类方法,该方

法利用 BERT 得到待分类短文本的词向量初始特征,进而构建边和节点的特征并将其输入 GCN,然后在图卷积层之间加入 ResNet 结构,最后将利用图卷积层和 ResNet 层得到的短文本表示输出至 softmax^[20],得到最终的分类结果。同时,为了防止过拟合,在图卷积层引入了 Dropout 技术,旨在有效提高短文本分类的准确率,得到较好的分类效果。

1 BERT-GCN-ResNet 模型设计

1.1 模型算法

在数据输入 BERT-GCN-ResNet 模型前,构建词节点和文档节点的异构图。根据词频逆文档频率(TF-IDF)和点互信息(PMI)分别计算文档节点、词节点之间的边的权重和两个词之间的权重:

$$A_{ij} = \begin{cases} \text{PMI}(i, j) & i, j \text{ 为词且 } i \neq j \\ \text{TF-IDF}(i, j) & i \text{ 为文档, } j \text{ 为词} \\ 1 & i = j \\ 0 & \text{其他} \end{cases} \quad (1)$$

式中对于 PMI 值定义为

$$\text{PMI}(i, j) = \log \frac{p(i, j)}{p(i)p(j)} \quad (2)$$

$$p(i, j) = \frac{\#W(i, j)}{\#W} \quad (3)$$

$$p(i) = \frac{\#W(i)}{\#W} \quad (4)$$

其中: $\#W(i)$ 为语料库中包含单词的滑动窗口的数目; $\#W(i, j)$ 为滑动窗口的数目,同时包含单词 i 和 j 的窗口; $\#W$ 为语料库中滑动窗口的总数。PMI 值为正数时,表示语料库中单词的语义相关性较高,反之则表示语料库中单词的语义相关性较低或者没有相关性。因此,只需要在 PMI 值为正数的词之间添加边。

在 BERT-GCN-ResNet 模型中,使用标识矩阵 H 作为初始节点特征,在 BERT 预训练模型中得到特征节点表示向量为

$$X = f(H) \quad (5)$$

GCN 是一个多层神经网络,适用于任意拓扑结构的节点与图,可同时学习节点特征与结构信息。这种特殊的拓扑结构通过一次卷积可使每一个节点都拥有其邻居节点的信息。因此,在节点分类与边预测任务中,GCN 效果远远优于其他方法。将其应用于文本处理,并根据节点的邻域属性进行节点嵌入向量。第 i 层 GCN 层输出特征矩阵计算式为

$$L^{(i)} = \rho(\tilde{A}L^{i-1}W_i) \quad (6)$$

其中: $\tilde{A} = D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ 为归一化邻接矩阵, 式中 D 为 BERT-GCN-ResNet 模型所生成的图结构 P 的度矩阵 (度矩阵为对角矩阵, 对角线上的元素为各个顶点的度, 顶点的度表示和该顶点相关联的边的数量); W_i 为权重矩阵, $W_i \in R^{b_{i-1} \times n_i}$, n 为特征矩阵嵌入维数, b 为特征向量的维度; ρ 为激活函数 Tanhshrink. 本文模型为两个 GCN 层, 计算式为

$$L^{(2)} = \rho(\tilde{A}L^{(1)}W_1) \quad (7)$$

其中: $L^1 = \tilde{A}XW_0 + X$, $L^0 = X$ 是该模型输入的特征矩阵, 在两个 GCN 层间形成 ResNet 的残差结构. GCN 的输出被视为文档的最终表示, 然后输入到 softmax 层进行分类, 为

$$Z = \text{softmax}(\tilde{A}\text{Tanhshrink}(\tilde{A}XW_0 + X)W_1) \quad (8)$$

损失函数为多分类合页损失函数

$$L = - \sum_{d \in y_D} \sum_{F=1}^F Y \ln G(X, A) \quad (9)$$

其中: G 为 BERT-GCN-ResNet 模型经过 softmax 后的最终输出表达式, y_D 为具有标签的文档索引集合, F 为输出特征的维度, Y 为标签矩阵. 权重参数 W_0 和 W_1 可以通过梯度下降训练.

