2024 年 第 24 卷 第 20 期
2024, 24(20):08598-08

科 学 技 术 与 工 程
Science Technology and Engineering

ISSN 1671—1815
CN 11—4688/T

DOI:10.12404/j.issn.1671-1815.2402303

引用格式:张昱, 冯亚寒, 丁千惠. 融合 Word2Vec 词嵌入的多核卷积神经网络音乐歌词多情感分类方法[J]. 科学技术与工程, 2024, 24(20): 8598-8605.

Zhang Yu, Feng Yahan, Ding Qianhui. Multi-sentiment classification of music lyrics by incorporating Word2Vec word embedding multi-core convolutional neural networks[J]. Science Technology and Engineering, 2024, 24(20): 8598-8605.

融合 Word2Vec 词嵌入的多核卷积神经网络音乐歌词多情感分类方法

张昱[1,2,3], 冯亚寒[1,2*], 丁千惠[1,2]

(1. 北京建筑大学电气与信息工程学院, 北京 100044; 2. 北京建筑大学建筑大数据智能处理方法研究北京市重点实验室, 北京 100044; 3. 中国矿业大学(北京)深部岩土力学与地下工程国家重点实验室, 北京 100083)

摘　要　目前,音乐歌词情感分类大多以二标签极性情感为主,多情感标签分类较少,并且对于情感性不确定的歌词而言,得到的分类性能不高。为了解决多情感标签研究分类的不足,以及提高分类准确性,提出一种利用 Word2Vec 词嵌入技术,并使用多核卷积神经网络作为分类器的音乐歌词多情感分类方法。该方法首先结合音乐歌词文本,进行数据预处理和可视化分析;其次利用 Word2Vec 词嵌入提取歌词局部特征,构建特征情感向量,挖掘歌词中情感信息,将歌词转化为更利于分类器模型输入的词向量;最后在分类器中,选用卷积神经网络模型,并在此基础上采用不同高度卷积核的方式构建新模型以此得到多情感分类。结果表明:音乐歌词多情感分类的结果达到 94.26%,与传统 CNN 相比,分类精确率提高了 6.86%,取得了良好性能。

关键词　自然语言处理; 情感分类; 卷积神经网络; 词嵌入; 文本分类; 音乐歌词

中图法分类号　TP391.1;　　**文献标志码**　A

Multi-sentiment Classification of Music Lyrics by Incorporating Word2Vec Word Embedding Multi-core Convolutional Neural Networks

ZHANG Yu[1,2,3], FENG Ya-han[1,2*], DING Qian-hui[1,2]

(1. Beijing Key Laboratory of Intelligent Processing for Building Big Data, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 2. School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 3. State Key Laboratory for GeoMechanics and Deep Underground Engineering, China University of Mining & Technology, Beijing 100083, China)

[**Abstract**]　Currently, most of the music lyrics emotion classification is based on two-label polar emotions, while multi-emotion label classification is rare, and the classification performance obtained is not high for lyrics with uncertain emotionality. To address the limitations of multi-sentiment labeling research in classification and to enhance classification accuracy, a multi-sentiment classification method was proposed for music lyrics using Word2Vec word embedding technology and a multi-core convolutional neural network as the classifier. The method initially integrated music lyrics text for data preprocessing and visualization analysis. Secondly, Word2Vec word embedding was utilized to extract local features of the lyrics, construct feature sentiment vectors, mine sentiment information within the lyrics, and converted the lyrics into word vectors that were more suitable for input into the classifier model. Finally, a convolutional neural network model was selected as the classifier, and upon this foundation, a novel model was constructed with various heights of convolutional kernels to achieve multi-emotion classification. The experimental results show that the result of music lyrics multi-sentiment classification reaches 94.26%, which improves the classification accuracy by 6.86% compared with the traditional CNN and achieves good performance.

[**Keywords**]　natural language processing; sentiment classification; convolutional neural network; word embedding; text classification; music lyrics

随着人们生活水平不断提高,情感领域也在不断地丰富,而音乐是一种善于表达情感的艺术载体,它所表达的情感往往来源于生活[1],人们希望可以根据心情来选择不同情感类型的音乐映射当

收稿日期: 2024-04-01　　**修订日期**: 2024-07-05

基金项目: 北京市深地空间科学与工程研究院基金(XD2021021);北京市教育部产学合作协同育人项目(221001576090901);北京市北京建筑大学研究生教育教学质量提升项目(J2022003)

第一作者: 张昱(1979—),男,汉族,内蒙古呼和浩特人,博士,副教授。研究方向:岩爆、人工智能、智慧城市。E-mail:zhangyuscholar@bucea.edu.cn。

* **通信作者**: 冯亚寒(2000—),女,汉族,山东潍坊人,硕士研究生。研究方向:大数据、人工智能。E-mail:991362201@qq.com。

下的心理活动。针对网络中产生的这些歌词信息,自动准确地提取这些信息和隐含的情感并进行分析,对用户、商家和研究者来说都具有重要意义[2]。情感分析和分类是计算机科学的重要分支之一,指通过运用自然语言处理技术,从而识别和分类文本中隐藏的情感和观点,这一过程中能够将文本内容系统地识别、提取、量化并深入研究,所以也被称作文本的上下文挖掘。情感分析的应用范围广泛,对于歌词来说,可以用来针对不同的情感特征来分析人们隐藏在内的的潜在情绪[3]。

