基于 ALBERT-CRNN 的弹幕文本情感分析

曾　诚[1,2,3]，　温超东[1]，　孙瑜敏[1]，　潘　列[1]，　何　鹏[1,2,3]

(1. 湖北大学 计算机与信息工程学院　湖北 武汉 430062；
2. 湖北省软件工程工程技术研究中心　湖北 武汉 430062；
3. 智慧政务与人工智能应用湖北省工程研究中心　湖北 武汉 430062)

摘要：提出一种结合 ALBERT 预训练语言模型与卷积循环神经网络(convolutional recurrent neural network，CRNN)的弹幕文本情感分析模型 ALBERT-CRNN。首先使用 ALBERT 预训练语言模型获取弹幕文本的动态特征表示，使得句子中同一个词在不同上下文语境中具有不同的词向量表达；然后利用 CRNN 对特征进行训练，充分考虑了文本中的局部特征信息和上下文语义关联；最后通过 Softmax 函数得出弹幕文本的情感极性。在哔哩哔哩、爱奇艺和腾讯视频三个视频平台的弹幕文本数据集上进行实验，结果表明，ALBERT-CRNN 模型在三个数据集上的准确率分别达到 94.3%、93.5%和 94.8%，相比一些传统模型具有更好的效果。

关键词：弹幕文本；情感分析；词向量；预训练语言模型；卷积循环神经网络

中图分类号：TP18　　**文献标志码**：A　　**文章编号**：1671-6841(2021)03-0001-08

DOI：10.13705/j.issn.1671-6841.2020359

0　引言

随着数字媒体技术的快速发展，弹幕成为人们表达观点的一个重要途径，深受年轻用户的欢迎，并在哔哩哔哩、爱奇艺和腾讯视频等多个视频平台逐渐流行。弹幕评论在发送后直接同步显示在视频播放过程中，并在屏幕上从右到左如子弹般缓缓飞过，故网友称之为弹幕。相比普通文本，弹幕文本有着其独特的风格，不仅包含较多的网络词语和字符表情，且存在大量“同词不同义”的情况，给情感分析带来较大挑战。如“喜欢”这个词，在“我很喜欢 up 主的风格”这个句子中表示一种正向的情感，但在“男主总是喜欢搬弄是非”这个句子中表示一种倾向，不带任何感情色彩。传统的弹幕文本情感分析方法大多利用情感词典来判断弹幕文本的情感极性。文献[1]采用情感词典对弹幕文本中的词汇进行情感分析，得到词汇的情感强度并将其累加，从而得出弹幕文本整体的情感极性。文献[2]针对弹幕文本口语化的特点，建立了网络弹幕常用词词典，能够更好地识别弹幕文本中出现的网络词汇，并通过该词典计算弹幕文本的情感值，最终利用情感值对弹幕文本进行分类。近年来，随着深度学习的发展，越来越多的研究者将深度学习应用在弹幕文本分析任务中。文献[3]提出一种基于注意力机制的长短期记忆网络情感分析模型，用于帮助用户准确地获取弹幕文本中所包含的情感信息。文献[4]使用双通道卷积神经网络对文本进行情感分析，其中一个通道为字向量，另一个通道为词向量，解决了单通道卷积神经网络视角单一以及不能充分学习到文本特征信息的问题。文献[5]则提出一种基于奇异值分解算法的卷积神经网络模型，使用奇异值分解的方法代替传统卷积神经网络模型中的池化层进行特征提取和降维，有效提升了弹幕文本情感分析的效果。

虽然以上研究取得了不错的效果，但是由于弹幕文本中存在着大量“同词不同义”的情况，以上方法在进行特征提取时无法区分句子中同一个词在不同上下文语境中的不同含义，且在训练过程中不能兼顾文本中的局部特征信息和上下文语义关联，导致其分类准确率相对较低。因此，本文结合 ALBERT 预训练语言模型和卷积循环神经网络(CRNN)方法来分析弹幕文本的情感极性，提出一种弹幕文本情感分析模型 ALBERT-CRNN。利用 ALBERT 预训练语言模型获取弹幕文本的动态特征表示，充分利用了句子中词的上

收稿日期：2020-11-03
基金项目：国家自然科学基金项目(61977021，61902114)；湖北省 2019 年技术创新专项(2019ACA144)。
作者简介：曾诚(1976—)，男，教授，主要从事人工智能与行业软件研发研究，E-mail：zc@hubu.edu.cn。