1.2 网络结构

本文提出的基于 BERT-GCN-ResNet 文本多分类模型由 BERT、GCN 和 ResNet^[21] 组成. 首先, 在文本输入之前, 对需要的文本进行数据清洗, 去除无用的停词和错误的数据, 目的在于提高数据准确率和节约计算成本; 其次, 使用 BERT-base 对数据处理后的短文本中每个词进行初始特征表示. BERT 使用

MaskedLM 可实现深层双向联合训练, 使其更易于理解文中上下两个句子之间的联系. 因此, 将 BERT 生成的词向量添加到模型中, 提升模型分类性能. 在特征表示层中, 为有效提高分类性能, 构建了文本图卷积网络 GCN, 根据词共现和文档词建立单个文本图的语料库, 同时学习单词嵌入和文档嵌入. 为了充分学习上下文的更多语义信息, 提高模型分类性能, 引入 ResNet 模块. 在分类输出层中, 使用多分类合页损失函数 MultiMarginLoss 作为训练模型的损失函数, 通过 softmax 函数生成每个类别的概率, 并根据概率值进行短文本类别预测.

BERT-GCN-ResNet 整体网络结构如图 1 所示, 其中左半部分为子结构 BERT 模型网络结构. 首先将数据清洗后的短文本处理成词向量的形式 $[H_1, \dots, H_n]^T$, 然后经过 12 个 Multi-layer Transformer 模块得到新的词向量文本表示 $[X_1, \dots, X_n]^T$, 最后构建图数据结构 P . Multi-layer Transformer 由 Multi-head self-Attention 模块和 Layer Normalization 模块构成, 并且引入了残差模块 R . 将 BERT 处理得到的图结构数据输入到 GCN-ResNet Layers 模块, 如图 1 右半部分所示. 在图卷积层之间引入 ResNet 模块 $L^1 = \tilde{A}XW_0 + X$, 这种跳跃连接方式将有助于解决梯度消失和爆炸问题; 在训练网络时, 可保证良好的性能, 进而提高网络的提取特征能力和模型的性能. 通过 softmax 函数得到最终的输出结果 Z , 生成每个类别的预测值, 并根据预测值的大小进行短文本类别判定.

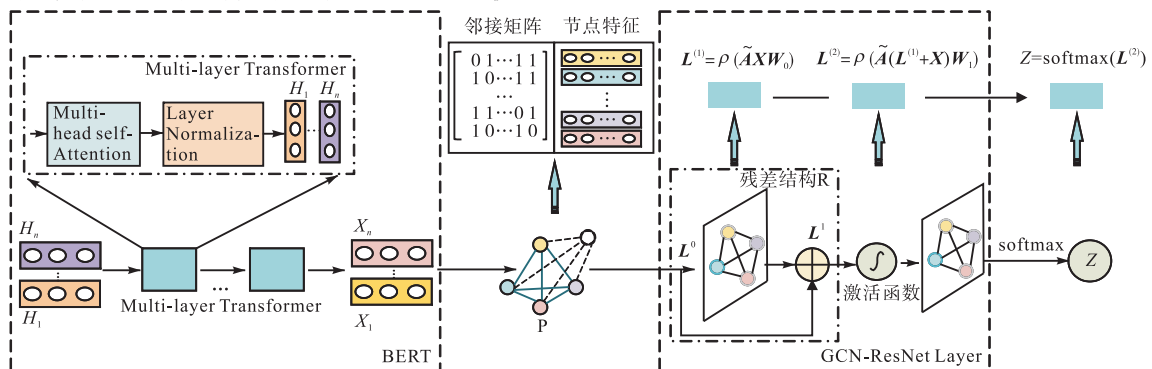


图 1 BERT-GCN-ResNet 网络结构

Fig. 1 BERT-GCN-ResNet network structure

2 实验设计

2.1 实验环境

实验在 Win 10 (64 位) 操作系统上进行, 以深度

学习框架 Pytorch 1.6.0 为基础, 编程语言采用 Python 3.7. 实验使用的 GPU 为 NVIDIA GeForce RTX 2080 Ti, 其显存为 11 GB, CPU 为 Intel Xeon CPU E5-2678 v3 六核, 其内存为 62 GB.