关于音乐歌词情感分类的研究已大量报道,然而,随着歌词情感分类的需求和重要性不断增加,传统的人工分类无法满足人们的要求,随之而来的机器学习和深度学习使得歌词分类效率显著提高。Trohidis 等[4]通过 6 组音乐分类实验,提出多标签建模,以及该算法和特征的预测质量。王小凤等[5]提出了一种基于短语情感分析方法,通过从句子中提取短语进行情感识别,依次分析结果为消极,还是积极,但这种算法忽略了句子本身的细节问题,得出来的结果并不准确。Kim[6]提出将卷积神经网络模型(convolutional neural networks, CNN)用于音乐分类,并针对机器学习和深度学习进行对比实验,结果表明,机器学习分类结果率要低于深度学习模型。龚安等[7]提出了跨媒体分词模型的情感分析方法,然而该方法无法处理歌词信息中包含与类别标签相同的情感词问题,并且情感识别率也比较低。王洁等[8]对中文歌词的音乐情感分类方法作出研究,并且提到了利用 Word2Vec 技术来构建中文情感词典,再利用 *K*-邻近算法(*K*-nearest neighbors, KNN)得到分类结果。Palanisamy 等[9]利用 ImageNet 预训练的 CNN 模型作为音频分类的网络,使模型适应音乐特征,通过定性可视化对不同的两个数据集分别进行分类实验,显示较好的性能和分类效果,实现了音乐的音频分类。Tong[10]设计了一种基于深度信念网络(deep belief net, DBN)的音乐情感分类方法,基于音乐的频率和频带能量对传统的 DBN 网络增加了微调节点,从而加强模型的可调性,并融合传统的支持向量机(support vector machine, SVM)分类算法使得音乐得到多情感分类。Yang 等[11]提出一种基于歌词和评论的音乐情感分类方法,在构建情感向量时综合考虑情感词的词频、情感强度和词性对情感分类的影响,并将音乐的歌词和评论中的实体词与构建的情感词典相匹配,最终利用 CNN 模型分类器输出分类结果。

前人研究都取得了一定的成果,并使得音乐分类进行了一定的发展,但仍存在以下不足:一是从数据集的选择来看,音乐分类模型的输入比较偏向于音乐的音频,少数注重音乐歌词本身的情感性;二是从输出结果来看,传统歌词情感分类方法主要通过关键词匹配技术来识别歌词中包含的情感特征词,随后计算情感分数以预测文本所表达的是正面还是负面情绪[12],因此大多输出情感结果为二标签分类,但是歌词中包含丰富的情感词,应该将音乐根据多种情感进行分类[13];三是从模型选择来看,传统的机器学习[如朴素贝叶斯(naive Bayes model, NB)、SVM 等]多存在分类时间长、鲁棒性差、易陷于局部最优等缺陷,而深度学习[如循环神经网络(recurrent neural network, RNN)]的出现很好地解决了以上问题。RNN 在自然语言处理中是常见的模型,但针对音乐歌词情感分类的研究中,CNN 是主流模型。在 CNN 模型上添加卷积核的方式在其他数据集上有所应用,如新闻、微博评论等,但至今为止并没有在结合音乐歌词的同时利用此方式来实现音乐的多情感分类。

鉴于此,构建符合音乐情感特征的词向量和选择适合音乐歌词多情感分类模型[14]。选用 Word2Vec 词嵌入技术构建音乐歌词领域中的情感词典,研究 CNN 在音乐歌词中实现多情感分类的性能,并进行改进,设计一种多核 CNN 模型,以解决 Word2Vec 词嵌入出现的情感强度问题,两者相互补充,以提高分类结果的准确率;设计一种融合 Word2Vec 词嵌入多核卷积神经网络的音乐歌词多情感分类方法,对音乐歌词进行包括欢快、伤感、宣泄、治愈在内的 4 种情感分类,以实现音乐多情感分类. 研究成果可为用户提供更加个性化和精准的音乐推荐,增强音乐欣赏的体验,为情感分析领域提供更加丰富的数据支持和算法改进方法。

1　多情感分类模型

利用歌词进行音乐分类是自然语言处理的重要分支,前人研究中常利用歌词挖掘和其他相关技术,将歌词以观点、态度或者其他不同类别作为指标进行情感分类[15-17]。

在提取音乐特征方面,将单词表示向量(称为词嵌入)作为自然语言处理的特征学习技术引起学者们的广泛关注[18],词嵌入作为神经概率语言模型的衍生产品,单词嵌入隐含了单词在向量各个维度上的语义和句法信息,很好地表达了语义相似度[19]。近年来,基于词嵌入的向量技术在自然语言处理中发挥着重要的作用。Garg 等[20]对词嵌入进行了研究,证实了 Word2Vec 嵌入技术的优越性。

黄仁等[21]基于 Word2Vec 建立情感词典对互联网商品评论进行情感研究,取得了一定的成果,但是所设计的情感词典未包含词语的情感强度,也没有考虑情感词本身对分类的影响程度,在一定程度上影响了 Word2Vec 输出特征要素的性能。由此可见,选择合适的词嵌入技术后,分类器的选择也至关重要。

CNN 作为深度学习的主流算法,应用于各大领域,选用 CNN 并将其进行改进,设置 3 层不同高度的卷积核对音乐歌词进行多情感分类,应用于音乐歌词作为音乐多情感分类模型[22]。情感分类包括:多标签或二标签分类问题,本次研究使用多标签分类。多情感分类方法结构图如图 1 所示,每个阶段具体内容如下。