下文信息,使得句子中同一个词在不同上下文语境中具有不同的词向量表达;使用 CRNN 对特征进行训练,充分考虑了文本中的局部特征信息和上下文语义关联,进一步提高了模型在弹幕文本情感分析任务中的准确率。

1 相关工作

深度学习技术在情感分析中应用的前提是解决词映射问题,即将文本转化为机器可以识别的数字,常用的方法是针对文本训练词向量。Mikolov 等[6-7]提出 Word2Vec 模型,其本质为一种神经网络概率语言模型,包括 CBOW 和 Skip-Gram 两种模型,但通过该模型训练出的词向量为静态词向量,舍弃了文本中大量词语的位置信息,不能表示出文本的完整语义。Devlin 等[8]在 ELMo[9]和 GPT[10]的基础上提出了 BERT 预训练语言模型,该模型通过使用双向 Transformer[11]编码器对语料库进行训练得到文本的双向编码表示,且训练出的词向量为动态词向量,使得句子中同一个词在不同上下文语境中具有不同的词向量表达。

增大 BERT 预训练模型的规模能提升下游任务的效果,但受计算资源的限制,所需的训练时间较长,且进一步提升模型的规模将导致显存或内存不足。为此,Lan 等[12]提出 ALBERT 模型,该模型是基于 BERT 模型的一种轻量级预训练语言模型,采用双向 Transformer 获取文本的特征表示,且大大减少了模型中的参数,并在多项自然语言处理(NLP)任务中取得了最佳效果。为准确判断弹幕文本的情感倾向,特征分类算法的选择也尤为重要。Kim[13]提出文本卷积神经网络模型(TextCNN),使用不同尺寸的卷积核对文本局部特征进行训练,实现了句子级别的分类任务,并取得了较好的分类效果。Socher 等[14]将循环神经网络(RNN)应用到 NLP 任务中,在进行文本特征训练时能够较好地利用上下文信息。然而,传统 RNN 存在梯度爆炸和消失问题,处理长序列文本的效果并不理想。Hochreiter 等[15]提出长短期记忆网络(LSTM),解决了传统 RNN 的梯度爆炸和消失问题。之后,Dey 等[16]提出了门控循环单元(GRU),在保持 LSTM 效果的同时使得模型结构更加简单。

卷积神经网络(CNN)虽然能够高效地利用文本的局部特征信息,但无法获取上下文信息;双向门控循环单元(BiGRU)虽然能够有效地获取文本的上下文信息,但由于自身循环递归的特性,网络结构复杂度较高,时间代价较大。文献[17]结合 CNN 和 BiGRU 的网络结构优势,提出基于注意力机制的 CRNN 文本分类算法,既能利用 CNN 训练局部特征的优势,又能利用 BiGRU 获取上下文语义信息的优势,提升了文本分类的准确率。然而,由于其在提取文本特征时采用的是传统的 Word2Vec 模型,导致其舍弃了大量词语的位置信息,无法表示出文本的完整语义。综合以上研究,本文将 ALBERT 预训练语言模型与 CRNN 相结合,提出了一种基于 ALBERT-CRNN 的弹幕文本情感分析模型,并通过与其他模型的对比实验证明了其在弹幕文本情感分析中的有效性。

2 基于 ALBERT-CRNN 的弹幕文本情感分析模型

本文提出的 ALBERT-CRNN 弹幕文本情感分析方法主要包括以下步骤。1) 对弹幕文本数据进行清洗和预处理,筛选出具有情感极性的弹幕文本数据,并打上相应的标签。2) 使用 ALBERT 预训练语言模型获取弹幕文本的动态特征表示。3) 使用结合 CNN 与 BiGRU 的神经网络模型 CRNN 对文本特征进行训练,获取每条弹幕文本的深层语义特征。4) 利用 Softmax 函数对文本深层语义特征进行分类,最终得出每条弹幕文本的情感极性。

2.1 ALBERT 预训练语言模型

2.1.1 ALBERT 模型结构 ALBERT 预训练语言模型采用双向 Transformer 获取文本的特征表示,其模型结构如图 1 所示。其中,$\boldsymbol{E}_1$, $\boldsymbol{E}_2$,…,$\boldsymbol{E}_N$ 表示序列中的每一个字符,经过多层双向 Transformer 编码器的训练,最终得到文本的特征向量表示 $\boldsymbol{T}_1$,$\boldsymbol{T}_2$,…,$\boldsymbol{T}_N$。Transformer 的模型结构为 Encoder-Decoder[18-19],ALBERT 采用的是其 Encoder 部分,该部分由多个相同的基本层组成。其中,每个基本层包含两个子网络层:第一个为多头自注意力机制层;第二个为普通前馈网络层。

