第44卷　第4期　　《新疆师范大学学报》(自然科学版)　　Vol.44,No.4
2025年12月　　Journal of Xinjiang Normal University　　Dec. 2025
(Natural Sciences Edition)

基于混合词嵌入的多通道MOOC评论情感分析模型

王奴建,贾　灿,申红旭,艾孜尔古丽·玉素甫*
(新疆师范大学 计算机科学技术学院,新疆 乌鲁木齐 830054)

摘　要:多通道词嵌入模型在情感分析方面显示出卓越的性能,但往往未能充分利用文本中的图结构,多将各种词嵌入技术进行组合应用。为了解决这一问题,本研究首先计算并归一化文本词汇的TFIDF权重,然后利用权重构建加权图,通过Node2vec训练获取节点嵌入向量。同时,将Word2vec和Bert词向量合并,并结合权值,实现多通道特征学习。最后通过R-Drop结构提升泛化能力。实验结果表明,模型在准确率和F1分数上分别比传统卷积模型提升了3.35%和3.80%,与其他多通道模型相比,也显示出优越性。

关键词:多通道;词嵌入;MOOC;R-Drop;节点词向量
中图分类号:TP391　　**文献标识码:**A　　**文章编号:**1008-9659(2025)04-0037-10

在线课程的质量除了能够通过亲身参与并直接体验得出外,还可以借助课程评论区中其他学习者的评价来了解。学生间的这种互动反馈对其他学习者的选择具有重要的指导意义。

Bansal等人[1]采用Word2vec模型将评论内容通过向量表示,并通过多种机器学习算法验证了词向量表示的多分类器,可以有效提高情感分析任务中的准确率。研究表明,通过结合语义词关系和上下文信息,分布式词向量能够显著提升情感分类的准确性。Joulin等人[2]提出的FastText模型通过训练单词,将其表示平均化作为文档输入,结合分类器进行情感分析。Devlin等人[3]开发了一种新型语言模型,该模型在大规模语料库上进行预训练,并通过微调技术应用于下游情感分析任务。这些预训练语言模型能够在网络中嵌入预训练词向量,从而有效提高模型的性能。除此之外,Glove[4]、Elmo[5]等也能完成单词的预训练工作,但是这些语言模型均建立在序列文本之上,输入形式单一。

为此,本研究提出了一种基于混合词嵌入的多通道MOOC评论情感分析模型。首先在利用传统词向量的同时,构建评论文本的图结构,通过Node2vec[6]训练取得节点词向量。融合两种不同的词向量,为模型提供更多特征,丰富词的语义。本研究模型利用R-Drop,进一步提高了模型泛化能力,从而可以更好地完成MOOC评论文本的情感分析。

1　相关工作

自卷积神经网络(CNN)用于文本分析任务以来,学者们一直利用CNN进行各种探索和尝试。Kim[7]在情感分析中首次运用CNN,并取得了显著的效果。在情感分析任务中,CNN比循环神经网络(RNN)运行时间短,能够利用更深层次的结构获得更多的分布式特征。Guo等人[8]通过为不同的词赋予多个权重并将这些权重应用于词的嵌入,将转换后的特征输入多通道CNN模型中得以实现文本分类。Liu等人[9]利用注意力机制学习两个上下文向量,结合双注意力机制和卷积神经网络进行句子表示和分类,并在基准文本分类

[收稿日期]2025-02-25　　[修回日期]2025-03-19
[基金项目]新疆维吾尔自治区创新环境(人才、基地)建设专项项目(2022D03001);新疆师范大学青年拔尖人才资助项目(XJNUQB2022-22);新疆师范大学智慧教育工程技术研究中心招标课题(XJNU-ZHJY202403);国家自然科学基金项目(61662081);国家社会科学基金项目(14AZD11)。
[作者简介]王奴建(1995-),男,硕士研究生,主要从事自然语言处理文本情感分析方面研究,E-mail:lxznjw@163.com.
*[通讯作者]艾孜尔古丽·玉素甫(1987-),女,副教授,主要从事自然语言处理、智慧教育方面研究,E-mail:Azragul2010@126.com.

