

# 基于字词融合特征的微博情绪识别方法

殷昊 徐健 李寿山 周国栋

(苏州大学计算机科学与技术学院 江苏 苏州 215006)

**摘要** 文本情绪识别是自然语言处理问题中的一项基本任务。该任务旨在通过分析文本判断该文本是否含有情绪。针对该任务,提出了一种基于字词融合特征的微博情绪识别方法。相对于传统方法,所提方法能够充分考虑微博语言的特点,充分利用字词融合特征提升识别性能。具体而言,首先将微博文本分别用字特征和词特征表示;然后利用LSTM模型(或双向LSTM模型)分别从字特征和词特征表示的微博文本中提取隐层特征;最后融合两组隐层特征,得到字词融合特征,从而进行情绪识别。实验结果表明,该方法能够获得更好的情绪识别性能。

**关键词** 情绪识别,融合特征,LSTM

**中图分类号** TP391 **文献标识码** A

## Emotion Recognition on Microblog Based on Character and Word Features

YIN Hao XU Jian LI Shou-shan ZHOU Guo-dong

(School of Computer Science & Technology, Soochow University, Suzhou, Jiangsu 215006, China)

**Abstract** Text emotion recognition is an important task in the community of nature language processing. This task aims to predict the involving emotion towards a piece of text. This paper proposed a novel emotion recognition approach based on character and word features. Compared to most traditional approaches, this approach employs both the character and word features by considering the characteristic of microblog text. Specifically, the feature presentations of microblog are extracted respectively from characters and words. Then, a LSTM model (or Bi-directional LSTM model) is employed to extract the hidden feature presentations from the above feature presentations. Third, the two groups of hidden character and word feature representations are merged to perform emotion recognition. Empirical studies demonstrate the effectiveness of the proposed approach for emotion recognition on SINA microblog.

**Keywords** Emotion recognition, Fusion features, LSTM

## 1 引言

随着社交网络的迅速发展, Twitter、Face-book、微博等社交媒体逐渐成为人们沟通情感和传达信息的重要途径。它们是一种新兴的基于用户关系的信息获取、传播及分享平台, 注册用户可以随时通过电脑、手机等设备发布信息。就国内而言, 微博以更新快、信息量大、传播广等特点吸引了越来越多的用户, 已逐渐成为人们分享和获取信息的核心社交平台。在微博海量的文本信息中, 有很大一部分文本带有微博用户的情感信息。微博用户通过微博发表自己的观点, 如对名人的喜欢或厌恶、对商品的评论、对社会热点的看法及生活中的喜怒哀乐等。处理和分析这些带有情绪的言论有助于舆情监控、突发事件预警及心理疾病治疗等实际应用。因此, 对微博等社交文本的情绪分析研究有着较高的研究意义和应用价值<sup>[1]</sup>。

情绪识别和情绪分类是情绪分析的两个基本任务<sup>[2]</sup>。情绪识别(Emotion Recognition)是指对目标文本进行分析, 判断该文本是否含有情绪。下面给出情绪识别任务的示例。

A: 的确非常精彩! 看好法国队晋级! 哈哈(有情绪)

B: 诚信互粉, 在线秒回! (无情绪)

A文本包含强烈的情绪表达, 属于有情绪文本; B文本只是客观陈述, 不含有情绪。

情绪分类(Emotion Classification)是针对文本表达的个人情绪(如高兴、伤心、惊讶等)进行自动分类的方法。情绪识别是情绪分类的前提和基础, 本文研究的重点集中在情绪识别任务上。

目前, 基于机器学习的情绪识别方法通常使用词袋(bag-of-words)模型, 选用词作为文本特征进行文本表示<sup>[3]</sup>。这种文本表示方法很大程度上依赖于分词算法的性能。微博语言是微博用户为了更好地表达和相互交流而使用的带有“草根性”“娱乐性”和“创新性”的有别于传统现代汉语的口语化书面语<sup>[4]</sup>。对于微博文本这种不规范的口语化书面语, 大多数分词算法很难得到正确的分词结果。