2.1.2 ALBERT 模型对 BERT 模型的改进 为减少 BERT 模型的参数和增强模型的语义理解能力,

ALBERT 模型在 BERT 模型的基础上主要进行了两点改进。首先,ALBERT 模型通过嵌入层参数因式分解和跨层参数共享方法有效减少了 BERT 模型中的参数,大大降低了训练时的内存花销,并有效提升了模型的训练速度。其次,为弥补 Yang 等[20]提出的 BERT 模型中 NSP 任务存在的缺点,ALBERT 模型通过使用 SOP 任务代替 BERT 模型中的 NSP 任务,提升了多句子输入的下游任务的效果。

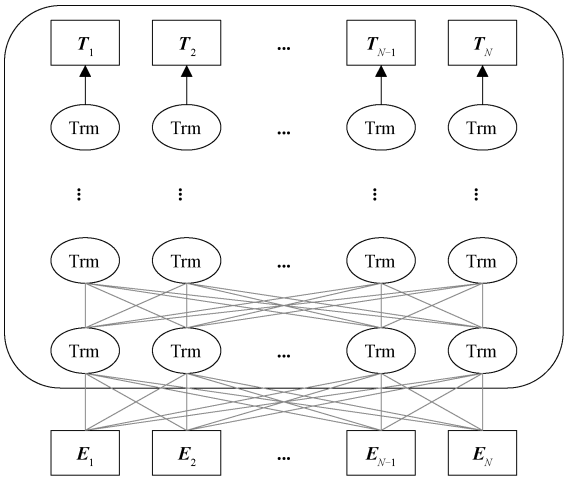

图 1　ALBERT 模型结构

Figure 1　ALBERT model structure

2.2　CNN 层

CNN 是一种前馈神经网络,由卷积层和池化层组成。在卷积层中,通过与文本特征表示矩阵 $\boldsymbol{S}$ 进行卷积得到新的特征 $\boldsymbol{c}_i$,其计算公式为

$$\boldsymbol{c}_i = f(\boldsymbol{w} \otimes \boldsymbol{S}_{i:i+m-1} + \boldsymbol{b}), \tag{1}$$

其中:i 表示第 i 个特征值;m 表示卷积计算中滑动窗口的大小;$\boldsymbol{S}_{i:i+m-1}$ 表示矩阵 $\boldsymbol{S}$ 中第 i 行至第 $i+m-1$ 行的文本特征矩阵;$\boldsymbol{w}$ 为卷积核;⊗表示卷积计算;f 为非线性激活函数;$\boldsymbol{b}$ 为偏置值。将滑动窗口应用于各局部文本特征矩阵 $\boldsymbol{S}_{1:m}, \boldsymbol{S}_{2:m+1}, \cdots, \boldsymbol{S}_{n-m+1:n}$,最终得到特征向量 $\boldsymbol{C}$,

$$\boldsymbol{C} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_{n-m+1})。 \tag{2}$$

另外,分别采用尺寸为 3、4、5 的卷积核对特征图进行卷积,对每个卷积核重复上述过程,得到各自的特征向量。在池化层中,通过最大池化方法保留权重最大的特征值,并舍弃其他特征值,计算公式为

$$\boldsymbol{p}_j = \max(\boldsymbol{c}_j), \tag{3}$$

其中:$\boldsymbol{p}_j$ 表示特征图中第 j 个池化区域内最大的特征值。

2.3　BiGRU 层

GRU 是 LSTM 模型的一个变体,其模型结构如图 2 所示。LSTM 模型包含三个门计算,即输入门、输出门和遗忘门。GRU 模型在 LSTM 模型的基础上进行了简化,只由 $\boldsymbol{z}_t$ 和 $\boldsymbol{r}_t$ 两个门控单元组成。其中:$\boldsymbol{z}_t$ 表示更新门,用于控制前一时刻的状态信息被带入到当前状态中的程度;$\boldsymbol{r}_t$ 表示重置门,用于控制忽略前一时刻的状态信息的程度。