数据集上进行测试,证明了模型的有效性。Lee等人[10]提出了一种基于半监督词级情感嵌入的对话情感分类方法,该方法利用NRC情感词汇表,通过添加单词级情感向量进一步获得话语级情感向量,最终基于LSTM实现情感分类。Park等人[11]通过创建单词嵌入层,基于CNN实现了文本情感分类,但该模型未考虑多个句子的上下文信息。陈波等人[12]利用粗糙集与自训练的两种词向量实现了对文本特征的提取,有效提高了向量表示的时间成本。Wang等人[13]利用Microsoft概念图和构建的多语义情感强度词典,获取词语的情感信息,从而为词语提供更准确的语义和情感表达。Pham等人[14]结合Word2vec、Glove的单词嵌入以及one-hot,提出了一种多卷积神经网络模型,在该模型中每个输入的单独表示均由一个卷积神经网络(CNN)处理,该模型在实际情感分类任务中表现良好。范昊等人[15]通过融合Bert、TextCNN、BiLSTM,结合词嵌入信息、文本特征以及上下文信息,显著提升了新闻标题类别识别的准确率。

上述研究中的词嵌入模型主要通过三种方式提升性能:一是通过加权方法改进词向量;二是利用情感词典增强语义表示;三是将先验知识融入情感分析中的单词嵌入。这些方法在理论和实验层面均证明了其有效性,但大多局限于单词的向量表示。图结构本身蕴含着大量丰富的信息,通过学习图的节点嵌入,可以为情感分析提供更多的特征信息。Yao等人[16]运用图卷积网络进行文本分类,通过构建单词文档图,将传统的分类问题转化为节点分类问题,获得了显著的效果。现有的图嵌入模型,如Deepwalk[17]使用均匀随机游动进行搜索,完成节点的训练。LINE[18]则通过邻域相似假设方式完成节点训练。相比之下,Node2vec[6]不仅考虑了深度优先搜索(DFS)邻域,还利用广度优先搜索(BFS)邻域,并且引入两个超参数控制随机游走策略完成节点的训练。本研究通过TFIDF技术构建加权图,利用Node2vec进行节点词向量的训练。这种方法与传统词向量相结合,同时为模型增添图结构的相关信息,显著提高了本研究模型的情感识别准确率。

2 基于混合词嵌入的多通道MOOC评论情感分析模型

2.1 模型构建流程

本研究提出一种基于混合词向量的多通道模型,模型通过Word2vec生成词向量,并结合TFIDF进行加权优化,同时利用Node2vec在依存句法分析和频繁序列构建的加权图上训练节点词向量,将两种向量进行位置嵌入融合。在此基础上,引入R-Drop层,结合Dropout机制和正则化损失函数,有效提升模型的泛化性能和鲁棒性。

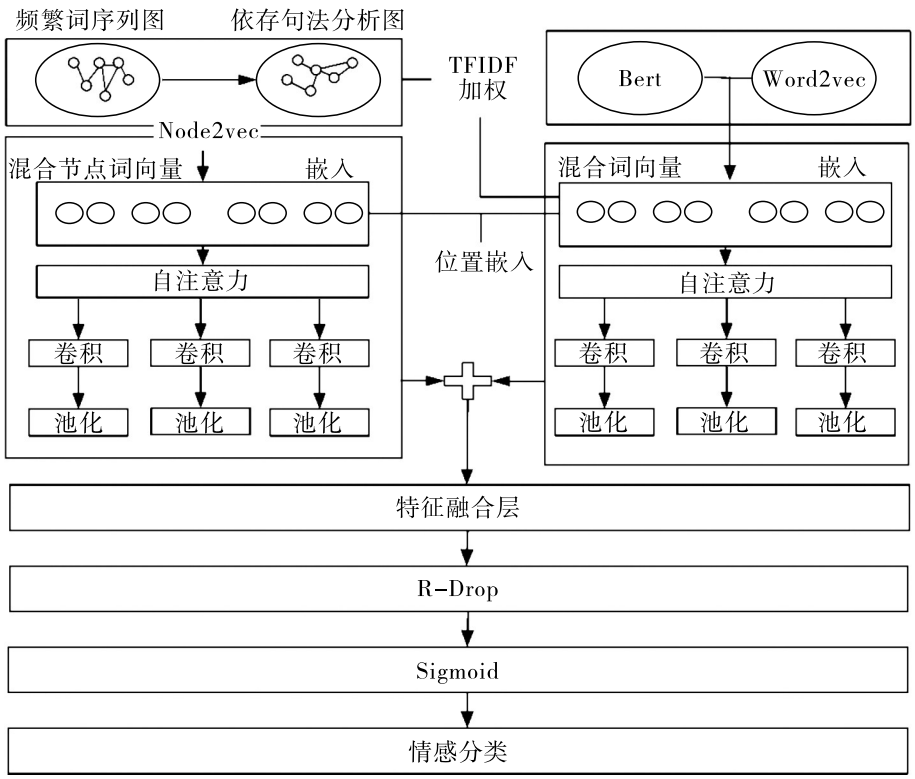

图1 混合词嵌入的多通道模型示意图

2.2 词嵌入层

2.2.1 节点词向量

在自然语言处理领域,图结构无处不在。例如,将一句话中的每个字按顺序连接起来,即可形成一个有向图。然而,这种简单的图结构所包含的信息量十分有限。为了解决这一问题,本研究构建了两种图结构,并将其进行融合,同时利用Node2vec算法进行图特征的学习。