(1)数据预处理。数据预处理包含收集数据、统计歌曲长度以及歌曲长度出现的频数在内的相关工作,并将它们分别进行可视化分析。

(2)特征提取。采用 Word2Vec 词嵌入特征提取技术,能够将歌词内容进行处理,添加一种取出停用词以及过滤低频词汇的功能,然后计算相似语义度,选取最优特征要素,最后将其情感要素转化为向量输入分类器中[23],从而能够保证歌词词语特征求取的高效性。

(3)分类器设计。利用多核 CNN 网络作为音乐多情感分类器模型,在设置利用 Word2Vec 技术提取音乐歌词内容之后,又在原来单层 CNN 的基础上,添加不同尺寸的卷积核(高度分别为 3×100、4×100、5×100)提取出不同种类的特征图。而后利用池化层提取出最优局部特征,进而完成情感分类[24],提高多情感分类的准确率。

(4)性能评估。性能评估的评价指标包括准确率 p_{re} 和召回率 r_{ec}[25],其定义分别见式(1)、式(2),其中,准确率用于衡量分类器准确性,召回率用于衡量分类器是否能找全该类样本。使用测度 F_1 均衡这两方面,定义见式(3)。

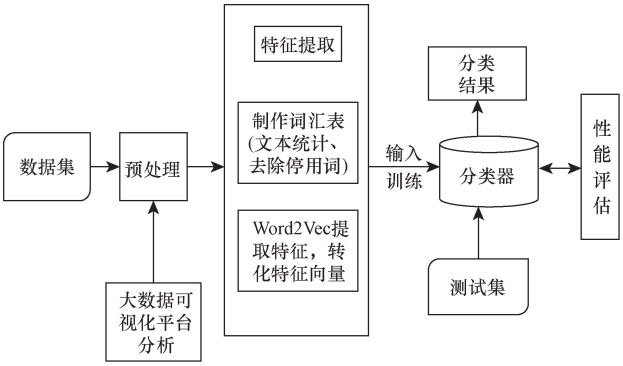


图 1 多情感分类方法结构图

Fig. 1 Structure of the multi-sentiment classification method

$$p_{re} = \frac{l_{yc}}{l_{yo}} \tag{1}$$

$$r_{ec} = \frac{l_{yc}}{m_{ul}} \tag{2}$$

$$F_1 = \frac{2p_{re}r_{ec}}{p_{re} + r_{ec}} \tag{3}$$

式中:l_{yc} 为实际正确的分类样本;l_{yo} 为预测正确的样本;m_{ul} 为全部正确的分类样本。

2 歌词多情感分类

音乐歌词包含丰富的情感内容,根据自然语言技术,首先要对歌词进行预处理,预处理包括数据预处理、文本特征表示。数据预处理主要是根据音乐歌词数据集的特点确定多情感分类输出的情感类型以及实现歌词文本序列长度可视化。文本特征表示分为去除停用词和构建情感词典。

2.1 数据预处理

本地音乐歌词数据来源于 Chinese Lyrics 中文歌词数据库,选取其中 8 000 首音乐,对比各音乐播放器的情感分类歌单,最终确定欢快、伤感、宣泄和治愈四类情感,其中每个类别对应的数据集数为欢快占据 2 400 条、伤感 2 800 条、治愈类型 1 200 条、宣泄类别 1 600 条。每首歌曲都伴有一段与其情感类别相契合的歌词,鉴于歌词本身蕴含的主观性特点,采用数据集历经多位独立个体的审慎确认,以确保其准确性。各类情感歌词在数据集中的分布占比,如图 2 所示。

将选取的歌词数据集导入大数据 AI 实验室自主研发的可视化平台中,对音乐歌词数据进行可视化分析[26],结果如图 3 所示。可以看出,大多数的歌词长度在 320 字以内。从歌词出现频数的累计分布函数图(图 4)可以看出,在数据中 90% 的分位点对应的歌词长度为 323 字。由此可知,本次研究中读取的音乐歌词长度为 320 字。