2.2 数据集

本文在 5 个基准短文本数据集 R8、R52、AGNews、TagMyNews、Abstext 上进行实验。

R8 是 Reuters21578 数据集的一个子集。R8^[19]有 8 个类别,被分为 6906 个训练文件和 768 个测试文件。

R52 是 Reuters21578 数据集的一个子集。R52^[19]有 52 个类别,训练集中文件 8190 个,测试集中文件 909 个。

AGNews^[22]是由 2000 多个不同的新闻来源搜集的超过 100 万篇新闻文章构成的。实验从中随机挑选 7600 条新闻并分为 4 类,其中训练集中文件 6840 个,测试集中文件 760 个。

TagMyNews 是 Vitale 于 2012 年^[23]发布的基准分类数据中的新闻标题数据集。该数据包含 7 类,实验从中随机挑选 9794 个数据并分为 7 类,其中训练集中文件 8814 个,测试集中文件 980 个。

Abstext 是使用网络爬虫技术在爱思唯尔数据库自动获取的英文文献摘要数据集^[24],经过数据处理后得到 3557 条数据,分为 3 个类别,其中训练集数据 3200 个,测试集数据 357 个。

对上述数据集进行预处理,标记文本数据,删除在 NLTK(natural language toolkit)库中定义的停词以及在 R8、R52、AGNews、TagMyNews、Abstext 中出现小于 5 次的低频单词。预处理后数据集的统计数据见表 1,其中平均长度指单词个数。

表 1 数据集的统计数据结果

Tab. 1 Statistical results of the data sets

数据集名称	训练集	测试集	类别	平均长度/个
R8	6 906	768	8	65.72
R52	8 190	909	52	69.82
AGNews	6 840	760	4	24.49
TagMyNews	8 814	980	7	21.65
Abstext	3 200	357	3	221.26

2.3 实验结果与分析

实验均采用 Adam^[25]优化器优化损失,epoch 为 200,学习率为 0.1,在 R8、R52、AGNews、TagMyNews 短文本数据集上进行短文本分类准确性检测,并按照训练集与测试集比例为 9:1 进行划分数据集。按照图 1 所设计的网络模型进行训练,将预处理后的短文本转换为 BERT 词向量并输入到 BERT-GCN-ResNet 网络。实验对比朴素贝叶斯(naive Bayesian, NB)、决策树(decision tree, DT)、随机森林(random forests, RF)、SVM、Fasttext、卷积

长短时记忆网络(convolutional long short-term memory, CLSTM)^[26]、卷积神经网络(convolutional neural network, CNN)、门控循环单元(gated recurrent unit, GRU)、长短时记忆网络(long short-term memory, LSTM)、双向长短时记忆网络(bidirectional long short-term memory, Bi-LSTM)和 GCN 短文本分类的平均准确率,实验结果见表 2。

表 2 短文本分类任务准确率

Tab. 2 Accuracy of short text classification task

模型	准确率/%			
	R8	R52	AGNews	TagMyNews
TF-IDF + NB ^[5]	91.73	79.09	85.52	85.47
TF-IDF + DT ^[7-8]	90.22	81.93	62.60	62.15
TF-IDF + RF ^[7-8]	90.46	82.96	80.00	80.47
TF-IDF + SVM ^[6]	94.63	89.82	85.65	85.51
Fasttext ^[11]	87.24	78.57	72.63	7.37
CNN ^[12]	92.57	84.91	64.34	58.43
GRU ^[13]	79.53	69.16	55.13	14.50
LSTM ^[14]	71.06	69.49	57.89	24.11
CLSTM ^[26]	51.18	25.20	25.00	25.33
Bi-LSTM ^[15]	71.71	67.40	55.53	31.26
GCN ^[19-20]	96.88	93.81	67.61	54.28
BERT-GCN-ResNet	97.01	94.07	86.71	85.54