(1)依存句法分析图。在传统的词嵌入模型预训练过程中,往往只是利用句子的上下文,通过遮蔽或者替换的方式生成词向量。句子除了顺序结构外,还包含许多其他关系。语法结构对于识别文本情感模型非常重要,尤其是对于包含相同词汇集但不同语法的课程评论。例如,“课程不好吗?”“课程不好!”这两句话中,前者表示惊讶,后者则表达消极的情感。依存句法分析能够依据特定的语法,分析句子中的句法单位及其之间的联系[19]。依存图用有向弧来表示,使用spacy进行依存句法分析产生的示例(图2)。

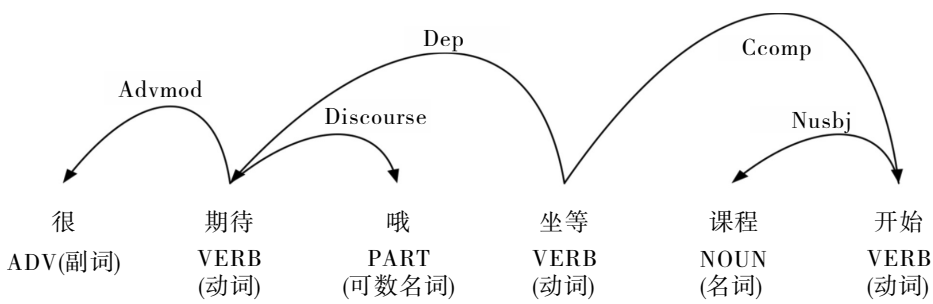

图2 依存图示例

以“很期待哦,坐等课程开始”为例,这句话主要包含五种依存关系,“很”与“期待”是状语关系,“期待”与“哦”是话语关系,“坐等”与“期待”是限定关系,“坐等”与“开始”是从句补充关系,“开始”与“课程”是名词性主语关系。利用spacy工具对文本评论数据进行句法依存分析,构建图结构,由此可以得到一个邻接矩阵 $A = X^{nxn}$.

(2)频繁词序列图。在情感分析中经过分词处理的每一句话均可视为一个项的集合。通过设定不同的支持度阈值,可以控制生成的频繁项,从而实现对评论数据的有效挖掘。支持度是指相关联的数据项 X 和 Y 在数据集中出现的频次与总数据集的比值,公式如下

$$Support(x,y) = \frac{number(x,y)}{num(Allsamples)} \tag{1}$$

在构建过程中,将节点w1和w2之间的边权设置为TFIDFw1与TFIDFw2之和。得到加权图后,将其输入Node2vec模型中进行节点词向量的训练,以获取节点词向量。

2.2.2 混合词向量

设计的词向量随机混合方法包含两部分,第一部分是利用Word2vec自训练的词向量,第二部分是Google基于Bert预训练的中文词向量,具体的混合方法如图3所示。

在词向量混合过程中,首先统计文本的高频词和低频词,然后利用TFIDF算法进一步确定文本特征词,其中特征词采用文章自训练的词向量,其余词则采用Bert预训练的中文词向量。将获得的混合词向量与相应词的归一化TFIDF值相乘,最终得到加权词向量。计算公式如下

$$W_{v_i'} = W_{tfidf} W_{v_i} \tag{2}$$

其中,W_{v_i} 是单词 W 的混合词向量,W_{tfidf} 是单词 W 的TFIDF值。

2.2.3 位置嵌入

由于模型未采用RNN结构,因此难以有效捕捉文本的位置信息。为了解决这一问题,将位置信息进一步融合到预训练的词向量中,从而得到融合词向量。采用Vaswani等人[20]提出的位置编码方式,计算公式如下

$$\mathrm{PE}(pos,2i) = \sin\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right) \tag{3}$$

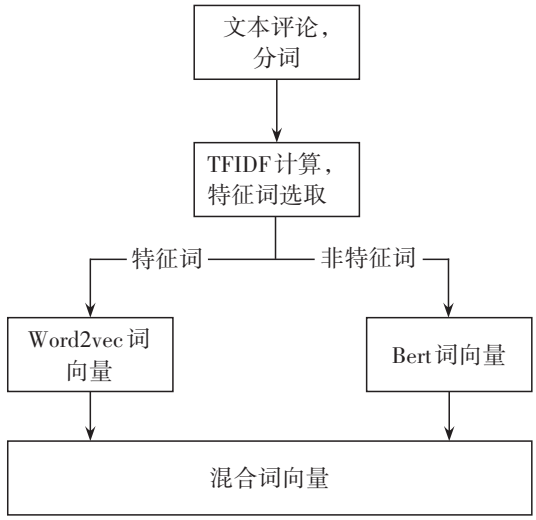

图3 混合词向量流程图

$$\mathrm{PE}(pos,2i+1)=\cos\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right) \tag{4}$$