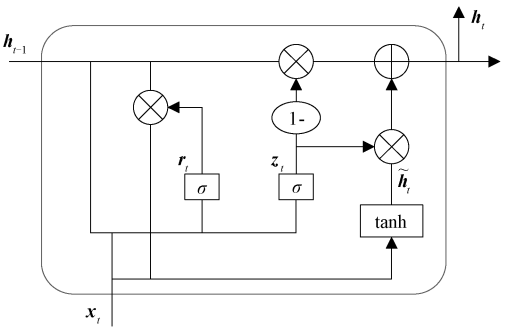

图 2　GRU 模型结构

Figure 2　GRU model structure

基于以上 GRU 的模型结构,可以得出 GRU 的前向传播计算公式为

$$\boldsymbol{z}_t = \sigma(\boldsymbol{W}_{zx}\boldsymbol{x}_t + \boldsymbol{W}_{zh}\boldsymbol{h}_{t-1} + \boldsymbol{b}_z), \tag{4}$$

$$\boldsymbol{r}_t = \sigma(\boldsymbol{W}_{rx}\boldsymbol{x}_t + \boldsymbol{W}_{rh}\boldsymbol{h}_{t-1} + \boldsymbol{b}_r), \tag{5}$$

$$\tilde{\boldsymbol{h}}_t = \tanh(\boldsymbol{W}_{\tilde{h}x}\boldsymbol{x}_t + \boldsymbol{W}_{\tilde{h}h}(\boldsymbol{r}_t \odot \boldsymbol{h}_{t-1}) + \boldsymbol{b}_{\tilde{h}}), \tag{6}$$

$$\boldsymbol{h}_t = (1 - \boldsymbol{z}_t) \odot \boldsymbol{h}_{t-1} + \boldsymbol{z}_t \odot \tilde{\boldsymbol{h}}_t, \tag{7}$$

其中:σ 表示 sigmoid 激活函数;$\boldsymbol{x}_t$ 表示当前时刻的输入,在文本分类中表示第 t 个词的词向量;$\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{h}_t$ 分别表示前一时刻隐藏层状态和当前时刻隐藏层状态,$\tilde{\boldsymbol{h}}_t$ 表示当前时刻新的记忆;⊙表示向量的点乘。

2.4　ALBERT-CRNN 模型

ALBERT-CRNN 模型结构如图 3 所示,主要由以下 6 个部分组成:输入层、ALBERT 层、CRNN 层(包含 CNN 层和 BiGRU 层)、全连接层、Softmax 层和输出层。该模型的工作流程如下。

步骤 1　利用输入层将弹幕文本数据输入到模型的 ALBERT 层中,输入的弹幕文本数据 $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N)$,其中 $\boldsymbol{X}_i$ 表示该条弹幕文本中的第 i 个词。

步骤 2　在 ALBERT 层对输入的文本数据进行序列化,将文本数据 $\boldsymbol{X}$ 中的每个词转化为其在字典中所对应的编号。序列化后的文本数据 $\boldsymbol{E} = (\boldsymbol{E}_1, \boldsymbol{E}_2, \cdots, \boldsymbol{E}_N)$,其中 $\boldsymbol{E}_i$ 表示文本中第 i 个词的序列化字符。利用

