



# 计算机科学

COMPUTER SCIENCE

## 基于弱化图卷积网络的文本分类

黄玉娇, 陈铭凯, 郑媛, 范兴刚, 肖杰, 龙海霞

引用本文

黄玉娇, 陈铭凯, 郑媛, 范兴刚, 肖杰, 龙海霞. 基于弱化图卷积网络的文本分类[J]. 计算机科学, 2023, 50(6A): 220700039-5.

HUANG Yujiao, CHEN Mingkai, ZHENG Yuan, FAN Xinggang, XIAO Jie, LONG Haixia. [Text Classification Based on Weakened Graph Convolutional Networks](#) [J]. Computer Science, 2023, 50(6A): 220700039-5.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [面向交通流量预测的时空Graph-CoordAttention网络](#)

Spatial-Temporal Graph-CoordAttention Network for Traffic Forecasting

计算机科学, 2023, 50(6A): 220200042-7. <https://doi.org/10.11896/jsjcx.220200042>

### [基于句间信息的图注意力卷积网络的文档级关系抽取](#)

Document-level Relation Extraction of Graph Attention Convolutional Network Based on Inter-sentence Information

计算机科学, 2023, 50(6A): 220800189-6. <https://doi.org/10.11896/jsjcx.220800189>

### [基于情感知识的双通道图卷积网络的方面级情感分析](#)

Aspect-based Sentiment Analysis Based on Dual-channel Graph Convolutional Network with Sentiment Knowledge

计算机科学, 2023, 50(5): 230-237. <https://doi.org/10.11896/jsjcx.220300008>

### [基于深度跨模态信息融合网络的股票走势预测](#)

Deep Cross-modal Information Fusion Network for Stock Trend Prediction

计算机科学, 2023, 50(5): 128-136. <https://doi.org/10.11896/jsjcx.220400089>

### [基于交互注意力和图卷积网络的方面级情感分析](#)

Aspect-level Sentiment Classification Based on Interactive Attention and Graph Convolutional Network

计算机科学, 2023, 50(4): 196-203. <https://doi.org/10.11896/jsjcx.220100105>

# 基于弱化图卷积网络的文本分类

黄玉娇<sup>1</sup> 陈铭凯<sup>1</sup> 郑媛<sup>1</sup> 范兴刚<sup>1</sup> 肖杰<sup>2</sup> 龙海霞<sup>2</sup>

1 浙江工业大学之江学院 浙江 绍兴 312030

2 浙江工业大学计算机科学与技术学院 杭州 310000

**摘要** 文本分类是自然语言处理领域中的经典问题。传统的文本分类模型存在需要人工提取特征,分类准确率不高,难以处理非欧氏空间数据等问题。为了解决上述问题,进一步提高文本分类的准确率,提出了 W-GCN 模型。该模型在 Text-GCN 模型的基础上加以改进,建立了全新的弱化结构模型,用以替换 Text-GCN 模型中对神经元的 Dropout 操作,并通过弱化权重,精确控制弱化力度大小,在一定程度上保留 Dropout 防止过拟合功能的基础上,避免了由直接丢弃神经元造成的特征丢失问题,因此提高了模型分类的准确率。与 Text-GCN 模型相比,基于弱化图卷积网络建立的 W-GCN 模型,在 R8 数据集上准确率提高了 0.38%,在 R52 数据集上准确率提高了 0.62%。实验结果证明了模型改进和弱化结构的有效性。

**关键词:** 图卷积网络;文本分类;文本图构建方法;弱化结构;Dropout

中图法分类号 TP391

## Text Classification Based on Weakened Graph Convolutional Networks

HUANG Yujiao<sup>1</sup>, CHEN Mingkai<sup>1</sup>, ZHENG Yuan<sup>1</sup>, FAN Xinggang<sup>1</sup>, XIAO Jie<sup>2</sup> and LONG Haixia<sup>2</sup>

1 Zhijiang College of Zhejiang University of Technology, Shaoxing, Zhejiang 312030, China

2 College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310000, China

**Abstract** Text classification is a classic problem in the field of natural language processing. The traditional text classification model needs to extract features manually, the classification accuracy is not high, and it is difficult to deal with non-European spatial data. In order to solve the above problems and further improve the accuracy of text classification, the W-GCN model is proposed. This model is improved on the basis of the Text-GCN model, and a new weakened structure model is established to replace the text-GCN model. The dropout operation of neurons, and by weakening the weight, accurately control the weakening strength, and on the basis of retaining the dropout to a certain extent to prevent overfitting, it avoids the loss of features caused by directly discarding neurons, thus improving the accuracy of model classification. Compared with the Text-GCN model, the W-GCN model based on the weakened graph convolutional network improves the accuracy by 0.38% on the R8 dataset and 0.62% on the R52 dataset. The experimental results prove that the model improve and weaken the effectiveness of the structure.