其中，pos表示词语在文本中的位置，d_{model}代表向量的嵌入维度，i表示嵌入向量的索引。

每个词位置嵌入后的表示如下

$$mv_i = wv_i + pv_i \tag{5}$$

其中，wv_i表示文本第i个位置的词向量，pv_i表示文本第i个位置的位置向量。

2.3 特征融合层

通过不同的卷积池化操作，获得了不同的特征。这些特征通过向量操作进行拼接，实现了特征融合。具体操作如下

$$X_m = \sum_{k\in 3,5,7}^{i=1} cni_k \tag{6}$$

其中，X_m表示融合后的向量，cni_k代表经过通道i、窗口为k的卷积池化后得到的特征向量。

2.4 R-Drop层

正则化是深度学习中用来提高模型泛化能力，改善过拟合现象的一种有效手段。Liang等人[21]通过引入一种简单的正则化策略R-Drop，在神经机器翻译、文本摘要和图像分类等深度学习任务中验证了其有效性。该研究主要基于Transformer模型，对其他模型进行实验，将R-Drop应用于CNN模型。R-Drop的工作原理如图4所示。

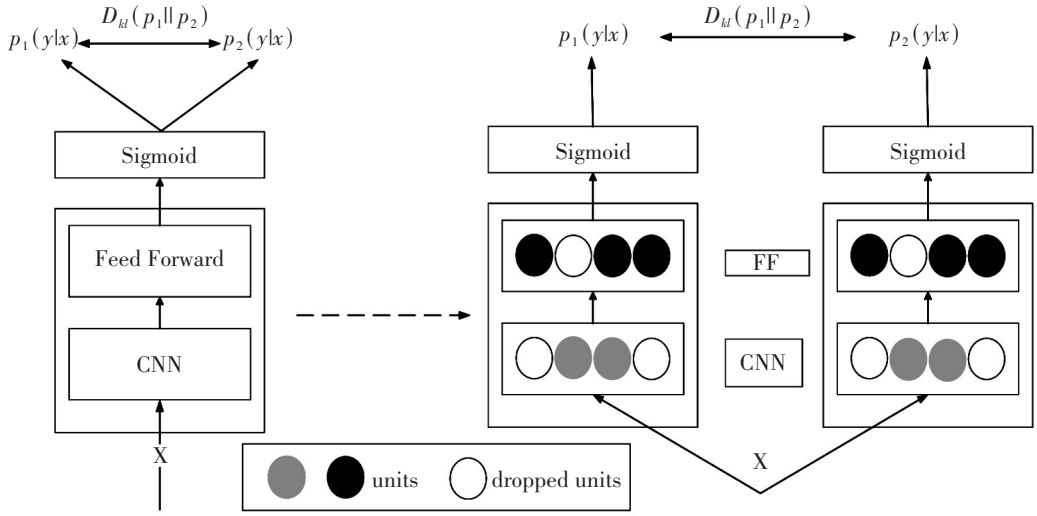

图4 R-Drop原理图

R-Drop通过最小化两个分布之间的双向Kullback-Leibler(KL)散度,促使两个子模型对相同数据样本的输出分布保持一致。

假设训练数据为$\{x_i, y_i\}$,模型为$p_q(x,y)$,采用交叉熵函数作为损失函数,如式(7)所示

$$L_i = -\log p_\theta(x_i, y_i) \tag{7}$$

R-Drop作用于通过Dropout采样的子模型的输出。具体而言,在每个小批量训练中,每个数据样本经过两次前向传递,每次传递由不同的子模型通过随机删除部分隐藏单元来处理。此时,R-Drop的损失函数分为两部分:一部分是常规交叉熵,如式(8)所示

$$L_i^{ce} = -\log p_\theta^{(1)}(x_i, y_i) - \log p_\theta^{(2)}(x_i, y_i) \tag{8}$$

另一部分是对称KL散度,如式(9)所示

$$L_i^{(kl)} = 1/2\left[KL(p_\theta^{(1)}(y|x_i)\|p_\theta^{(2)}(y|x_i)) + KL(p_\theta^{(2)}(y|x_i)\|p_\theta^{(1)}(y|x_i))\right] \tag{9}$$