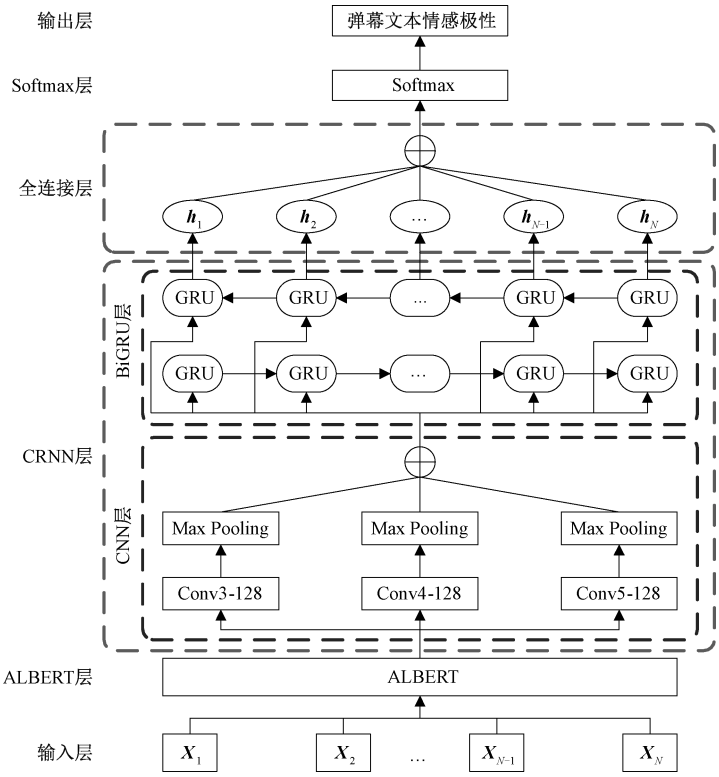

图 3 ALBERT-CRNN 模型结构

Figure 3 ALBERT-CRNN model structure

多层双向 Transformer 编码器对序列化后的弹幕文本进行训练,得到弹幕文本的动态特征表示。文本特征表示 $\boldsymbol{T}=(\boldsymbol{T}_1,\boldsymbol{T}_2,\cdots,\boldsymbol{T}_N)$,其中 $\boldsymbol{T}_i$ 表示文本中第 i 个词的特征向量。在使用 Transformer 编码器获取弹幕文本特征时,计算当前句子中每个词与其他词之间的相互关系,然后利用这些相互关系去调整每个词的权重,从而获得句子中每个词的新的表达。通过此方式训练出的文本特征表示 $\boldsymbol{T}$ 充分利用了句子中词的上下文信息,使得句子中同一个词在不同上下文语境中具有不同的词向量表达,较好地区分了同一个词在不同上下文语境中的不同含义。

步骤 3 将文本特征表示 $\boldsymbol{T}$ 输入到 CNN 层中,分别采用尺寸为 3、4、5 的卷积核对文本特征进行训练,经池化层降维后分别得到三个文本向量 $\boldsymbol{F}_{c_1}$、$\boldsymbol{F}_{c_2}$ 和 $\boldsymbol{F}_{c_3}$,将三者进行叠加得到向量 $\boldsymbol{F}_c$。为保证池化后的三个文本向量能够相互叠加,在池化层采用全填充的方式使得池化后的三个文本向量形状相同。

步骤 4 将 CNN 层的输出 $\boldsymbol{F}_c$ 分别传给 BiGRU 层的前向 GRU 层和后向 GRU 层,经过多个 GRU 隐藏单元的训练,最终得到两个文本向量表示,分别记作 $\boldsymbol{F}_{g_0}$ 和 $\boldsymbol{F}_{g_1}$。

步骤 5 将 $\boldsymbol{F}_{g_0}$ 和 $\boldsymbol{F}_{g_1}$ 进行叠加得到向量 $\boldsymbol{F}_g$,$\boldsymbol{F}_g$ 的维度为 $2h$,其中 h 为 GRU 隐藏单元个数。通过全连接层对 $\boldsymbol{F}_g$ 进行两次全连接,全连接层的输出维度为 s,s 表示情感标签的个数。

步骤 6 使用 Softmax 函数对全连接层的输出结果进行归一化,得到弹幕文本情感极性的概率分布矩阵 $\boldsymbol{L}$,对 $\boldsymbol{L}$ 按行取最大值的索引,最终得到弹幕文本的情感极性。

上述流程中的步骤 3 和步骤 4 充分利用了 CNN 训练局部特征的优势以及 BiGRU 获取上下文语义信息的优势,从而能够充分考虑文本中的局部特征信息和上下文语义关联,进一步提升了模型的效果。

3 实验部分与结果分析

3.1 实验环境与数据

实验环境如下:操作系统为 Ubuntu16.04,CPU 是 Intel® Xeon® Gold 5218,GPU 为 NVIDIA Tesla V100,内存为 48 GB,使用 Python3.6 进行算法编程,并使用 Tensorflow1.12.0 作为深度学习框架。利用网络爬虫技

术分别从哔哩哔哩、爱奇艺和腾讯视频三个视频网站上爬取弹幕文本数据,形成不同的弹幕文本数据集。对弹幕数据进行清洗,剔除掉不具有情感倾向的弹幕,并将具有情感倾向的弹幕分为负向和正向情感弹幕。数据清洗完成后,哔哩哔哩弹幕数据集共包含 5 037 个负向情感样本和 5 109 个正向情感样本,爱奇艺弹幕数据集共包含 5 014 个负向情感样本和 5 115 个正向情感样本,腾讯视频弹幕数据集共包含 5 024 个负向情感样本和 5 218 个正向情感样本。对以上数据进行预处理,并按照 7∶3 的比例划分为训练集和测试集。