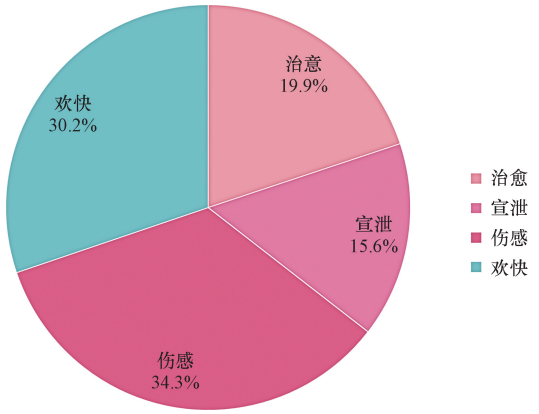


图 2 多情感分类数据占比图

Fig. 2 Proportion chart of sentiment text classification data

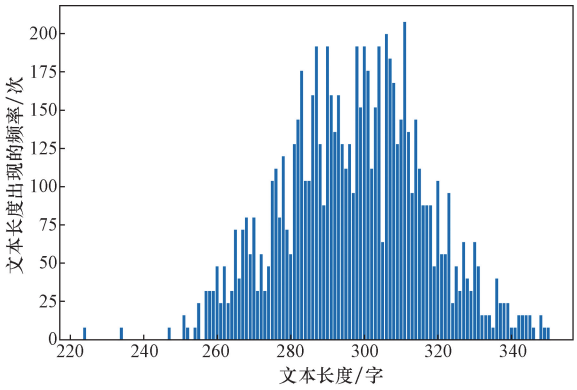


图 3　歌词长度频数统计图

Fig. 3　Text length frequency statistics chart

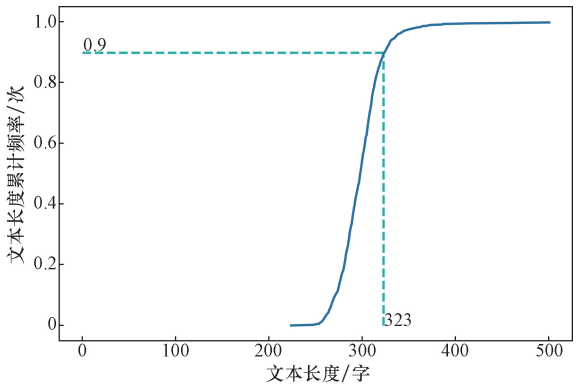


图 4　歌词频数累计分布函数图

Fig. 4　Cumulative distribution function of text frequency

2.2　文本特征提取

2.2.1　去除停用词

众所周知,在音乐歌词中有大量的重复片段,一些歌词使用频率很高但是意义不大,因此在歌词预处理阶段,针对一些副词、连词、数词在歌词中无实际意义的字(如“了”“会”“要”“与”等),在制作词汇表时将其去除,可降低对预处理的影响,同时提高自然语言处理的速度。因此,在停用词词汇表中设置去除音乐歌词中 25 个频繁停用词,如图 5 所示。

```
#去除停用词
stop_words = ['的', '了', '在', '是', '我', '有', '和', '就', '都', '一',
              '个', '上', '也', '到', '要', '去', '你', '会', '着', '这',
              '呀', '得', '成', '歌', '她', '样', '他', '舞', '地', '与']
```

图 5　去除停用词

Fig. 5　Remove stop words

2.2.2　特征提取

Word2Vec 是一种高效创建单词嵌入的技术,将各个词使用一个定长的向量来表示,并采用两层神经网络。Word2Vec 技术主要包括:连续词袋(CBOW)模型和跳字(Skip-gram)模型,这两个模型可以提供高质量的单词嵌入向量[27]。其中,CBOW 预测给定上下文的单词,Skip-gram 预测给定单词的上下文[28],它的本质特征是将一个词标识成一个向量。因此 Word2Vec 词向量可以较好地表达不同词之间的相似和类比关系,采用 Word2Vec 预训练词向量也会考虑词语之间顺序及上下文之间的关系,考虑到情感词本身的分布表示,它能够将语义相近的词语映射到相近的向量空间中,从而实现词语之间的语义关联。相较于现下流行的 BERT(bidirectional encoder representation from transformers)模型,Word2Vec 词嵌入更简单、高效。因此,利用 Word2Vec 技术构建词典。

本次研究中,Word2Vec 的输入是歌词语料库,输出是反映该语料库中词语的一组向量(即音乐歌词情感特征向量)。Word2Vec 技术计算语义相似度,筛选相似度足够的词向量汇总成新的情感词典[29]。Word2Vec 能够将歌词转化为深度神经网络容易理解的数字形式[30],适用于多核 CNN 模型的输入方式,因此选择 Word2Vec 技术作为前期数据预处理模型,模型图如图 6 所示。

3　多核 CNN 的音乐多情感分类模型

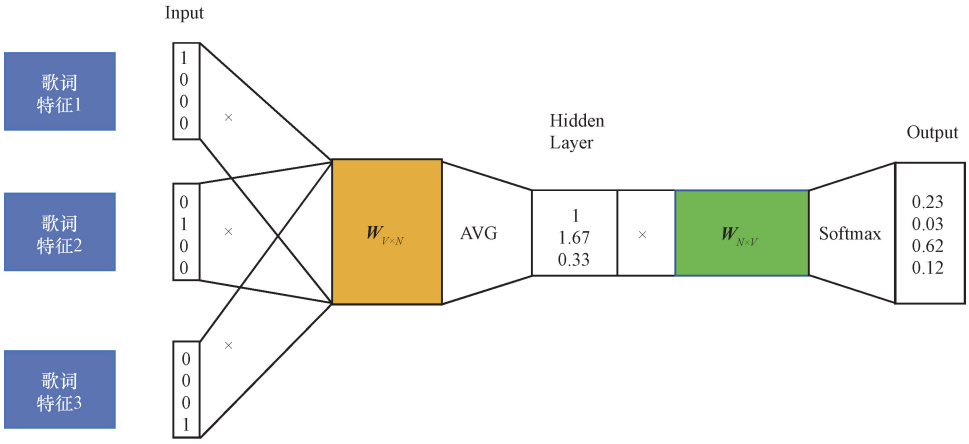
CNN 具有权重共享和稀疏连通性的特点,一般由卷积层、池化层和全连接层组成。普通单层 CNN 对于音乐歌词情感分类而言分类性能不高,为了提高音乐多情感分类的准确性,提出一种在音乐歌词情感分类研究基础上,新的卷积神经网络音乐歌词多情感分类模型(称为多核 CNN)。该模型在单层 CNN 的基础上,首先将上述 Word2Vec 特征提取的词向量作为输入,然后将卷积层部分使用不同尺寸的卷积核,更好地选取音乐歌词中的情感特征,最后得到多情感分类结果的高准确率。本文算法模型如图 7 所示。