最终损失函数为这两部分的加权和,如式(10)所示

$$L_i = L_i^{(ce)} + \alpha L_i^{(kl)} \tag{10}$$

3 实验

3.1 实验参数设置和数据集

3.1.1 实验数据集

本实验使用的语料源于中国大学MOOC网站的课程评论。收集的评论评分范围在1~5分之间,因此将一星评论定义为消极情感,五星评论定义为积极情感。经过筛选,共获得4万条评论,其中积极评论和消极评论各2万条。浏览后发现语料标注存在不准确的情况,因此需要人工对语料进行再次标注。经过人工再次标注的数据被用作情感分类实验的语料,其中部分评论信息如表1所示。

表1 MOOC评论数据集示例

标签	评论
1	好的大学没有围墙,能有自己喜欢的很开心。
0	第二个男老师鼻音很重,听都听不清,讲得也不太好。

3.1.2 实验参数设置

实验参数设置如表2所示。其中,Node2vec_Size、Bert_Size和Word2vec_Size表示词向量的嵌入维度。

表2 实验参数

参数	参数值
Node2vec_Size	125
Node2vec_W	5
p	0.5
q	2
α	5
Bert_Size	125
Epoch	20
Word2vec_Size	125
Word2vec_W	5
Batch_size	32
CNN_W	[3,5,7]
Dr	0.4
Lr	0.001

图5展示了Word2vec与Node2vec在不同词嵌入维度下TextCNN模型获得的准确率。结果表明当维度为125时,模型能够取得较高的准确率。Node2vec_W和Word2vec_W表示训练的窗口大小;Batch_size表示数据输入的批次大小;CNN_W表示卷积窗口大小;Dr表示Dropout的比率;Lr表示学习率;Epoch表示训练轮数;α表示KL散度的权值。

图6展示了不同α取值下,本研究所提模型在采用Dropout正则化策略时的准确率。结果显示,当α取值为5、Dropout取值为0.4时,模型能够获得较高的准确率。

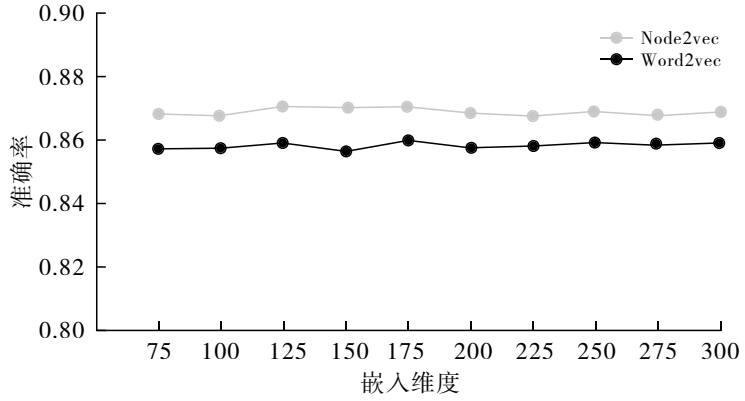

图5 不同嵌入维度模型准确率

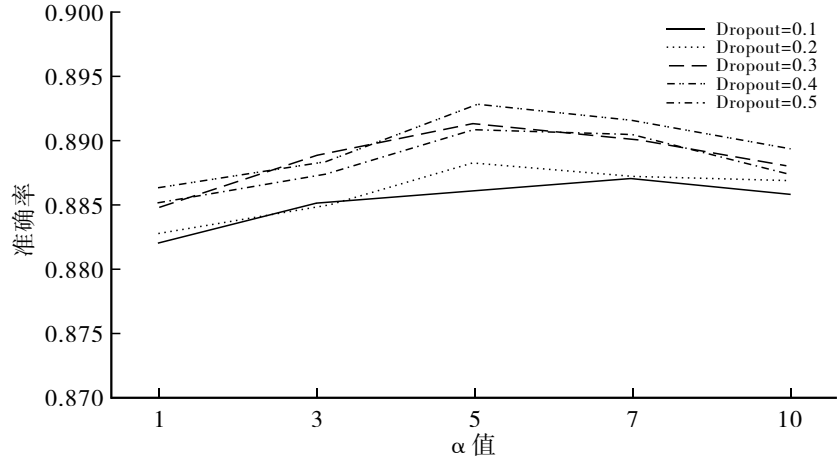

图6 不同α值与Dropout模型准确率

参数p被称为进出参数,参数q被称为返回参数。为探究p和q的最佳参数值,本实验将节点的邻域W设置为5,从节点出发游走序列数目设置为10,游走长度设置为80,节点向量维度设置为125. 采用TextCNN作为基线模型,训练10轮。经测试,当参数在p=0.5、q=2时,模型获得最高准确率,相关数据如表3所示。