3.2　评价指标

为评价模型的分类效果,采用混淆矩阵对分类结果进行统计。使用 TP 表示实际为正样本且预测为正样本,FP 表示实际为负样本但预测为正样本,TN 表示实际为负样本且预测为负样本,FN 表示实际为正样本但预测为负样本。根据混淆矩阵统计的结果,采用准确率(Acc)、精确率(P)、召回率(R)以及精确率与召回率的调和平均值($F1$)对模型效果进行评价,计算公式为

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}, \tag{8}$$

$$P = \frac{TP}{TP + FP}, \tag{9}$$

$$R = \frac{TP}{TP + FN}, \tag{10}$$

$$F1 = \frac{2 \times P \times R}{P + R}。 \tag{11}$$

3.3　实验参数

实验参数主要包括 ALBERT 模型和 CRNN 模型的参数。其中 ALBERT 采用 Google 发布的预训练模型 ALBERT-Base,其模型参数如下:嵌入层尺寸为 128,隐藏层尺寸为 768,隐藏层的层数为 12,注意力头的个数为 12,并且使用 ReLU 作为模型的激活函数。另外,在模型训练的过程中对该预训练模型进行微调,以更加适用于本文的情感分析任务。CRNN 模型参数如下:CNN 中的卷积核尺寸分别为 3、4、5,且每种尺寸卷积核的个数均为 128,另外在池化层采用最大池化的方法对特征进行降维,且池化尺寸为 4。BiGRU 中的 GRU 隐藏单元个数为 128,模型的层数为 1,采用 ReLU 作为激活函数,并在训练阶段将 Dropout 的比例设置为 0.5。ALBERT-CRNN 模型训练参数如下:设置批次大小为 64,迭代轮数为 30,由于弹幕文本一般较短,设置最大序列长度为 30,采用交叉熵损失函数,选取 Adam 作为模型的优化器,并将学习率设置为 5×10^{-5}。

3.4　对比实验设置

为验证 ALBERT-CRNN 弹幕文本情感分析模型的有效性,将 ALBERT-CRNN 模型与 SVM、CNN、BiGRU、CRNN 以及 ALBERT 模型进行对比,在哔哩哔哩、爱奇艺和腾讯视频三个视频平台的弹幕文本数据集上分别进行实验。其中 SVM、CNN、BiGRU 和 CRNN 模型均基于 Word2Vec 模型构建词向量;ALBERT 和 ALBERT-CRNN 模型则采用 Google 发布的中文预训练模型 ALBERT-Base 来进行文本特征表示,并将此预训练模型在本文数据集下进行微调。

3.5　实验结果及分析

不同模型在三个弹幕文本数据集上的精确率、召回率和 $F1$ 值结果如表 1 所示。可以看出,相比 SVM、CNN、BiGRU、CRNN 和 ALBERT 模型,ALBERT-CRNN 模型在哔哩哔哩数据集上的 $F1$ 值分别提高了 8.5、5.5、5.6、4.9 和 0.6 个百分点,在爱奇艺数据集上的 $F1$ 值分别提高了 8.1、5.0、5.5、3.8 和 0.7 个百分点,在腾讯视频数据集上的 $F1$ 值分别提高了 8.1、5.9、5.7、5.0 和 1.9 个百分点。由此可以得出,相比其他基于 Word2Vec 构建词向量的模型,ABERT 和 ALBERT-CRNN 模型在弹幕文本情感分析中有着明显的优势,证明了由预训练语言模型获取的文本特征能够充分利用句子中词的上下文信息,较好地区分了句子中同一个词在不同上下文语境中的不同含义,从而使得弹幕文本情感分析的效果得到了提升。另外,ALBERT-CRNN 模型相比 ALBERT 模型在弹幕文本情感分析中具有更优的表现,证明了 CRNN 模型能够充分考虑文本中的局部特征信息和上下文语义关联,进一步提升了模型的性能。

图 4 给出了不同模型在三个弹幕文本数据集上的准确率对比。可以发现,相比 SVM、CNN、BiGRU、CRNN

表1 不同模型在三个数据集上的精确率、召回率和*F*1值结果

Table 1 Precision, recall and *F*1-value results of different models on three datasets

数据集	模型	评价指标/%		
		P	*R*	*F*1
哔哩哔哩	SVM	82.2	89.7	85.8
	CNN	88.6	88.9	88.8
	BiGRU	89.9	87.5	88.7
	CRNN	91.0	87.9	89.4
	ALBERT	90.1	97.6	93.7
	ALBERT-CRNN	93.9	94.7	94.3
爱奇艺	SVM	84.7	85.9	85.3
	CNN	87.4	89.5	88.4
	BiGRU	88.0	87.7	87.9
	CRNN	88.6	90.7	89.6
	ALBERT	91.5	93.8	92.7
	ALBERT-CRNN	93.8	93.1	93.4
腾讯视频	SVM	86.5	87.1	86.8
	CNN	88.9	89.1	89.0
	BiGRU	85.9	92.8	89.2
	CRNN	90.4	89.5	89.9
	ALBERT	94.2	91.8	93.0
	ALBERT-CRNN	93.7	96.1	94.9