例 1

原始文本: 哇...晓哥~~想念那玫瑰啊...哈哈...Gyoku 钰  
分词文本: 哇/...晓哥~/~/想念/那/玫瑰/啊/...哈哈...·

/Gyoku 钰

本文受国家自然科学基金(61375073, 61672366)资助。

殷昊(1994-), 男, 硕士生, 主要研究方向为自然语言处理、情感分析, E-mail: 20155227031@stu.suda.edu.cn; 徐健(1992-), 男, 硕士生, 主要研究方向为自然语言处理、情感分析; 李寿山(1980-), 男, 教授, 主要研究方向为自然语言处理、情感分析; 周国栋(1967-), 男, 教授, 主要研究方向为自然语言处理、句法分析等。

从例1可以看出,原始文本中“哈哈”表达了强烈的高兴情绪,因此可得出该条微博含有情绪。然而在分词后的文本中,“…哈哈…”被误分为一个词。在词特征空间中,“哈哈”和“…哈哈…”属于两个不同的特征,分类器很难对该微博进行情绪识别。一种简单的方法是仅使用字特征来表示微博文本。虽然字特征表示可以避免分词错误对情绪识别的影响,但是字特征表示往往不能完全表达原始文本的信息。

#### 例2

原始文本:每次心情不好就胃疼…

字特征表示:每次心情不好就胃疼…

在例2中,原始文本中“心情不好”表达了难过的情绪,因此可知该条微博含有情绪。但在字特征表示中,“心”“情”“不”和“好”等单字并不能很好地表达情绪。

因此,针对上述问题,本文尝试通过LSTM(或双向LSTM)神经网络融合微博文本的字特征和词特征来提升情绪识别的性能。具体而言,首先将微博文本分别使用字特征和词特征表示;然后利用LSTM(或双向LSTM)神经网络分别从字特征和词特征表示的微博文本中提取微博文本的隐层特征;最后融合两组隐层特征得到字词融合特征,通过全连接层和Sigmoid分类器进行情绪识别。实验结果表明,使用字词融合特征能够显著提升微博情绪识别的性能。

本文第2节介绍情绪分析的相关工作;第3节给出实验语料的描述;第4节介绍本文提出的情绪识别方法;第5节介绍实验结果和相关分析;最后给出总结,并对下一步工作进行展望。

## 2 相关工作

随着互联网及社交媒体的兴起,文本情绪分析逐渐成为自然语言处理领域的一项热点研究课题,受到了国内外研究人员的广泛关注。情绪识别和情绪分类是情绪分析的两个基本任务,下面分别介绍这两项工作的研究现状。

### 2.1 情绪识别

Aman等<sup>[2]</sup>收集了博客文章并对其进行了词语级/短语级情绪类别和强度的标注工作,使用基于知识的方法识别句子级文本的情绪状况。Maeda等<sup>[5]</sup>首先定义了3种写作风格方式,即一种正式的表达和两种非正式的表达,并在Twitter语料库上进行情绪识别研究。刘欢欢等<sup>[6]</sup>在中文情绪语料库(Ren-CECps)上研究了句子级的情绪识别方法,并比较了不同特征(词、词+词性、词+词)对情绪识别的影响。姚源林等<sup>[7]</sup>介绍了NLP&.CC2013中文微博情绪分析评测任务中关于情绪识别的相关结果。黄磊等<sup>[3]</sup>提出了一种基于句法信息的微博文本情绪识别方法,该方法充分考虑了微博文本的句法信息,能够明显提高微博情绪识别的性能。

### 2.2 情绪分类

Pang等<sup>[1]</sup>首次将机器学习的方法应用到情感分析任务中,并对电影评论进行了情感极性的判别。Li等<sup>[8]</sup>提出利用句子的标签因子图和上下文标签因子图进行句子级的情绪分类,很好地解决了数据稀疏和情绪的多标签问题。Liu等<sup>[9]</sup>通过新闻文本和评论文本之间的联系,研究了作者情绪和读者情绪的分类。Rana等<sup>[10]</sup>将神经网络方法应用于带有噪音的文本情感分类,实验证明了该方法能够很好地处理噪音问

题。Tripathi等<sup>[11]</sup>探索了深度神经网络和卷积神经网络在情绪分类任务上的性能,通过实验证明了神经网络方法比传统方法更加优越。

以往的研究大都采用词特征来对文本进行特征表示。然而,由于微博文本表达较为口语、写作不规范,通过分词程序并不能得到理想的词特征。字特征虽然可以避免上述缺点,但在中文文本中,字特征所能表达的信息往往少于词特征。因此,本文提出一种基于字词融合特征的情绪识别方法,以进一步提升情绪识别的性能。