表3 不同p、q参数下的准确率 单位:%

p	q				
	0.25	0.5	1	2	4
0.25	86.72	86.70	86.52	86.68	86.26
0.5	86.20	86.72	86.41	87.07	86.82
1	86.65	86.02	86.60	86.70	86.57
2	86.78	86.90	86.82	86.75	86.50
4	86.21	86.96	86.20	86.83	85.96

3.2 实验分析

3.2.1 对比实验

为验证该模型在MOOC评论情感分类上的有效性,通过对比以下模型的实验结果进行分析。选取的对比模型如下:

SVM:支持向量机,一种经典的机器学习分类方法。

TextCNN[7]:使用卷积网络进行文本情感分析,采用随机初始化进行词嵌入。

TextRNN:利用循环神经网络进行文本情感分类。

TextCNN-Att:在卷积网络基础上添加注意力机制,用于文本情感分析。

Fasttext[2]:基于自训练的词嵌入模型,使用单词嵌入均值表示文档,利用分类器进行情感分析。

MWCNN[8]:为单词设置不同权重并应用于单词嵌入中,最后输入CNN中实现情感分类。

MCNN[14]:基于CNN,利用多种单词嵌入方式实现文本情感分析。

Our model:本研究模型。

表4 MOOC评论数据集实验结果

Model	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
SVM	84.70	84.90	83.97	84.43
TextCNN	85.92	85.71	86.11	85.91
TextRNN	87.09	87.75	86.11	86.92
TextCNN-Att	87.32	86.32	88.61	87.45
Fasttext	87.42	85.88	89.22	87.52
MWCNN	88.20	87.26	88.94	88.09
MCNN	88.26	87.27	89.56	88.40
Our model	89.27	87.08	92.50	89.71

由表4可知,各模型在MOOC评论数据集上的表现存在显著差异。其中,SVM模型的各项指标明显低于深度学习模型,这表明深度学习通过学习更深层次的特征信息,能够显著提升模型的性能。在深度模型中,模型在所有评价指标上均取得最优结果,与TextCNN、TextRNN和TextCNN-Att三个基础模型的准确率相比,其准确率分别提高了3.35%、2.18%和1.95%. 具体而言,TextCNN-Att在CNN基础上引入注意力机制,有效增强了模型的学习能力。与Fasttext模型相比,多通道模型通过构建加权词向量并添加位置向量,显著提高了模型准确率。此外,相较于标准的TextCNN模型,MWCNN对单词进行加权处理,性能得到了明显提升,这验证了单词加权策略的有效性。MCNN模型通过多通道的优势取得了仅次于本模型的效果,但其仅采用传统的词向量,未考虑其他向量形式,因此未能充分发挥多通道的潜力。多通道模型结合经典的Bert词向量与自训练的Word2vec词向量,构建了混合词向量,并结合节点词向量,有效扩展了模型的特征输入形式,充分发挥了多通道的优势,从而获得了理想的模型性能。此外,模型的召回率偏高,这应该与输入文本的长度有关,积极评论比消极文本的平均长度要短4个字符。

3.2.2 不同向量初始化实验

为探究初始向量方式对模型性能的影响,选用TextCNN作为基线模型,分析词嵌入层参与和不参与神经网络训练时的各项指标。其中,freeze参数设置为False,表示Embedding层参与神经网络的训练,freeze值为True,则表示Embedding层不参与训练。采用以下词向量表示方式:R代表随机词嵌入,B代表Bert词向量,W代表Word2vec词向量,H代表混合词向量,N代表节点向量。

从表5可以看出,采用随机初始化方式进行词嵌入时,各项指标均低于其他方法,进一步验证了预训练模型的有效性。通过Word2vec训练得到的词向量准确率仅比随机初始化方式高0.3个百分点,这可能与预训练语料库的规模有关。相比之下,虽然Bert词向量的准确率比Word2vec高0.05个百分点,但其F1值却低0.34个百分点。由于Bert词向量已经过降维处理,差异可能与正面和负面情感文本的长度以及词向量维度有关。将两者混合的词向量结合了各自的优点,因此各项指标较为均衡。Node2vec在所有方法中表现最佳,其准确率比随机初始化方式高1.14个百分点,这表明节点嵌入的有效性。由此可见,针对小规模语料,图结构所蕴含的特征比单纯的序列语言模型能够挖掘出更多信息。此外,文本的构图方式也可能存在一定影响,当在图中添加频繁词序列特征时,准确率略有提升。