和ALBERT模型,ALBERT-CRNN模型在弹幕文本情感分析中具有更佳的效果,在三个数据集上的准确率分别达到94.3%、93.5%和94.8%,再次证明了ALBERT-CRNN模型在弹幕文本情感分析任务中的有效性。

利用ALBERT-CRNN模型对单个弹幕文本进行预测,展示了模型真实的运用功能。为方便了解输出结果的含义,将预测值大于0.5的定义为正向情感,其余的定义为负向情感。单个样例分析结果如表2所示,可以看出,ALBERT-CRNN模型对随机选取的单个弹幕文本的预测结果都是正确的,能够成功应用于弹幕文本情感分析中。

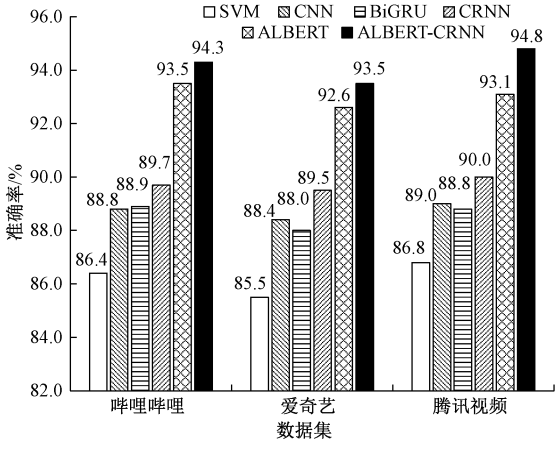

图4 不同模型在三个数据集上的准确率对比

Figure 4 Accuracy comparison of different models on three datasets

表2 单个样例分析结果

Table 2 Single sample analysis results

句子	预测值	情感
喜欢它的主题风格,很喜欢放歌的时候那个唱片旋转很有感觉	0.993	正向
他喜欢挺肚子,体态就没那么好看了	0.135	负向
说实话不喜欢广场 没以前好	0.087	负向
如晶真实自然,这就是辩论的最高智慧	0.952	正向
跟着up主可以学到好多东西,爱了爱了	0.976	正向
回顾这集看到那么多黑阿薰傻x还是好生气呢	0.063	负向

4　结语

本文提出一种结合 ALBERT 与 CRNN 的弹幕文本情感分析模型 ALBERT-CRNN。通过 ALBERT 预训练语言模型获取弹幕文本的动态特征表示,解决了传统弹幕情感分析方法无法区分句子中同一个词在不同上下文语境中含义不同的问题;使用结合 CNN 与 BiGRU 的神经网络 CRNN 对特征进行训练,充分利用了文本中的局部特征信息和上下文语义关联。在哔哩哔哩、爱奇艺和腾讯视频三个视频平台的弹幕文本数据集上进行对比实验,证明了 ALBERT-CRNN 模型在弹幕文本情感分析任务中的有效性。由于 ALBERT 模型在使用过程中的参数量仍然较大,导致训练所花费的时间较长。在下一步研究工作中,将对 ALBERT 模型进行压缩,在模型精度不受较大损失的情况下尽可能降低模型的复杂度,从而提高模型的训练效率。

参考文献:

[1] 郑飏飏, 徐健, 肖卓. 情感分析及可视化方法在网络视频弹幕数据分析中的应用[J]. 现代图书情报技术, 2015(11): 82-90.
ZHENG Y Y, XU J, XIAO Z. Utilization of sentiment analysis and visualization in online video bullet-screen comments[J]. New technology of library and information service, 2015(11): 82-90.

[2] 洪庆, 王思尧, 赵钦佩, 等. 基于弹幕情感分析和聚类算法的视频用户群体分类[J]. 计算机工程与科学, 2018, 40(6): 1125-1139.
HONG Q, WANG S Y, ZHAO Q P, et al. Video user group classification based on barrage comments sentiment analysis and clustering algorithms[J]. Computer engineering & science, 2018, 40(6): 1125-1139.