图 7 中第 1 层为输入层,在模型的输入处理中,在本次实验研究中读取的音乐歌词长度为 320,所以 word embedding 设置为 100 维,在该层上采用了 Word2Vec 技术特征提取情感方法生出 f 维的矩阵,然后将生成的矩阵向量作为输入传至卷积层。歌词采用 Word2Vec 技术特征提取后,得到 f 维的词向量 $\boldsymbol{R}^f$,假设它是的第 i 个字的向量表示,因此长度为 n 的句子可表示为

$$\boldsymbol{x}_{i:n} = \boldsymbol{x}_1 \oplus \boldsymbol{x}_2 \oplus \cdots \oplus \boldsymbol{x}_n \tag{4}$$

式(4)中:$\oplus$ 为连接符,将长度为 n 的句子中的每个向量全部连接起来。

第 2 层为卷积层,卷积层用于使用卷积核来提取特征映射。卷积核可以看作是矩阵。为了避免层数过多而导致特征过于抽象的问题,在单层 CNN



Input 为输入;AVG 为扩展;Hidden Layer 为隐藏层;Output 为输出;Softmax 为全连接;Target 为目标函数;$\boldsymbol{W}_{V\times N}$为 $V\times N$ 的矩阵;$\boldsymbol{W}_{N\times V}$为 $N\times V$ 的矩阵

图 6 Word2Vec 歌词嵌入模型图

Fig. 6 Word2Vec lyrics embedding model diagram

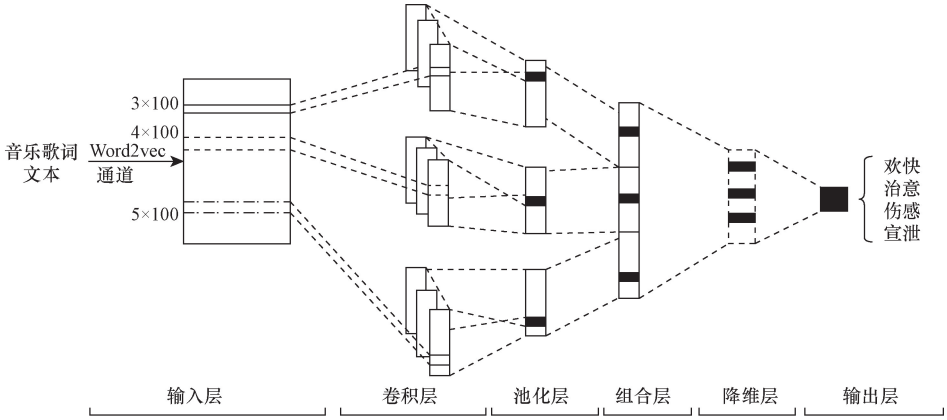


图 7 多核 CNN 模型

Fig. 7 Multi-core CNN model

的基础上设计了不同尺寸的卷积核。卷积的详细操作如下。

设计 3 种不同高度的卷积核,3 种卷积核的高度设为 h,分别为 3×100、4×100、5×100,每种尺寸的卷积核有 r 个,值设置 128,权值矩阵 $\boldsymbol{w}_1$ 是一个实数矩阵,对 h 个字的内容块进行一个特征提取的操作。由 $\boldsymbol{x}_{i:i+h-1}$ 提取到的特征 o_i 可表示为

$$o_i = f(\boldsymbol{w}_1\boldsymbol{x}_{i:i+h-1} + \boldsymbol{b}_1) \tag{5}$$

式(5)中:$\boldsymbol{x}_{i:i+h-1}$ 为输入的第 i 个到第 $i+h-1$ 个窗口内的词向量矩阵;$f(\cdot)$ 为一个非线性激活函数;$\boldsymbol{b}_1$ 为偏置项,属于实数。

整个卷积操作完成后会得到特征图,可表示为

$$\boldsymbol{o} = [o_1, o_2, \cdots, o_{n-h+1}] \tag{6}$$

式(6)中:$\boldsymbol{o}$ 为大小为 $n-h+1$ 的实数向量。

第 3 层为池化层,池化层保留了卷积以后的数据特征。采用 1-Max 池化方法,从卷积后的每个窗口产生的特征向量中选择一个最优情感特征向量,最后将其拼接起来,降低了数据所在的高维数据空间的维数,避免了维数灾难,同时提取到更为重要的音乐情感特征要素。

池化层的操作可表示为

$$\hat{\boldsymbol{o}} = \max\{\boldsymbol{o}\} \tag{7}$$

式(7)中:$\hat{\boldsymbol{o}}$ 为 $\boldsymbol{o}$ 的最大值。

本次研究中进行 3 种不同尺寸的卷积核操作,因此将池化后的数据特征拼接到一起,最终形成特征向量 $\boldsymbol{\theta}$,$\boldsymbol{\theta}\in\mathbf{R}^{1+3r}$,可表示为

$$\boldsymbol{\theta} = \hat{\boldsymbol{o}}^3 + \hat{\boldsymbol{o}}^4 + \hat{\boldsymbol{o}}^5 \tag{8}$$

式(8)中:$\hat{\boldsymbol{o}}^3$、$\hat{\boldsymbol{o}}^4$、$\hat{\boldsymbol{o}}^5$ 分别为卷积核高度为 3、4、5 最大池化后的输出情感向量。