**Keywords** Graph convolutional neural networks, Text classification, Construction method of text map, Weakened structure, Dropout

## 1 引言

文本分类是自然语言处理(Natural Language Processing, NLP)中的一项基本任务。随着网络技术在 21 世纪的迅猛发展,互联网所产生的文本数据也逐渐呈现出了指数增长的趋势。如何将互联网中所产生的巨量文本数据如新闻和垃圾邮件等进行准确、迅速的归类<sup>[1-2]</sup>,以便挖掘其中有价值的信息,是目前文本分类任务所面临的一个重要问题。

在机器学习发展初期,基于统计的机器学习算法是构建文本分类模型的主要方式,如朴素贝叶斯<sup>[3]</sup>、支持向量机<sup>[4-5]</sup>等,这些模型在文本分类应用中具有一定的稳定性和准确性。然而由于这些模型普遍需要人工提取特征,且特征提取工作十分繁琐,因此人工成本较高。

近十年来,深度学习被广泛研究并应用于文本分类任务

中。基于深度学习的文本分类模型相较于传统的机器学习模型,其最大的特点在于无需人工设置节点及提取特征,模型本身通过不断迭代学习实现特征的自动提取。这一改进节省了文本分类任务在提取文本数据特征时所需的人工成本,同时避免了因人工提取特征导致的特征不完全性问题。因此,深度学习模型的准确率普遍高于基于统计的机器学习模型。近年来,随着深度学习和神经网络领域研究的不断深入,涌现出了大量高精度、低成本的优秀模型。其中,卷积神经网络(Convolutional Neural Network, CNN)<sup>[6]</sup>由于具有优秀的特征提取能力和建模能力,被广泛地应用于各个领域。但由于传统的卷积方式在非欧氏空间数据上无法保持数据的平移不变性,因此在文本数据、图数据等数据上表现不佳。为了能够将卷积这一思想运用到非欧氏空间数据上,研究人员通过不断研究和实验,提出了基于谱域的卷积方法和基于空域的卷积方法,实现了在非欧氏

基金项目:国家自然科学基金(61972354,62106225);浙江省自然科学基金(LY20F020024, LZ22F020011)

This work was supported by the National Natural Science Foundation of China(61972354,62106225) and Natural Science Foundation of Zhejiang Pvince, China(LY20F020024, LZ22F020011).

通信作者:黄玉娇(huangyujiao@zjut.edu.cn)

数据上的卷积神经,并运用这些方法构建并提出了图卷积神经网络(Graph Convolutional Network,GCN)<sup>[7-10]</sup>。Defferrard 等则首先将图卷积神经网络运用到文本分类任务中<sup>[11]</sup>。

为了进一步提高文本分类的准确率,本文通过改进在文本分类领域拥有较高分类准确率的 Text-GCN<sup>[12]</sup> 模型,提出了名为弱化图卷积网络(Weaken Graph Convolutional Network, W-GCN)的文本分类模型。主要创新点包括:

(1)提出一种弱化结构方法,替换了 Text-GCN 模型中的 Dropout 操作<sup>[13-14]</sup>。将原本被 Dropout 直接丢弃的神经元改进为对这些神经元的弱化,这一改进不仅在一定程度上保留了 Dropout 防止过拟合的特性,而且还使模型提取到原来被丢弃部分神经元的特征。