表5 不同初始化方式比较

向量初始化方式	freeze	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
R	False	85.92	85.71	86.11	85.91
	True	77.91	77.10	77.60	77.35
B	False	86.07	85.86	86.17	86.01
	True	78.51	78.99	76.39	77.67
W	False	86.02	83.81	89.04	86.35
	True	77.96	71.29	81.74	76.16
H	False	86.65	86.38	86.73	86.55
	True	78.67	78.38	78.46	78.42
N(只有依存图结构)	False	86.96	85.42	89.04	87.19
	True	84.65	83.86	84.59	84.22
N(添加序列图结构)	False	87.06	85.91	88.36	87.12
	True	84.88	82.91	87.00	84.91

3.2.3 不同权值归一化方式比较

不同归一化处理方式对词汇权值的影响各异。在利用词汇权值处理经典词向量和构建图结构时,这些处理方式会对模型产生一定影响,实验分析不同条件下模型的性能。

由表6可知,进行归一化处理后模型性能的提升程度不同。当采用Min-Max标准化处理TFIDF值时,相较于未经归一化处理模型性能分别提高了0.23和0.34个百分点。这一提升可能是由于原始数据分布较为集中,归一化处理后的结果表现更佳。然而,Z-score标准化处理反而降低了模型性能,这可能是由于该处理使数据符合正态分布,从而削弱了TFIDF的特性。相比之下,百分比归一化方法在一定程度上缓解了TFIDF值过大的问题,并使模型性能有所提升。综上所述,通过TFIDF计算词的权重特征,并采用Min-Max标准化处理后构建加权词向量和加权图,有效提升了模型性能。

表6 不同权值归一化比较

	归一化方式	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
Model	无	89.04	86.54	92.40	89.37
	Z-score标准化	89.08	86.51	92.54	89.43
	百分比归一化	89.11	87.44	91.28	89.32
	Min-Max标准化	89.27	87.08	92.50	89.71

3.2.4 通道与特征消融实验

为验证通道数目与特征添加的有效性,本研究开展以下实验。单通道模型采用TextCNN,多通道模型则保持模型结构,但采用不同嵌入方式进行实验。其中,R表示随机词嵌入,H表示混合词向量,G表示图节点向量。

表7 通道与特征消融实验

通道数	Model	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
单通道	R	85.92	85.71	86.11	85.91
	H	86.65	86.38	86.73	86.55
	G	87.06	85.91	88.36	87.12
多通道	R+R	85.90	84.35	88.52	86.38
	H+H	86.67	84.16	91.01	87.45
	G+G	87.18	84.45	90.54	87.39
	R+H	86.55	84.71	89.68	87.12
	R+G	87.07	84.09	88.67	86.32
	H+G	88.65	86.10	92.31	89.09
	H+G+R-Dorp	89.27	87.08	92.50	89.71

由表7可知,相较于采用随机词嵌入的多通道模型,单通道模型的准确率和F1值相差并不明显。若只是设置多通道而没有构建不同特征输入,只会增加模型的复杂度。而采用多通道模型和随机词嵌入方式与其他方式的组合,模型性能仍取决于另一词向量的好坏。本研究利用TFIDF构建了混合词向量,依据依存分析和频繁序列构建了加权图,通过训练获得了节点词向量。在多通道实验中,采用节点词向量与混合向量的形式比其他组合方式取得了更高的准确率和F1值。相较于随机词嵌入,准确率和F1值分别提高了2.75%和2.71%,可见本研究构建词向量的有效性。通过添加R-Drop,模型的准确率和F1值有了进一步的提高,可见R-Drop通过改进Dropout,进一步提升了模型的性能。

4 总结

构建节点词向量与混合词向量,通过加权、添加位置信息,进一步完善词向量,利用多通道和分层R-Drop实现了对MOOC评论文本的情感分析。在实验分析中,首先通过对比分析模型,分析了模型的优点和性能;针对不同嵌入方式进行了分析比较;接下来通过分析不同权值归一化方式,进行分析比较;最后进行通道与特征消融实验,验证了多通道的优势与最佳向量组合的方式。本模型从词向量的角度出发,完成了对MOOC评论数据集的分析,这对后续的教学评价情感分析具有一定的现实意义。本研究模型使用TFIDF方法计算词的权重,未来将尝试其他加权方法。

参考文献:

[1] BANSAL B, SRIVASTAVA S. Sentiment Classification of Online Consumer Reviews Using Word Vector Representations[J]. Procedia Computer Science, 2018, 132: 1147-1153.

[2] JOULIN A, GRAVE E, BOJANOWSKI P, et al. Bag of Tricks for Efficient Text Classification[J]. arXiv preprint arXiv: 1607.01759, 2016.

[3] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. arXiv preprint arXiv: 1810.04805, 2018.