下一层为降维层(又称为隐藏层),经过非线性降维之后,上述特征向量 $\boldsymbol{\theta}$ 变成特征向量 $\boldsymbol{\theta}\in\mathbf{R}^{1\times m}$,其中 m 为降维层中神经元节点数。

最后一层为 Softmax 输出层,又称为全连接层,将最后的特征结果输出至该层,通过 Softmax 层输

出 4 种情感类别的概率分布,得到的预测类别的标签值,对应的表达式为

$$c_i = \mathrm{Softmax}(\boldsymbol{w}_2 n + \boldsymbol{b}_2) \tag{9}$$

式(9)中:$\boldsymbol{w}_2$ 为一个实数矩阵;n 为类别数;$\boldsymbol{b}_2$ 为偏置项。

4　可视化实验分析

4.1　实验环境搭建

对音乐歌词多情感分类在相同的实验环境下进行不同类型的对比实验,体现不同网络对音乐多情感分类的性能。本次实验环境如表 1 所示。

表 1　环境搭建

Table 2　Environmental setup table

实验环境	环境配置
操作系统	Windows 11 系统
处理器	Intel(R) Core(TM) i5-10210U CPU@1.60 GHz、2.11 GHz
机带 RAM/GB	8.00
编程语言	Python 3.6.5
编程软件	Jupyter notebook
机器学习库	Tensorflow-1.14、sklearn
Word embedding 训练软件	Word2Vec

4.2　可视化实验结果与分析

使用测试集文本对本文方法进行测试,随机选取欢快、伤感、治愈、宣泄测试集各 50 个文本进行情感分类后的结果展示,其分类结果如图 8 所示,结果表明,欢快的准确率最高,而其他三类分类准确相较欢快来说略低,且大多数错误的分类歌曲都被分到喜悦类别里。

为了更好体现出此方法对于多情感分类的性能,依据中外学者对音乐歌词多情感分类的研究又设置多个对比实验,以 SVM、KNN 为代表的机器学习模型,和深度学习中的 RNN、经典 CNN(包括单层 CNN、多层 CNN)和多核 CNN。不同对比实验的准确率如表 2 所示。

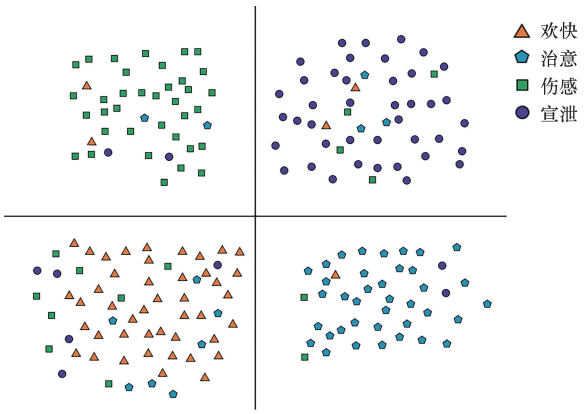


图 8　分类结果图

Fig. 8　Classification results chart

表 2　不同网络对比实验结果

Table 2　Comparison experiment results of different networks

算法模型	p_{re}	r_{ec}	F_1
KNN	0.715 6	0.713 4	0.709 8
SVM	0.762 4	0.768 6	0.767 6
RNN	0.813 2	0.811 3	0.808 6
单层 CNN	0.874 0	0.875 8	0.876 3
多层 CNN	0.838 3	0.833 4	0.832 5
多核 CNN	0.942 6	0.943 6	0.943 2

从表 2 中可以看出,单层 CNN 的分类结果准确率为 86.4%,然而在单层 CNN 的基础上添加卷积神经网络层数获得的多情感分类结果为 83.83%,是比单层 CNN 要低的,说明多层 CNN 未达到预期效果,因此多层多核 CNN 没有设置相应的对比实验。所提出的融合 Word2Vec 词嵌入多核卷积神经网络的音乐歌词多情感分类方法在分类中准确率达到 94.26%,在这几类实验中准确率最高,表明采用上述方法对于音乐歌词多情感分类有着不错的性能。

为了进一步研究不同情感分类模型,在上述分类中选出 3 种最优的模型进行可视化分析,分别对多核 CNN、单层 CNN、RNN 模型进行对比,测试模型迭代 100 轮次输出一组测试精度值和损失值。随着每轮迭代次数的变化,三类模型的输出结果如图 9 所示。

由图 9 可知,三类模型的精度值随着迭代次数的不断增加而趋于平缓,最终靠向收敛状态,而损失值也在不断下降最终在小区间内稳定波动。其中,多核 CNN 模型整体的精度值曲线高于其他两类模型,说明多核 CNN 在普通 CNN 基础上改进之后取得了较好的情感分类效果。多核 CNN 的损失值浮动最大,但整体损失值低于其他两类模型,并且明显好于 RNN 模型。由此可证明本文方法在音乐歌词情感分类的有效性。

5　结论

提出一种融合 Word2Vec 词嵌入多核卷积神经网络的音乐歌词多情感分类方法,该方法采用 4 种不同的音乐情感对音乐歌词进行分类。利用 Word2Vec 词嵌入技术,计算上下文中的语义相似度得到最优情感词,并将其转化为便于多核 CNN 分类器模型输入的情感向量,之后在单层卷积网络的基础上添加 3 种不同尺寸的卷积核的方式来提高音乐歌词多情感分类的准确率。得到以下结论。