(2)设置了弱化权重,通过调整弱化权重的大小,来准确控制弱化结构对于被选中神经元的弱化力度大小。

(3)将本文提出的 W-GCN 模型分别运用于 R8 和 R52 这两个文本分类数据集上。与一些经典模型相比,W-GCN 模型具有更高的准确率。

## 2 研究方法

### 2.1 Dropout

Dropout 可以使模型在训练的过程中随机将一部分神经元丢弃,其主要实现方法为将被丢弃的神经元输出设置为 0。在效果上,Dropout 的提出者 Hinton 认为 Dropout 可以较为有效地防止模型的过拟合,在一定程度上达到了正则化的效果。具体公式<sup>[14]</sup>如下

$$r_i^{(j)} = \text{Bernoulli}(p) \tag{1}$$

$$\hat{y}^{(l)} = r^{(l)} * y^{(l)} \tag{2}$$

其中, Bernoulli 函数表示以概率  $p$ , 随机生成一个 0,1 组成的向量;  $y^{(l)}$  表示第  $l$  层的输出。

Dropout 的可视化效果如图 1 所示。

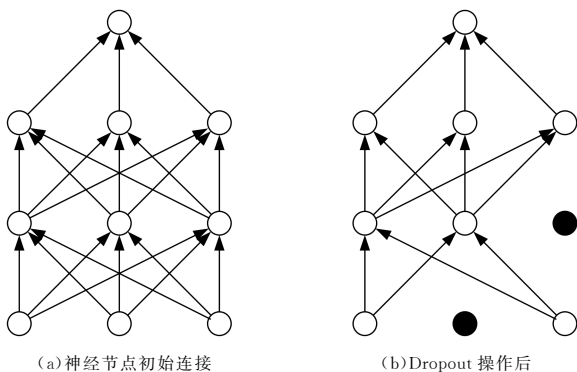


图 1 Dropout 效果对比图  
Fig.1 comparison of dropout effect

### 2.2 弱化结构

弱化结构与 Dropout 的不同之处在于,弱化结构以概率  $p$ , 随机生成一个 0,1 组成的向量之后,用全为 1 的向量减去该向量,并将弱化权重  $W_w$  与被减后的向量相乘,得到一个全新向量,之后该全新向量与该层的输出相乘,并与使用最初向量相乘的该层输出进行叠加处理,最终得出弱化结构的输出。具体公式如下。

$$R^{(j)} = (1 - \text{Bernoulli}(p)) \times W_w \tag{3}$$

$$\hat{y}^{(l)} = (r^{(l)} + R^{(l)}) * y^{(l)} \tag{4}$$

由此,弱化结构构建完成,具体可视化效果如图 2 所示。

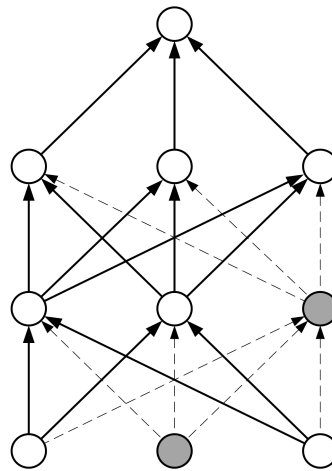


图 2 弱化效果图  
Fig.2 weakening effect

图 2 中,灰色节点代表被随机选中的弱化节点,线条粗细代表节点输出比重。

### 2.3 邻接矩阵构建

#### 2.3.1 Document-Word 的权重构建

Document-Word 之间的权重构建运用了 TF-IDF(Term Frequency-Inverse Document Frequency)<sup>[15]</sup> 这一计算方法,其中 TF 表示词频率。IDF 表示逆向文档频率。主要思想在于如果一个单词或短语在一篇文档中出现的频率较高,但在其他文档中出现的频率较低,那么我们可以认为该词或者该短语拥有很好的区分能力,可以突显文档本身的特征,从而进行文本分类。TF 和 IDF 的具体计算公式如下所示:

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{5}$$

$$idf_i = \log \frac{|D|}{|j:t_i \in d_j|} \tag{6}$$

$$TF-IDF = tf_{i,j} \times idf_i \tag{7}$$

其中,  $|D|$  表示数据集中的文档总数,  $n_{i,j}$  表示单词  $i$  在文档  $j$  中出现的次数,  $\sum_k n_{k,j}$  表示文档  $j$  中所有词出现的次数,  $|j:t_i \in d_j|$  表示包含词  $t_i$  的文档数。

#### 2.3.2 Word-Word 的权重构建

Word-Word 之间的权重采用 PMI(Pointwise Mutual Information) 计算公式。简单概括,PMI 就是在一定的区域内两个词同时出现的频率。区域的具体大小由滑动窗口(window)的大小决定,两个词一起出现的次数越多,PMI 值越大,而 PMI 值越大,也就代表两个词之间的相关性越高。