[3] 庄须强, 刘方爱. 基于 AT-LSTM 的弹幕评论情感分析[J]. 数字技术与应用, 2018, 36(2): 210-212.
ZHUANG X Q, LIU F A. Emotional analysis of bullet-screen comments based on AT-LSTM[J]. Digital technology and application, 2018,36(2): 210-212.

[4] 李平, 戴月明, 吴定会. 双通道卷积神经网络在文本情感分析中的应用[J]. 计算机应用, 2018, 38(6): 1542-1546.
LI P, DAI Y M, WU D H. Application of dual-channel convolutional neural network in sentiment analysis[J]. Journal of computer applications, 2018, 38(6): 1542-1546.

[5] 邱宁佳, 丛琳, 周思丞, 等. 结合改进主动学习的 SVD-CNN 弹幕文本分类算法[J]. 计算机应用, 2019, 39(3): 644-650.
QIU N J, CONG L, ZHOU S C, et al. SVD-CNN barrage text classification algorithm combined with improved active learning [J]. Journal of computer applications, 2019, 39(3): 644-650.

[6] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[C]// Proceedings of the 27th International Conference on Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2013: 3111-3119.

[7] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. [2020-08-17]. https://www.researchgate.net/publication/234131319.

[8] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [EB/OL]. [2020-08-17]. https://www.researchgate.net/publication/328230984.

[9] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: Association for Computational Linguistics, 2018:2227-2237.

[10] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training [EB/OL]. [2020-08-17]. https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf.

[11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. [2020-03-02]. https://www.researchgate.net/publication/317558625.

[12] LAN Z Z, CHEN M D, GOODMAN S, et al. ALBERT: a lite BERT for self-supervised learning of language representations [EB/OL]. [2020-02-13]. https://www.researchgate.net/publication/336084032.

[13] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2014: 1746-1751.

[14] SOCHER R, LIN C Y, NG A Y, et al. Parsing natural scenes and natural language with recursive neural networks[C]//Proceedings of the 28th International Conference on Machine Learning. New York: ACM Press, 2011: 129-136.
[15] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
[16] DEY R, SALEM F M. Gate-variants of gated recurrent unit (GRU) neural networks[C]//Proceedings of the 60th IEEE International Midwest Symposium on Circuits and Systems. Piscataway: IEEE Press, 2017: 1597-1600.
[17] 陈榕, 任崇广, 王智远, 等. 基于注意力机制的CRNN文本分类算法[J]. 计算机工程与设计, 2019, 40(11): 3151-3157.
CHEN R, REN C G, WANG Z Y, et al. Attention based CRNN for text classification[J]. Computer engineering and design, 2019, 40(11): 3151-3157.
[18] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [EB/OL]. [2020-05-24]. https://www.researchgate.net/publication/265252627.
[19] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[EB/OL]. [2020-05-24]. https://www.researchgate.net/publication/262877889.
[20] YANG Z L, DAI Z H, YANG Y M, et al. XLNet: generalized autoregressive pretraining for language understanding[EB/OL]. [2020-05-26]. https://www.researchgate.net/publication/333892322.

Barrage Text Sentiment Analysis Based on ALBERT-CRNN

ZENG Cheng[1,2,3], WEN Chaodong[1], SUN Yumin[1], PAN Lie[1], HE Peng[1,2,3]
(1. *School of Computer Science and Information Engineering, Hubei University, Wuhan* 430062, *China*;
2. *Hubei Province Engineering Technology Research Center for Software Engineering, Wuhan* 430062, *China*;
3. *Hubei Engineering Research Center for Smart Government and Artificial Intelligence, Wuhan* 430062, *China*)

Abstract: The barrage text sentiment analysis model ALBERT-CRNN based on ALBERT pre-training language model and convolutional recurrent neural network (CRNN) was proposed. Firstly, the ALBERT pre-training language model was used to obtain the dynamic feature representations of barrage texts, so that the same word had different word vector expressions in different contexts. Then, these feature vectors were trained by CRNN, which made the local features and context semantic correlation to be fully considered. Finally, the sentiment polarity of barrage texts was obtained by the Softmax function. Experiments were carried out on the barrage text datasets of Bilibili, iQiYi and Tencent video platforms. The experimental results showed that the accuracy of ALBERT-CRNN on the above three datasets reached 94.3%, 93.5% and 94.8% respectively, which were better than some traditional models.

Key words: barrage text; sentiment analysis; word vector; pre-training language model; convolutional recurrent neural network

(责任编辑:孔 薇 王浩毅)