(1)RNN、不同层数 CNN 和多核 CNN 的情感分类效果高于传统分类器 KNN、SVM,说明深度学习可

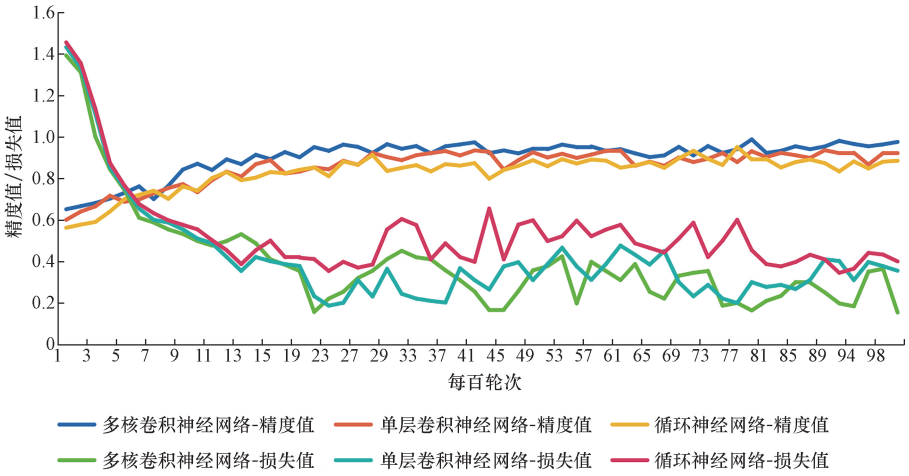


图 9 不同模型的测试精度和损失

Fig. 9 Test accuracy and loss for different models

以提取到得更多的情感分类特征,情感分类训练效率更高效。

(2)音乐歌词多情感分类的 CNN 模型学习到的情感分类特征更多,展示的性能更高效,比 RNN 更适用于音乐歌词。

(3)本文方法得到的音乐情感分类结果达到 94.26%,与传统 CNN 相比,分类精确率提高了 6.86%,取得了较好的效果。

参 考 文 献

[1] Peter D, Stefanie N, André B, Cornelia L, et al. Content-based mood classification for photos and music: a generic multi-modal classification framework and evaluation approach[C]//Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval. Quebec: ACM, 2008: 97-104.

[2] Zhou M, Liu D, Zheng Y, et al. A text sentiment classification model using double word embedding methods[J]. Multimedia Tools and Applications, 2020, 81(14): 18993-19012.

[3] 诸林云, 范菁, 曲金帅, 等. 基于 BERT 与多通道卷积神经网络的细粒度情感分类[J]. 科学技术与工程, 2023, 23(31): 14264-14270.
Zhu Linyun, Fan Jing, Qu Jinshuai, et al. Finegrained emolion classilication based on BERT and muli-channel convolutional neund networks[J]. Seienee Technology and Engineering, 2023, 23(33): 14264-14270.

[4] Trohidis K, Tsoumakas G, Kalliris G, et al. Multi-label classification of music by emotion[J]. EURASIP Journal on Audio, Speech, and Music Processing, 2011, 11: 325-330.

[5] 王小凤, 耿国华, 李鹏, 等. 一个基于相关反馈的神经网络音乐情感分类器构建算法[J]. 西北大学学报(自然科学版), 2012, 42(1): 30-35.
Wang Xiaofeng, Di Guohua, Li Peng, et al. A music emotion classifier construction algorithm of neural network based on relevant feedback[J]. Journal of Northwest University(Natural Science Edition), 2012, 42(1): 30-35.

[6] Kim Y. Convolutional neural networks for sentence classification[J]. Computing Research Repository, 2014, 14: 1746-1751.

[7] 龚安, 丁明波, 窦菲. 基于 DBN 的多特征融合音乐情感分类方法[J]. 计算机系统应用, 2017, 26(9): 158-164.
Gong An, Ding Mingbo, Dou Fei. Music mood classification method based on deep belief network and multi-feature fusion[J]. Computer Systems & Application, 2017, 26(9): 158-164.

[8] 王洁, 朱贝贝. 面向中文歌词的音乐情感分类方法[J]. 计算机系统应用, 2019, 28(8): 24-29.
Wang Jie, Zhu Beibei. Classification of musical emotions oriented to Chinese lyrics[J]. Computer Systems & Applications, 2019, 28(8): 24-29.

[9] Palanisamy K, Singhania D, Yao A. Rethinking CNN models for audio classification[J]. arXiv Preprint, 2020: arXiv. 2007. 11154.

[10] Tong G. Music emotion classification method using improved deep belief network[J]. Mobile Information Systems, 2022, 2022: 1-7.

[11] Yang Q, Liu S, Gong T. Improve the application of reinforcement learning and multi-modal information in music sentiment analysis[J]. Expert Systems, 2023, 21(2): 1154-1160.

[12] 杨秀璋, 郭明镇, 候红涛, 等. 融合情感词典的改进 BiLSTM-CNN + Attention 情感分类算法[J]. 科学技术与工程, 2022, 22(20): 8761-8770.
Yang Xiuzhang, Guo Mingzhen, Hou Hongtao, et al. Improved BiLSTM-CNN + Attention sentiment classifiealion alorithm fused with sentiment dictionary[J]. Science Technology and Engineering, 2022, 22(20): 8761-8770.