要计算 Word 和 Word 之间的权重,首先需要构建一个长度大小为  $L$  的 window,在将文档中的词去重之后,通过 window 在文档句子上不断地单向滑动,把一个完整的句子切成一个个长度为  $L$  的片段,由此我们得到了很多长度为  $L$  的词片段  $\#W$ ,之后统计单词  $i$  在 window 里出现的频率  $\#W(i)$ ,再统计 window 中单词  $i, j$  同时出现的频率  $\#W(i, j)$ ,最后通过 PMI<sup>[12]</sup> 的公式得到 Word-Word 之间的权重。

$$PMI(i, j) = \log \frac{p(i, j)}{p(i)p(j)} \tag{8}$$

$$p(i, j) = \frac{\#W(i, j)}{\#W} \tag{9}$$

$$p(i) = \frac{\#W(i)}{\#W} \tag{10}$$

### 2.3.3 总体构建

最后通过结合 Document-word 和 Word-Word 对于图的边权重的构建方法,并给节点本身加上自连接,其余部分补 0,最终构建出邻接矩阵  $A(i,j)$ 。

$$A(i,j) = \begin{cases} TF-IDF_{i,j}, & i \text{ is document, } j \text{ is word} \\ PMI(i,j), & i, j \text{ are words, } PMI(i,j) > 0 \\ 1, & i=j \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

### 2.4 特征矩阵构建

假设图中有  $N$  个节点,每个节点都有  $M$  维特征,提取  $N$  个节点的所有特征构建  $N \times M$  维矩阵  $X$ ,则称  $X$  为特征矩阵。例如假设有 3 个节点,其特征向量分别为  $(1,0,0), (0,1,0), (0,0,1)$ ,则由它们构成的特征矩阵为:

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (12)$$

### 2.5 图卷积神经网络

传统的卷积模型在欧氏空间数据上具有良好的表现,但现实生活中大部分数据为非欧氏空间数据,传统的卷积模型在非欧氏空间数据如图数据上的表现不佳。其主要原因在于,传统的卷积模型在非欧氏空间的数据上无法保持“平移不变性”。为了能够将卷积这一方式推广到图(Graph)等非欧氏空间数据上,学者们对图神经网络进行了深入的研究,并提出了 GCN 这一经典模型。GCN 中最为典型的是二层 GCN 模型,Text-GCN 也是二层 GCN 模型,与普通 GCN 相比仅在图构建方面有所差别,因此本文选择普通二层 GCN 模型进行构建作为样例。

$$Z = f(X, A) = softmax(\hat{A}ReLU(\hat{A}XW^{(0)}))W^{(1)} \quad (13)$$

$$\hat{A} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \quad (14)$$

$$\tilde{A} = A + I \quad (15)$$

其中,  $W^{(0)}$  代表第 1 层的权重,ReLU 为 GCN 的激活函数,  $I$  为单位矩阵,  $\tilde{D}$  为  $\tilde{A}$  对应的度矩阵,  $X$  表示特征矩阵。通过上述公式可以看出,GCN 对输入节点的特征矩阵  $X$  和  $\hat{A}$  以及  $W^{(0)}$  做线性变换,得到第一层的输出,再经过激活函数 ReLU 做第二次线性变换,得到第二层的输出,最后经过 Softmax 得到各类别的概率,从而确定节点的分类结果。损失函数采用交叉熵损失函数。由此,典型的二层 GCN 模型构建完成。之后利用梯度下降法训练 GCN 的权重  $W^{(0)}$  和  $W^{(1)}$ ,在多次迭代之后达到局部最优。具体结构如图 3 所示。