由表 2 可知,BERT-GCN-ResNet 模型的分分类准确率在 R8、R52、AGNews、TagMyNews 数据集上分类性能表现良好,其中在 R8 数据集上准确率高达 97.01%,比 GCN 模型有所提高,说明本文改进的模型具有更加理想的短文本分类性能。本文模型在 R52、AGNews、TagMyNews 数据集分类效果也明显优于其他方法,尤其在 TagMyNews 数据集上。TF-IDF+NB、TF-IDF+DT 和 TF-IDF+SVM 等传统的分类模型的分分类效果不理想,在数据集 R8 和 R52 分类准确率较低,特别是在数据集 R52 上,本文模型分类准确率比其提高 15%。相比于 GCN 模型,本文模型分类性能最高提升 31%,即使是在 CNN、Fasttext、CLSTM 等其他基准模型分类效果表现不好的数据集 TagMyNews 上,也能够表现出良好的分类性能,准确率为 85.54%。实验结果表明,本文提出的 BERT-GCN-ResNet 网络模型在短文本数据集上的分类效果显著。

为了验证本文提出 BERT-GCN-ResNet 网络模型的有效性,采用上述评价指标对各模块进行消融实验,并计算在每个数据集上的准确率,实验结果见表 3。与 GCN 相比,本文模型加入 BERT 模块后,在数据集 AGNews 和 TagMyNews 上的准确率分别提高 18.96% 和 30.22%;在数据集 R8 上,准确率略微提

升;在数据集 R52 上,准确率略微下降. 这表明 BERT 可获取上下文相关的双向特征表示,在分类任务上优势明显. 引入 ResNet 模块,本文模型在数据集 AGNews 和 TagMynews 上,准确率分别提高 19.62% 和 29.71%;在数据集 R52 上,准确率略微提升;在数据集 R8 上,准确率略微下降. 实验结果证明,引入 ResNet 可以提高网络提取特征能力,提升分类效果. 由于个别短文本语义相近,在文本预训练时出现文本表示相同的情况,造成语义混淆,导致在单独引入 BERT 或 ResNet 模块时,在数据集 R52 和 R8 上的准确率会有下降;同时引入 BERT 和 ResNet 模块后,在所有实验数据集上准确率均有所提高. 由此可见,本文模型融合了 BERT、GCN 和 ResNet 的优势,增强了文本语义特征提取能力,又凭借 GCN 适用于任意拓扑结构的独特优势,在节点分类任务中效果明显优于其他模型,因此, BERT-GCN-ResNet 模型能够使短文本分类性能提升.

表 3 BERT-GCN-ResNet 模型消融实验

Tab. 3 BERT-GCN-ResNet model ablation experiment

模型	准确率/%			
	R8	R52	AGNews	TagMyNews
GCN ^[14-15]	96.88	93.81	67.61	54.28
BERT-GCN	96.93	93.61	86.57	84.50
GCN-ResNet	96.57	93.85	87.23	83.99
BERT-GCN-ResNet	97.01	94.07	86.71	85.54

将该模型在短文本摘要数据集 Abstext 上进行验证,结果见表 4.