[13] 张昱, 高凯龙, 苏仡琳, 等. 双通道多核卷积神经网络中文歌词情绪分类方法[J]. 内蒙古大学学报(自然科学版), 2021, 52(5): 508-513.
Zhang Yu, Gao Kailong, Su Yilin, et al. Chinese text sentiment classification method with dual-channel multi-core convolutional neural network[J]. Journal of Inner Mongolia University(Natural Science Edition), 2021, 52(5): 508-513.

[14] He J, Zhou L. Liu Z, et al. Digital empirical research of influencing factors of musical emotion classification based on pleasure-arousal musical emotion fuzzy model[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2020, 24(7): 872-881.

[15] 郑承宇, 王新, 王婷, 等. 基于 Stacking-Bert 集成学习的中文短文本分类算法[J]. 科学技术与工程, 2022, 22(10): 4033-4038.

Zheng Chengyu, Wang Xin, Wang Ting, et al. Chinese short text classification algorithm based on Stacking-Bert ensemble learning [J]. Science Technology and Engineering, 2022, 22(10): 4033-4038.

[16] 黄卫春, 邹瑶, 熊李艳, 等. 基于卷积神经网络的多通道特征表示文本分类模型[J]. 科学技术与工程, 2021, 21(16): 6764-6771.
Huang Weichun, Zou Yao, Xiong Liyan, et al. Multi-channel feature representation text classification model based on convolutional neural network[J]. Science Technology and Engineering, 2021, 21(16): 6764-6771.

[17] Mutinda J, Mwangi W, Okeyo G. Sentiment analysis of text reviews using Lexicon-enhanced Bert embedding(LeBERT) model with convolutional neural network[J]. Applied Sciences, 2023, 13(3): 1445-1445.

[18] Rahimi Z, Homayounpour M. TensSent: a tensor based sentimental word embedding method[J]. Applied Intelligence, 2021, 51(8): 6056-6071.

[19] Yu K, Zhang Y, Wang X. Topic model over short texts incorporating word embedding[C]//Proceedings of the 2018 2nd International Conference on Advances in Energy, Environment and Chemical Science. New York: AEECS, 2018: DOI: 10.2991/aeecs-18.2018.34.

[20] Garg S, Subrahmanyam V. Sentiment analysis: choosing the right word embedding for deep learning model[M]. Berlin: Springer International Publishing, 2022.

[21] 黄仁, 张卫. 基于 Word2Vec 的互联网商品评论情感倾向研究[J]. 计算机科学. 2016, 43(S1): 387-389.
Huang Ren, Zhang Wei. Study on sentiment analyzing of internet commodities review based on Word2Vec[J]. Computer Science, 2016, 43(S1): 387-389.

[22] 张浩然, 谢云熙, 张艳荣. 基于 TextCNN 的歌词情感分类系统[J]. 哈尔滨商业大学学报(自然科学版), 2022, 38(3): 285-292.
Zhang Haoran, Xie Yunxi, Zhang Yanrong. Text sentiment classification system based on TextCNN[J]. Journal of Harbin University of Commerce (Natural Sciences Edition), 2022, 38(3): 285-292.

[23] Baek J, Chung K. Multimedia recommendation using Word2Vec-based social relationship mining[J]. Multimedia Tools and Applications, 2020, 80(26/27): 34499-34515.

[24] 蔡慧苹, 王丽丹, 段书凯. 基于 word embedding 和 CNN 的情感分类模型[J]. 计算机应用研究, 2016, 33(10): 2902-2905, 2909.
Cai Huiping, Wang Lidan, Duan Shukai. Sentiment classification model based on word embedding and CNN[J]. Application Research of Computers, 2016, 33(10): 2902-2905, 2909.

[25] 丁森华, 邵佳慧, 李春艳, 等. 歌词情感分析方法对比研究[J]. 广播电视信息, 2020, 2020(4): 92-96.
Ding Senhua, Shao Jiahui, Li Chunyan, et al. Comparative study on text sentiment analysis methods[J]. Radio and Television Information, 2020, 2020(4): 92-96.

[26] 张昱, 刘开峰, 张全新, 等. 基于组合-卷积神经网络的中文新闻歌词分类[J]. 电子学报, 2021, 49(6): 1059-1067.
Zhang Yu, Liu Kaifeng, Zhang Quanxin, et al. A combined-convolutional neural network for Chinese news text classification[J]. Chinese Journal of Electronics, 2021, 49(6): 1059-1067.

[27] Rezaeinia S, Rahmain R, Ghodsi A, et al. Sentiment analysis based on improved pre-trained word embeddings[J]. Expert Systems with Applications, 2019, 117: 139-147.

[28] Gu J, Li G, Jung J, et al. Contextual Word2Vec model for understanding chinese out of vocabularies on online social media[J]. International Journal on Semantic Web and Information Systems, 2022, 18(1): 1-14.

[29] 杨帆, 余华平. 中文歌词文本的情感分类综述[J]. 电脑知识与技术, 2021, 17(3): 207-208, 215.
Yang Fan, Yu Huaping. A review of the emotional classification of Chinese lyrics[J]. Computer Knowledge and Technology, 2021, 17(3): 207-208, 215.

[30] Priyadarshini I, Cotton C. A novel LSTM-CNN grid search-based deep neural network for sentiment analysis[J]. The Journal of Supercomputing. 2021, 77(12): 13911-13932.