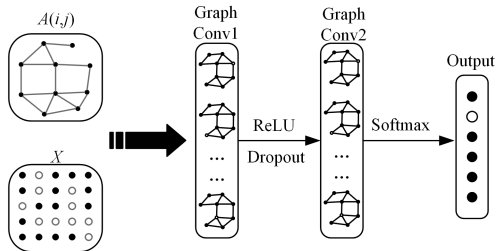


图 3 GCN 结构图

Fig. 3 Structure diagram of GCN

### 2.6 基于 GCN 的改进模型 W-GCN

本节在 Text-GCN 模型的基础上,将第一层图卷积结果

由原本的 Dropout 替换为第 2.2 小节的弱化结构。其原因是 Dropout 为了防止过拟合会随机丢弃一部分神经元,导致一些特征的丢失,使得模型无法很好地学习这一部分被丢失的特征,而使用弱化结构可以使原本被随机丢弃的神经元变为随机效果减弱,因此不仅在一定程度上保留了 Dropout 防止过拟合的作用,又使模型能够学习到被丢弃的那一部分特征,减轻了特征丢失对模型的负面影响,提高了模型分类的准确率。最终提出了 W-GCN 模型,其具体流程描述如下:

(1)将前文构建的异构图作为 Graph Conv1 的输入,对特征矩阵  $X$  做图卷积操作,利用激活函数做非线性变换,得到第一层的输出;

(2)将第一层输出结果送入弱化结构;

(3)将弱化结构的输出结果作为 Graph Conv2 的输入,并对其做图卷积操作;

(4)将第二次卷积的结果运用 Softmax 函数进行映射,得到分类预测结果,利用交叉熵损失函数对预测结果进行计算;

(5)利用梯度下降算法训练 W-GCN 的权重。

W-GCN 的具体流程图如图 4 所示。

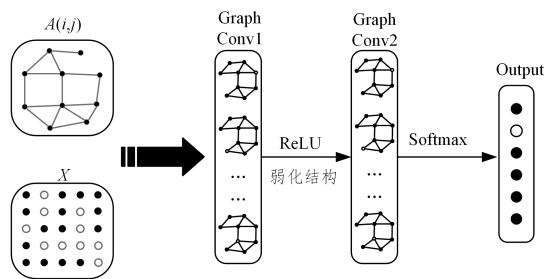


图 4 W-GCN 结构图

Fig. 4 Structure diagram of W-GCN

## 3 实验方法

### 3.1 实验器材

本文实验环境和机器配置如表 1 所列。

表 1 配置表

Table 1 Configuration table

实验环境	环境配置
操作系统	Linux
GPU	NVIDIA T4
编辑语言	Python3.85
深度学习框架	Pytorch

### 3.2 实验数据集

本实验采用 R8 和 R52 数据集。R8 和 R52 数据集是路透社新闻的两个子集,R8 数据集共有 8 个类别,分别为粮食类、收购类、原油类等新闻,并有 5485 个训练文档和 2189 个测试文档。R52 数据集是 R8 数据集的进一步扩充,共有 52 个类别,并有 6532 个训练文档和 2568 个测试文档。R8 和 R52 数据集皆是检测模型文本分类效果的常用数据集,对于模型文本分类效果的检验具有一定的代表性。两个数据集概况如表 2 所列。

表 2 数据集概况统计

Table 2 Data set overview statistics

Dateset	Doces	Training	Test	Words	Classes
R8	7674	5485	2189	7688	8
R52	9100	6532	2568	8892	52

### 3.3 实验评价指标

本实验中,模型的评分标准为常用的分类模型质量评价指标,即准确率(Accuracy),具体公式如下所示:

$$Accuracy = \frac{TP + TN}{N} \quad (16)$$

其中,  $N$  为样本的总体数量;  $TP$  为模型预测样本标号正确,且该样本为正样本的样本数量;  $TN$  为模型预测样本标号正确,且该样本为负样本的数量。

### 3.4 实验超参数设置

本实验运用相同设备,分别在 R8 和 R52 数据集上进行测试。具体超参数设置如表 3 所列。

表 3 超参数设置

Dateset	R8	R52
max_epoch	400	1000
Dropout	0.5	0.5
lr	0.02	0.02
early_stopping	50	1000
nhid	200	200
Weaken_weight	0.3	0.3

### 3.5 实验结果

为准确计算 W-GCN 在两个数据集上的准确率,本文在 R8 和 R52 数据集上分别进行了 5 次实验,之后计算 5 次实验的平均分类准确率和标准差,将计算结果作为 W-GCN 在 R8 和 R52 数据集上的准确率。具体实验结果如表 4 所列。