表 4 Abstext 实验结果

Tab. 4 Abstext experimental results

模型	准确率/%	模型	准确率/%
TF-IDF + NB ^[5]	94.31	LSTM ^[14]	57.01
TF-IDF + DT ^[7-8]	92.46	CLSTM ^[26]	41.51
TF-IDF + RF ^[7-8]	95.44	Bi-LSTM ^[15]	73.45
TF-IDF + SVM ^[6]	95.21	GCN ^[19-20]	93.12
Fasttext ^[11]	84.95	BERT-GCN	93.70
CNN ^[12]	87.70	GCN-ResNet	92.73
GRU ^[13]	80.51	BERT-GCN-ResNet	96.85

同 TF-IDF + NB、TF-IDF + DT、TF-IDF + RF、TF-IDF + SVM、Fasttext、CNN、GRU、LSTM、CLSTM、Bi-LSTM、GCN 等基准模型对比,本文模型的文献摘要自动分类结果准确率为 96.85%,比 GCN 提高了 3.73%,比 TF-IDF + DT 网络的准确率提高了 4.39%,明显优于其他基准模型,表明该模型具有良好的泛化能力. 由于 BERT 和 ResNet 模块可以提取短文本更多特征信息,进而提高短文本分类的准确

率. 实验结果表明,引入 BERT 预训练模型到 GCN 和 ResNet 融合网络中对于短文本分类效果的提升具有一定优势,尤其是在语义较为稀疏的短文本中表现出更好效果. 例如在 AGNews 和 TagMyNews 数据集上, BERT-GCN-ResNet 网络模型比 GCN 的实验结果分别提高了 19.10% 和 31.26%. 这表明该模型可以充分地学习文本中包含的长距离依赖和上下文信息,能够获得更加丰富的语义表示,可以极大促进短文本文献摘要分类性能的提升.

3 结论

本文提出一种基于 BERT-GCN-ResNet 的文献摘要自动分类方法. 将文献摘要分类问题转化为短文本节点分类问题. BERT 可以完成深层双向联合训练的任务,获得上下文相关的双向特征表示,因此,该方法利用 BERT 模型获得预训练的词向量,进而构建边和节点特征信息,将其输入融合了 ResNet 模块的两层 GCN 网络中,进一步提高网络对文本语义特征的提取能力,实现浅层网络训练,即可达到深层网络训练的效果. 将 GCN 和 ResNet 层得到的短文本表示输出至 softmax 分类器,得到最终分类结果. 该方法在 4 种不同短文本数据集上进行准确性验证,准确率最高为 97.01%,优于基准模型. 在本文构建的文献摘要数据集 Abstext 上进行验证,结果表明该模型具有良好的泛化性能,提高了文献摘要自动分类的准确性.

在未来的研究工作中,本文将从如何构建更富含语义信息的图特征出发,进一步提升模型提取语义特征信息的能力. 同时,本文是将多标签数据处理成单标签进行多分类,具有一定的局限性,后续将探索在面对多标签的分类问题上,如何实现较好的分类效果.

参考文献:

- [1] 孔洁. 基于深度学习与《中国图书馆分类法》的文献自动分类系统研究[J]. 新世纪图书馆, 2021(5): 51-56.
- [2] 雷兵,刘小,钟镇. 基于题录信息的领域学术文献细粒度分类方法研究[J]. 图书情报工作, 2021, 65(14): 1-10.
- [3] 罗鹏程,王一博,王继民. 基于深度预训练语言模型的文献学科自动分类研究[J]. 情报学报, 2020, 39(10): 1046-1059.
- [4] 赵旻,张智雄,刘欢,等. 基于 BERT 模型的中文医学