## 3 语料收集与分析

本文的实验语料来自新浪微博,共计10819条。我们按照Huang等<sup>[12]</sup>提出的微博情绪标注规范对这些微博文本进行了标注,情绪分类体系如图1所示。

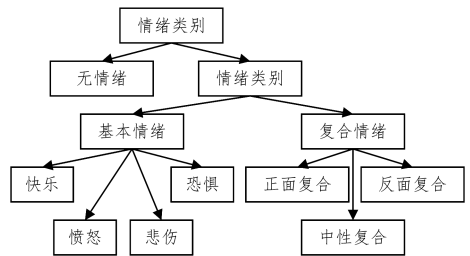


图1 情绪分类体系

其中,有情绪微博3046条,无情绪微博7773条。表1给出了标注结果中微博文本有无情绪的实例。

表1 微博有无情绪的实例

微博文本	有无情绪
我是微博达人咯~恭喜恭喜~偷笑鼓掌	有
又要上班了,郁闷!只想睡觉!	有
邓华德签约新疆男篮—新浪体育	无
【12 星座守信用程度】巨蟹座信用程度95%…	无

## 4 基于字词融合特征的微博情绪识别

传统的基于机器学习的微博情绪识别方法通常使用词袋(BOW)模型对文本进行特征表示<sup>[3]</sup>。BOW模型使用词作为文本的特征,将微博文本看成是若干词的集合,这种做法很大程度上依赖于分词算法的性能,尤其是对于微博这种口语化书面语,分词程序很难将其中的不规范用语、网络词语等正确切分。如果使用字特征对微博文本进行特征表示,虽然可以避免上述问题,但字特征所表达的信息远少于词特征所表达的信息。词是语言中最小的能够独立使用的单位,将词拆成字表示将在一定程度上丢失文本的信息。

本文提出的基于字词融合特征的微博情绪识别方法在一定程度上融合了字特征与词特征所表达的信息,使得词特征和字特征能够互相补充,更好地表示微博文本。具体而言,首先分别使用词特征和字特征对微博文本进行特征表示;接着使用长短时记忆神经网络(Long Short-term Memory,LSTM)分别从词特征和字特征所表示的文本中提取更深层次的隐层特征;最后将两组隐层特征进行融合得到字词融合特征,并输入全连接层和Sigmoid层,实现微博文本的情绪识别。

### 4.1 长短时记忆神经网络(LSTM)

LSTM是对传统循环神经网络(RNN)的改进,其使用记忆单元来对历史信息进行更新,可以解决RNN在反向传播

过程中的梯度消失和梯度爆炸问题。Alex Graves 对 LSTM 进行了改进,使得 LSTM 被广泛应用于自然语言处理、语音识别等领域中<sup>[13]</sup>。

使用 LSTM 可以学习到文本的长期依赖关系,历史信息的更新和利用受到 3 个门的控制。LSTM 单元在  $t$  时刻的更新过程如下:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

其中,  $x_t$  为 LSTM 在  $t$  时刻的输入,  $h_{t-1}$  为 LSTM 在  $t-1$  时刻的输出,  $\tilde{c}_t$  为记忆单元在  $t$  时刻的候选值。  $W, U$  是参数矩阵,  $b$  是偏置向量。  $\sigma$  表示 Logistic 函数, 符号  $\odot$  表示向量的 Hadamard 积。遗忘门  $f_t$  用来决定当前时刻哪些旧信息需要从记忆单元中抛弃; 输入门  $i_t$  决定当前时刻记忆单元中需要存储哪些新的信息; 输出门  $o_t$  决定了当前时刻记忆单元中哪些信息需要输出。遗忘门  $f_t$  与上一时刻记忆单元的值  $c_{t-1}$  相乘即可得到上一时刻记忆单元保存的信息, 输入门  $i_t$  与当前时刻记忆单元的候选值  $\tilde{c}_t$  相乘即可得到当前时刻新加入记忆单元的信息, 将两者相加即可得到记忆单元在当前时刻的值  $c_t$ 。使用  $\tanh$  函数将记忆单元  $c_t$  的值映射到  $-1$  与  $1$  之间, 起到了稳定数值的作用。最后将输出门  $o_t$  与  $\tanh(c_t)$  相乘, 得到 LSTM 单元在  $t$  时刻的输出  $h_t$ 。