表 4 W-GCN 的准确率

Table 4 Accuracy of W-GCN

R52	R8
0.9400	0.9753
0.9408	0.9740
0.9424	0.9735
0.9435	0.9740
0.9424	0.9758

为了测试 W-GCN 在文本分类任务上的效果,本文选取了 7 类常用于文本分类任务的机器学习模型进行比较。

(1)TF-IDF+LR:使用词频反比文档频率和词袋模型对单词进行加权,然后使用 Logistic 回归作为分类器。

(2)CNN:一种是随机初始化嵌入词的卷积神经网络模型,另一种是预先训练嵌入词的卷积神经网络模型。

(3)LSTM:一种循环神经网络。相比 RNN,LSTM 通过增加存储单元、输入门、输出门和遗忘门,来有效地保持长序列的历史信息。

(4)Bi-LSTM:一种双向 LSTM,可以将预先训练好的单词嵌入到 Bi-LSTM 模型中进行文本分类。

(5)TextRCNN:第一部分采用 Bi-RNN 来获取文本的上下文特征表示,第二部分采用 max-pooling 进一步提取特征信息,第三部分采用全连接层来获取类别概率。

(6)Graph-CNN-C:使用 Chebyshev 滤波器的图卷积神经网络模型,首先嵌入单词相似图,然后进行卷积运算。

(7)Text-GCN:通过构建基于 TF-IDF 和 PMI 的文本图作为图卷积神经网络的输入,来达到较高的分类准确率。

本文所有实验 window 大小皆取为 20,具体对比测试结果如表 5 所列。加粗表示该模型分类准确率最高。

表 5 W-GCN 实验对比

Table 5 W-GCN Test comparison

Model	R8	R52
TF-IDF+LR <sup>[17]</sup>	0.9374±0.0000	0.8695±0.0000
CNN-rand <sup>[17]</sup>	0.9402±0.0057	0.8537±0.0047
CNN-non-static <sup>[18]</sup>	0.9571±0.0052	0.8759±0.0048
LSTM <sup>[16]</sup>	0.9368±0.0082	0.8554±0.0113
LSTM-(pretra) <sup>[18]</sup>	0.9609±0.0019	0.9048±0.0086
Bi-LSTM <sup>[17]</sup>	0.9631±0.0033	0.9054±0.0091
Text-RCNN <sup>[17]</sup>	0.9553±0.0015	0.9147±0.0011
Grap-CNN-C <sup>[17]</sup>	0.9669±0.0012	0.9257±0.0022
Text-GCN	0.9707±0.0010	0.9356±0.0018
W-GCN	<b>0.9745±0.0008</b>	<b>0.9418±0.00125</b>

通过上述对比实验可以看出,W-GCN 在 R8 和 R52 数据集上的分类准确率高于大多数文本分类的模型。相较于使用相同构图方法,准确率较高且没有使用弱化结构的 Text-GCN,W-GCN 在使用弱化结构替换 Dropout 后,在 R8 数据集上分类准确率提高了 0.38%,在 R52 数据集分类准确率提高了 0.62%。对比实验验证了 W-GCN 模型中弱化结构的有效性。

### 3.6 消融实验

为了进一步探究弱化结构对模型的提升效果,本文依照上文超参数设置,分别对  $r^{(l)}$ ,  $R^{(l)}$  以及  $(r^{(l)} + R^{(l)})$  进行单独实验,探究使用三者的模型在 R8 和 R52 数据集上的分类准确率,具体结果如表 6 所列。

表 6 消融实验对比

Table 6 Comparison of ablation

Model	R8	R52
Text-GCN( $r^{(l)}$ )	0.9707±0.0010	0.9356±0.0018
Text-GCN( $R^{(l)}$ )	0.9676±0.0046	0.9004±0.0076
Text-GCN( $r^{(l)} + R^{(l)}$ )	<b>0.9745±0.0008</b>	<b>0.9418±0.00125</b>

通过实验可以看出,单独使用  $R^{(l)}$  的 Text-GCN 模型,准确率相比使用  $r^{(l)}$  的 Text-GCN 模型,准确率出现了下降的情况,其原因为单独使用  $R^{(l)}$  的 Text-GCN 模型只学习到了被丢弃部分的特征,且这些特征还被进一步弱化,使得模型学习到的特征不够充分,因此模型准确率下降;而使用  $(r^{(l)} + R^{(l)})$  的 Text-GCN 模型,即 W-GCN,则同时学习到了被丢弃部分和未被丢弃部分的两部分特征,相较于前两者,W-GCN 模型学习到的特征更加完善,因此模型准确率高于单独使用  $R^{(l)}$  和  $r^{(l)}$  的 Text-GCN。该实验更好地体现了弱化结构的具体效果。