[4] 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 (EMNLP), 2014: 1532-1543.

[5] PETERS M E. Deep Contextualized Word Representations[J]. arXiv preprint arXiv: 1802.05365, 2018.

[6] GROVER A, LESKOVEC J. Node2vec: Scalable Feature Learning for Networks[C]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 855-864.

[7] KIM Y. Convolutional Neural Networks for Sentence Classification[C]. 2014: 1746-1751.

[8] GUO B, ZHANG C, LIU J, et al. Improving Text Classification with Weighted Word Embeddings Via a Multi-channel TextCNNModel[J]. Neurocomputing, 2019, 363: 366-374.

[9] LIU T, YU S, XU B, et al. Recurrent Networks with Attention and Convolutional Networks for Sentence Representation and Classification[J]. Applied Intelligence, 2018, 48(10): 3797-3806.
[10] LEE Y J, PARK C Y, CHOI H J. Word-level Emotion Embedding based on Semi-supervised Learning for Emotional Classification in Dialogue[C]. 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), 2019: 1-4.
[11] PARK S H, BAE B C, CHEONG Y G. Emotion Recognition from Text Stories Using an Emotion Embedding Model[C]. IEEE International Conference on Big Data and Smart Computing (BigComp), 2020: 579-583.
[12] 陈波,谢珺,苗夺谦,等. 基于粗糙集和多通道词向量的中文文本情感特征分析[J]. 中文信息学报, 2020, 34(08): 94-104.
[13] WANG Y, HUANG G, LI J, et al. Refined Global Word Embeddings based on Sentiment Concept for Sentiment Analysis[J]. IEEE Access, 2021, 9: 37075-37085.
[14] PHAM D H, LE A C. Exploiting Multiple Word Embeddings and One-hot Character Vectors for Aspect-based Sentiment Analysis[J]. International Journal of Approximate Reasoning, 2018, 103: 1-10.
[15] 范昊,何灏. 融合上下文特征和BERT词嵌入的新闻标题分类研究[J]. 情报科学, 2022, 40(06): 90-97.
[16] YAO L, MAO C, LUO Y. Graph Convolutional Networks for Text Classification[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019: 7370-7377.
[17] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk: Online Learning of Social Representations[C]. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014: 701-710.
[18] TANG J, QU M, WANG M, et al. LINE: Large-scale Information Network Embedding[C]. Proceedings of the 24th International Conference on World Wide Web, 2015: 1067-1077.
[19] 陈强,何炎祥,刘续乐,等. 基于句法分析的跨语言情感分析[J]. 北京大学学报(自然科学版), 2014, 50(01): 55-60.
[20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is All You Need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
[21] LIANG X, WU L, LI J, et al. R-Drop: Regularized Dropout for Neural Networks[J]. arXiv preprint arXiv: 2106.14448, 2021.
[22] PORIA S, HAZARIKA D, MAJUMDER N, et al. Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research[J]. IEEE Transactions on Affective Computing, 2023.
[23] 郑智雄,刘建华,孙水华,等. 面向方面级情感分析的交互关系图注意力网络[J]. 计算机工程与应用, 2023, 59(15): 187-195.
[24] LIANG S, WEI W, MAO X L, et al. BiSyn-GAT+: Bi-Syntax Aware Graph Attention Network for Aspect-based Sentiment Analysis[J]. arXiv preprint arXiv: 2204.03117, 2022.
[25] 李峰. 一种支持否定授权的RBAC优化模型[J]. 新疆师范大学学报(自然科学版), 2025, 44(03): 11-16.

A Multi-channel MOOC Review Sentiment Analysis Model based on Mixed Word Embeddings

WANG Nu-jian, JIA Can, SHEN Hong-xu, AIZIERGULI·Yusufu*

(*College of Computer Science and Technology, Xinjiang Normal University, Urumqi, Xinjiang, 830054, China*)

Abstract: Multi-channel word embedding models demonstrate excellent performance in sentiment analysis, but they often fail to fully exploit the graph structure in text, they simply combine various word embedding techniques. To address this issue, the method proposed in this paper first calculates and normalizes the TFIDF weights of the text vocabulary, then uses these weights to construct a weighted graph. Node2vec is used to train and obtain node embedding vectors. Meanwhile, Word2vec and Bert word vectors are merged and combined with weights to achieve multi-channel feature learning. Finally, Rdrop structure is utilized to enhance generalization ability. Experimental results show that, compared to traditional convolutional models, the proposed method achieves 3.35% improvement in accuracy and 3.80% enhancement in F1 score. Additionally, it also exhibits superior performance compared to other multi-channel models.

Keywords: Multi-channel; Word embedding; MOOC; R-drop; Node word vector