#### 4.2 双向长短时记忆神经网络 (BLSTM)

传统的循环神经网络 (RNN) 是从序列的一端向另一端逐个读取输入数据, 因此在任意时刻神经网络里存储的数据只有当前时刻之前的信息。为了在当前时刻加入未来的上下文信息, Schuster 等<sup>[14]</sup> 提出了双向循环神经网络 (BRNN)。

BRNN 训练两个 RNN, 一个 RNN 从前往后计算隐层向量  $\vec{h}_t$ , 另一个 RNN 从后往前计算隐层向量  $\overleftarrow{h}_t$ 。两个 RNN 连接着同一个输出层, 在该输出层输入序列中, 每个节点都包含完整的过去和未来的上下文信息。在本文中, 两个隐层向量按式 (7) 进行结合。

$$h_t = \vec{h}_t + \overleftarrow{h}_t \quad (7)$$

图 2 所示为沿时间展开的循环神经网络。从图中可以看出, 隐藏层融合了前向层和后向层的信息, 在每一个时刻都能够考虑上下文信息。将双向循环神经网络中的前向层和后向层分别替换为 LSTM 单元, 即可实现双向长短时记忆神经网络 (BLSTM)。BLSTM 能够学习文本的长期依赖关系, 又能够同时利用历史信息和未来信息, 在自然语言处理、语音识别等领域应用十分广泛<sup>[14]</sup>。

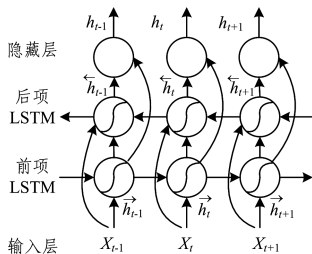


图 2 在时间上展开的 BRNN

#### 4.3 基于字词融合特征的情绪识别方法

为了能够充分利用文本的词特征和字特征信息, 我们设计了一种双通道的 LSTM 神经网络模型。该模型融合文本的词特征信息和字特征信息, 能够学习到文本更深层次的特征表示。模型的结构如图 3 所示。

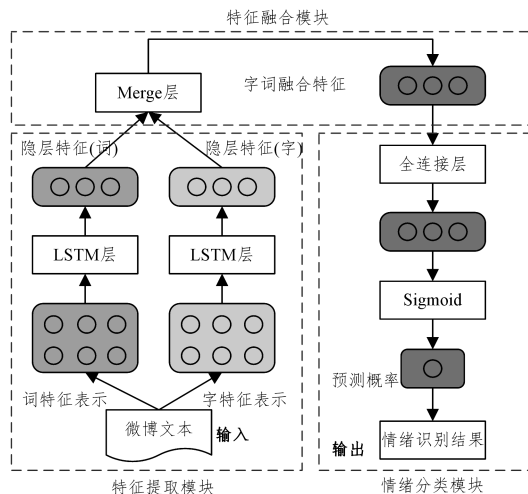


图 3 LSTM 模型结构

首先将文本分别表示为词特征表示  $T_w$  和字特征表示  $T_c$ , 接着使用 LSTM 从两组特征表示中学习更高维度的隐层特征。

$$h_w = LSTM_w(T_w) \quad (8)$$

$$h_c = LSTM_c(T_c) \quad (9)$$

其中,  $h_w$  和  $h_c$  分别表示由 LSTM 学习到的词隐层特征和字隐层特征。

为了得到文本的字词融合特征, 将  $h_w$  和  $h_c$  进行拼接融合, 并将结果作为 Merge 层的输入。

$$h_{wc} = h_w \oplus h_c \quad (10)$$

其中,  $h_{wc}$  即为文本的字词融合特征, 符号  $\oplus$  表示向量的拼接。

模型接下来是一个全连接层, 用于学习字词融合特征的更深层表示。为了降低模型的复杂度, 防止网络过拟合训练样本, 我们在全连接层上面加入了 Dropout 层。Dropout 层可以在模型训练时随机让网络中的某些隐层节点不工作<sup>[15]</sup>。

$$h^* = \varphi(W^T h_{wc} + b) \quad (11)$$

$$h^d = h^* \cdot D(p) \quad (12)$$

其中,  $h^*$  是字词融合特征  $h_{wc}$  经过全连接层后的输出,  $W^T$  和  $b$  分别为相应的权重矩阵和偏置向量。  $\varphi$  是一个激活函数, 本实现中该层采用 Relu 函数作为激活函数<sup>[16]</sup>。Relu 会使得网络中一部分神经元的输出为 0, 减少了参数之间的依赖性, 更接近生物学的激活模型, 缓解了过拟合的发生。Relu 的激活函数如式 (13) 所示:

$$\varphi(x) = \max(0, x) \quad (13)$$

Relu 将小于 0 的输出值全部置 0, 具有引导适度稀疏的能力。  $D(p)$  表示 dropout 操作, 其中  $p$  是一个可调节的超参数, 代表模型每次迭代训练时需要丢弃的隐层单元节点的比率。  $h^d$  是  $h^*$  经过 dropout 操作后的输出。

最后, 模型通过 Sigmoid 层完成微博文本的情绪识别任务。Sigmoid 层接受上一层的输出作为输入, 输出长度为 1 的

一维向量。通过 Sigmoid 函数将该向量的值映射到 0~1 之间,作为模型对微博文本是否含有情绪的预测概率  $p$ ,其计算过程如下:

$$p = \text{sigmoid}(W^d h^d + b^d) \quad (14)$$

其中, $p$  是预测概率, $W^d$  和  $b^d$  分别表示该层的权值矩阵和偏置向量。Sigmoid 函数如下:

$$\text{sigmoid}(x) = 1/(1 + e^{-x}) \quad (15)$$

选用最小化交叉熵代价函数作为该模型的目标损失函数,该函数具体如下:

$$C = -\frac{1}{m} \sum_{i=1}^m [y^i \ln a^i + (1 - y^i) \ln(1 - a^i)] + \frac{\lambda}{2m} \sum_w w^2 \quad (16)$$

其中, $m$  为训练样本的个数, $y^i \in \{0, 1\}$  为期望输出, $a^i \in (0, 1)$  为模型的实际输出。式(16)中的第一项为原始代价函数,后一项为正则化项, $\lambda$  为正则项系数, $w$  为网络权值。正则化项可以降低模型的复杂度,防止网络过拟合的发生。

在训练过程中,使用反向传播算法(Back Propagation, BP)对权值  $w$  进行更新:

$$w \leftarrow w - \eta \frac{\partial C}{\partial w} \quad (17)$$

其中, $\eta$  为学习率。

## 5 实验

### 5.1 实验设置

本实验选用新浪微博语料作为实验语料,语料的来源和分布已在第 3 节中介绍。我们分别选取 3000 条有情绪微博和无情绪微博作为实验样本。其中,80% 样本作为训练集,其余的 20% 样本作为测试集。

选用 LSTM 或 BLSTM 神经网络作为基本分类算法。实验中所用模型使用深度学习开源框架 Keras<sup>1)</sup> 搭建,后端使用 Theano<sup>2)</sup> 进行计算。在进行实验之前,首先采用复旦大学自然语言处理实验室公布的分词工具 FudanNLP<sup>3)</sup> 对中文微博文本进行分词。

实验中所用分类特征为一元词特征(Unigram)的 one-hot 表示。具体而言,首先将训练集中的词特征按出现频率降序构建词典;然后将微博文本中出现的词特征用该特征在词典中的下标来表示,由此构建微博文本的特征向量。LSTM 神经网络模型的具体参数设置如表 2 所列。

表 2 LSTM 神经网络的参数设置

参数描述	参数值
最大特征数	30000
LSTM 层输出维数	128
全连接层输出维数	64
Dropout rate	0.5
优化器	Adam <sup>[17]</sup>
迭代次数	15

实验采用正确率(Accuracy)作为衡量分类效果的标准。Accuracy 的计算方法如下:

$$\text{Accuracy} = \frac{1}{m} \sum_{i=1}^m 1_{y_i = y_i'} \quad (18)$$

其中, $m$  是测试样本的个数, $y_i$  是样本真实标签, $y_i'$  是样本的预测标签。

### 5.2 实验结果

实验实现了以下几种微博文本情绪识别方法。

1) LSTM(或 BLSTM)+字特征:统计训练样本中出现的所有字特征,并按字特征出现的频率降序排序,生成字特征列表。将微博文本中出现的字特征用其在字特征列表中对应的下标表示,由此构建样本的字特征向量。使用 LSTM(或 BLSTM)从字特征向量中学习隐层特征,通过全连接层和 Sigmoid 层进行微博情绪的识别。