**结束语** 针对传统的文本分类模型需要人工提取特征,准确率不高,在非欧氏空间数据上分类效果不好等问题,本文在 Text-GCN 的基础上,针对 Dropout 进行了改进,提出了弱化结构,并构建了 W-GCN 模型。与 7 类典型的文本分类模型在 R8 和 R52 数据集上进行的对比实验,表明了 W-GCN 具有较高的分类准确率,实验结果证明了弱化结构的有效性。

W-GCN 准确率的提升也让我们思考,弱化结构是否只能运用于文本分类问题?在其他任务中,如图片识别分类、弱化结构是否依旧有效,能否提高模型的准确率?

当今,互联网已逐渐成为了信息传递的主要媒介,每天都会产生难以计数的文本信息,对这些文本信息进行分类,既有益于互联网的管理,也使普通大众能够快速检索到自身需要的信息,因此文本分类的研究前景依然十分广阔,但文本分类也有着许多难题有待解决。接下来我们对文本分类的研究

将会从以下几方面入手:

(1)随着网络社交的蓬勃发展,短文本、抽象语言、网络热词的不断增多和变化为文本分类带来了新的机遇和挑战,我们将针对不同数据,对模型进行单独调整,以提高对特殊文本的分类准确率;

(2)对弱化工结构进行进一步探究,测试其在不同深度学习领域的效果,并进行进一步的改进。

### 参 考 文 献

- [1] CARRERAS X, MARQUEZ L. Boosting trees for anti-spam email filtering[C]//Proceedings of European Conference Recent Advances in NLP. 2001;58-64.
- [2] AGGARWAL C C, ZHAI C X. A survey of text classification algorithms[M]. US: Mining Text Data, Springer, 2012.
- [3] MCCALLUM A, NIGAM K. A comparison of event models for naive bayes text classification[C] // AAAI-98 Workshop on Learning for Text Categorization. 1998;41-48.
- [4] JOACHIMS T. Text categorization with support vector machines: Learning with many relevant features[C]//Proceedings of the 10th European Conference on Machine Learning. 1998: 137-142.
- [5] WANG X, CHEN R, JIA Y, et al. Short text classification using wikipedia concept based document representation[C]//International Conference on Information Technology and Applications. 2013;471-474.
- [6] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition[C] // Proceedings of the IEEE. 1998; 2278-2324.
- [7] ZHOU J, CUI G, ZHANG Z, et al. Graph neural networks: A review of methods and applications[J]. arXiv:1812.08434, 2018.
- [8] WU Z, PAN S, CHEN F, et al. A Comprehensive Survey on Graph Neural Networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(1): 4-24.
- [9] ZHANG Z, CUI P, ZHU W. Deep Learning on Graphs: A Survey[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(1): 249-270.
- [10] ZHOU J, CUI G, HU S, et al. Graph neural networks: A review of methods and applications[J]. AI Open, 2020, 1: 57-81.
- [11] DEFFERRARD M, BRESSON X, VANDERGHEYNST P, et al. Convolutional neural networks on graphs with fast localized spectral filtering[C]//Neural Information Processing Systems. 2016;3844-3852.
- [12] YAO L, MAO C, LUO Y. Graph Convolutional Networks for Text Classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019;7370-7377.
- [13] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.
- [14] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [15] BLEID M, NG A Y, JORDAN M I. Latent dirichlet allocation [J]. Journal of Machine Learning Research, 2003, 3(1): 933-1022.
- [16] LAI S, XU L, LIU K, et al. Recurrent convolutional neural networks for text classification[C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2015;2267-2273.
- [17] CHEN B. Graph convolution neural network algorithm for text classification[D]. Guilin: Guangxi Normal University, 2021.
- [18] WANG T, ZHU X F, TANG G. Text classification based on graph convolution neural network with knowledge enhancement [J]. Journal of Zhejiang University (Engineering Edition), 2022, 56(2): 322-328.



**HUANG Yujiao**, born in 1985, Ph.D, associate professor. Her main research interests include deep learning, text data analysis and dynamic characteristics of neural networks.