- 文献分类研究[J]. 数据分析与知识发现, 2020, 4(8): 41-49.
- [5] 贺鸣, 孙建军, 成颖. 基于朴素贝叶斯的文本分类研究综述[J]. 情报科学, 2016, 34(7): 147-154.
- [6] 王昊, 叶鹏, 邓三鸿. 机器学习在中文期刊论文自动分类研究中的应用[J]. 现代图书情报技术, 2014(3): 80-87.
- [7] BOUAZIZ A, DARTIGUES-PALLEZ C, DA COSTA PEREIRA C, et al. Short text classification using semantic random forest[C]//International Conference on Data Warehousing and Knowledge Discovery. Berlin: Springer, 2014: 288-299.
- [8] YIN C, XIANG J, ZHANG H, et al. A new classification method for short text based on SLAS and CART[C]//IEEE. Computational Intelligence Theory, Systems and Applications (CCITSA). New York: IEEE, 2015: 133-135.
- [9] ZHANG S, JIN X, SHEN D, et al. Short text classification by detecting information path[EB/OL]. [2021-08-01]. <https://doi.org/10.1145/2505515.2505638>.
- [10] 金燕, 黄杰. 基于信息熵与词长信息改进的 TFIDF 算法[J]. 浙江工业大学学报, 2021, 49(2): 203-209.
- [11] 张焱博, 郭凯. 基于 Fasttext 和多融合特征的文本分类模型[J]. 计算机仿真, 2021, 38(7): 461-466.
- [12] 万齐斌, 董方敏, 孙水发. 基于 BiLSTM-Attention-CNN 混合神经网络的文本分类方法[J]. 计算机应用与软件, 2020, 37(9): 94-98.
- [13] 邵良杉, 周玉. 基于语义规则与 RNN 模型的在线评论情感分类研究[J]. 中文信息学报, 2019, 33(6): 124-131.
- [14] 袁景凌, 丁远远, 盛德明, 等. 基于视觉方面注意力的图像文本情感分析模型[J/OL]. 计算机科学: 1-10 [2021-08-01]. <http://kns.cnki.net/kcms/detail/50.1075.TP.20210819.1727.042.html>.
- [15] 贺琪, 杨巧青, 黄冬梅, 等. Self-Att-BiLSTM: 一种面向业务流程活动与时间的多任务预测方法[J]. 激光与光电子学进展, 2021, 58(4): 109-116.
- [16] 徐翔, 靳菁. 基于 Word2vec 的信息窄化测度及影响因素研究[J]. 吉林大学学报(信息科学版), 2021, 39(3): 339-347.
- [17] 潘成龙. 基于深度学习的文本分类问题研究[D]. 武汉: 华中科技大学, 2019.
- [18] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[EB/OL]. [2021-08-01]. <https://arxiv.org/abs/1810.04805v2>.
- [19] YAO L, MAO C, LUO Y. Graph convolutional networks for text classification[EB/OL]. [2021-08-01]. <http://export.arxiv.org/abs/1809.05679>.
- [20] 陈俊芬, 赵佳成, 韩洁, 等. 基于深度特征表示的 Softmax 聚类算法[J]. 南京大学学报(自然科学), 2020, 56(4): 533-540.
- [21] DING H, GU Z, DAI P, et al. Deep connected attention (DCA) ResNet for robust voice pathology detection and classification[J]. Biomedical signal processing and control, 2021, 70: 102973.
- [22] ZHANG X, ZHAO J, LECUN Y. Character-level convolutional networks for text classification[EB/OL]. [2021-08-01]. <https://arxiv.org/abs/1509.01626>.
- [23] VITALE D, FERRAGINA P, SCAIELLA U. Classification of short texts by deploying topical annotations[EB/OL]. [2021-09-20]. https://doi.org/10.1007/978-3-642-28997-2_32
- [24] 马永军, 杨海波. 一种融合本体和最小二乘支持向量机的主题爬行方法[J]. 天津科技大学学报, 2015, 30(3): 72-77.
- [25] 杨观赐, 杨静, 李少波, 等. 基于 Dropout 与 ADAM 优化器的改进 CNN 算法[J]. 华中科技大学学报(自然科学版), 2018, 46(7): 122-127.
- [26] 张鹏, 张再跃. 基于 Attention-CLSTM 模型的商品评论分类[J]. 软件导刊, 2020, 19(2): 84-87.

责任编辑: 郎婧