2) LSTM(或 BLSTM)+词特征:统计训练样本中出现的所有词特征,并按词特征出现的频率降序排序,生成词特征列表。将微博文本中出现的词特征用其在词特征列表中对应的下标表示,由此构建样本的词特征向量。使用 LSTM(或 BLSTM)从词特征向量中学习隐层特征,通过全连接层和 Sigmoid 层进行微博情绪的识别。

3) LSTM(或 BLSTM)+字+词(混合特征):统计训练样本中出现的所有字特征和词特征,将二者混合并按出现频率降序排序,生成字词混合特征列表。将微博文本中出现的字特征和词特征用其在字词混合特征列表中对应的下标表示,由此构建样本的字词混合特征向量。使用 LSTM(或 BLSTM)从字词混合特征向量中学习隐层特征,通过全连接层和 Sigmoid 层进行微博情绪的识别。

4) LSTM(或 BLSTM)+字+词(融合特征):首先使用上述方法得到样本的字特征向量和词特征向量。分别使用 LSTM(或 BLSTM)从两种特征向量中学习更高层次的隐层特征。将两组隐层特征拼接融合得到字词融合特征,通过全连接层和 Sigmoid 层进行微博情绪的识别。

图 4 比较了几种基于 LSTM 的微博情绪识别方法的性能。从图中可以看出,使用字特征的情绪识别方法的性能最低,在训练样本取不同比例时,使用字特征时的 Accuracy 值都低于使用词特征的情绪识别方法。由于词特征是表达信息的基本单位,其包含的信息量要大于字特征,因此在情绪识别实验中使用词特征的性能较高。

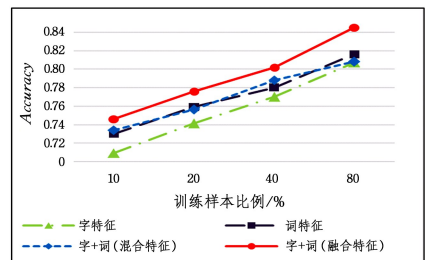


图 4 基于 LSTM 的情绪识别方法的性能比较

接下来比较使用字+词(混合特征)的情绪识别方法与仅使用字或词特征的情绪识别方法的性能。从图 4 中可以看出,字+词(混合特征)的性能明显优于使用字特征的性能,但与词特征的性能相比,并无显著提高。在训练样本取 10% 和 40% 时,字+词(混合特征)的性能优于词特征的性能;训练样

<sup>1)</sup> <http://Keras.io/>

<sup>2)</sup> <http://deeplearning.net/software/theano/>

<sup>3)</sup> <https://code.google.com/archive/p/fudanlpl/>

本取 20% 和 80% 时,词特征的性能优于字+词(混合特征)的性能。出现该现象的主要原因是字+词(混合特征)中,字特征与词特征并未显著区分,未能起到相互辅助的作用。

为了解决字+词(混合特征)方法中字特征与词特征难以区分的问题,我们实现了字+词(融合特征)的情绪识别方法。使用 LSTM 分别从微博文本的字特征表示和词特征表示中学习隐层特征,然后将两组隐层特征拼接融合得到字词融合特征,将其作为微博文本新的高维特征并用于分类。从图 4 中可以看出,该方法的性能明显优于其他方法,从而验证了该特征融合方式对情绪识别任务的有效性。

此外,我们还实现了基于 BLSTM 的微博情绪识别方法,其性能如图 5 所示。从图 5 可以看出,使用字特征的性能要低于使用词特征的性能;使用字词混合特征的性能与使用词特征的性能相当,并无显著提高;本文提出的基于字词融合特征的情绪识别方法的性能明显优于其他方法。这些结论和上述基于 LSTM 实验得到的结论一致,从而进一步验证了本文方法的有效性。

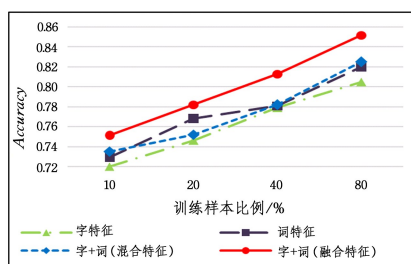


图 5 基于 BLSTM 的情绪识别方法的性能比较

图 6 比较了在基于字词融合特征的情绪识别方法中分别使用 LSTM 和 BLSTM 提取特征的微博情绪识别性能。在训练样本取不同比例的对比实验中, BLSTM 的性能都优于 LSTM 的性能。这些结果表明, BLSTM 通过结合文本的正向信息和反向信息,能够学习到文本更长的依赖关系,从而获得更好的文本表示。

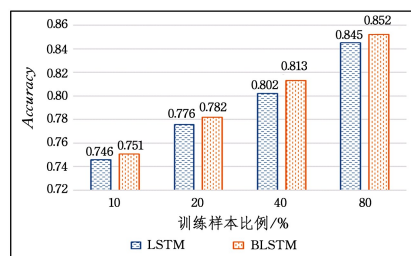


图 6 基于 LSTM 和 BLSTM(融合特征)的情绪识别方法的性能比较

**结束语** 本文提出了一种基于字词融合特征的微博文本情绪识别方法,该方法能够同时考虑微博文本的字特征和词特征。首先,将微博文本分别使用字特征和词特征表示;其次,使用 LSTM(或 BLSTM)从两组特征表示中分别学习隐层特征;最后,融合两组隐层特征得到字词融合特征,通过全连接层和 Sigmoid 层进行分类,得到情绪识别结果。实验结果表明,该方法能够有效融合字特征和词特征,性能上明显优于仅使用字特征或词特征的情绪识别方法。

下一步工作中,我们将考虑使用微博文本的其他特征(如

词性特征、句法特征)与字词特征融合来进一步提升微博情绪识别任务的性能。此外,我们也将考虑如何针对微博文本的特点设计更适合微博文本的分词算法,并利用更合适的分词算法获取更佳的词特征用于情绪识别任务。

## 参考文献

- [1] PANG B, LEE L, VAITHYANATHAN S. Thumbs up?: Sentiment Classification Using Machine Learning Techniques[C]// Proceedings of Empirical Methods in Natural Language Processing (EMNLP 2010). 2002:79-86.
- [2] AMAN S, SZPAKOWICZ M S. Identifying Expressions of Emotion in Text[C]// Text, Speech and Dialogue, International Conference. 2007:196-205.
- [3] 黄磊,李寿山,周国栋.基于句法信息的微博情绪识别方法研究[J]. 计算机科学, 2017, 44(2):244-249.
- [4] 王攸然. 微博语言特点初步分析[J]. 河南广播电视大学学报, 2015, 28(1):70-72.
- [5] MAEDA H, SHIMADA K, ENDO K. Twitter Sentiment Analysis Based on Writing Style[M]// Advances in Natural Language Processing. Springer Berlin Heidelberg, 2012:278-288.
- [6] 刘欢欢,李寿山,周国栋,等. 中文情绪识别方法研究[J]. 江西师范大学学报(自然版), 2013, 37(2):120-124.
- [7] 姚源林,王树伟,徐睿峰,等. 面向微博文本的情绪标注语料库构建[J]. 中文信息学报, 2014, 28(5):83-91.
- [8] LI S S, HUANG L, WANG R, et al. Sentence-level Emotion Classification with L-label and Context Dependence[C]// ACL (1). 2015:1045-1053.
- [9] LIU H H, LI S S, ZHOU G D, et al. Joint Modeling of News Reader's and Comment Writer's Emotions[C]// ACL (2). 2013:511-515.
- [10] RANA R. Emotion Classification from Noisy Speech-A Deep Learning Approach[J]. arXiv preprint arXiv:1603.05901, 2016.
- [11] TRIPATHI S, ACHARYA S, SHARMA R D, et al. Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset[C]// Twenty-Ninth IAAI Conference. 2017.
- [12] GRAVES A. Supervised Sequence Labeling with Recurrent Neural Networks[M]. Springer Berlin Heidelberg, 2012.
- [13] HUANG L, LI S S, ZHOU G D. Emotion Corpus Construction on Microblog Text[M]// Chinese Lexical Semantics. Springer International Publishing, 2015.
- [14] SCHUSTER M, PALIWAL K K. Bidirectional Recurrent Neural Networks[J]. IEEE Transactions on Signal Processing, 1997, 45(11):2673-2681.
- [15] 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.
- [16] GLOROT X, BORDES A, BENGIO Y. Deep Sparse Rectifier Neural Networks[C]// International Conference on Artificial Intelligence and Statistics. 2011.
- [17] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[J]. arXiv:1412.6980, 2017.