

# 基于多特征融合的混合神经网络模型讽刺语用判别\*

---

\* **收稿日期:**                      **定稿日期:**

**基金项目:** 安徽省自然科学基金项目(1508085QF119)、国家自然科学基金重点项目(61432004)、模式识别国家重点实验室开放课题(NLPR)(201407345), 中国博士后科学基金项目(2015M580532)

**作者简介:** 孙晓 (1980-), 男, 副教授, 博士, 人机交互, 自然语言处理, 情感计算; 何家劲 (1992-), 男, 硕士, 自然语言处理; 任福继 (1959-), 男, 教授, 博士, 主要研究方向人工智能与多媒体计算。

孙晓<sup>1</sup>, 何家劲<sup>1</sup>, 任福继<sup>1,2</sup>

(1. 合肥工业大学, 安徽 合肥 230009; 2. 德岛大学 日本 7700855)

**摘要:** 在社交媒体中, 存在大量的反讽和讽刺等语言现象, 这些语言现象往往表征了一定的情感倾向性。然而这些特殊的语言现象所表达的语义倾向性, 通常与其浅层字面含义相去甚远, 因此加大了社交媒体中文本情感分析的难度。鉴于此, 本文主要研究中文社交媒体中的讽刺语用识别任务, 构建了一个覆盖反讽、讽刺两种语言现象的语料库, 基于此挖掘反讽和讽刺的语言特点。本文通过对比一些有效领域特征, 验证了在反讽和讽刺文本的识别中, 其结构和语义等深层语义特征的重要性。基于此, 提出了一种有效的多特征融合的混合神经网络判别模型, 融合了卷积神经网络与 LSTM 序列神经网络模型, 通过深层模型学习深层语义特征和深层结构特征, 该模型获得了较好的识别精度, 优于传统的单一的神经网络模型和 BOW (Bag-of-Words) 模型。

**关键词:** 讽刺; 神经网络; 多特征融合; 情感分析

**中图分类号:** TP391

**文献标识码:** A

## **Ironically linguistic judgement based on hybrid neural network model via multi-feature fusion**

Xiao Sun<sup>1</sup>, Jiajin He<sup>1</sup>, Fuji Ren<sup>1,2</sup>

(1. HeFei University of Technology ,Hefei, Anhui 230009, China ; 2.The University of Tokushima, 7700855, Japan)

**Abstract:** In social media, there are a lot of ironically linguistic phenomenon of irony and satire, it implies a certain amount of emotional tendencies. However, the semantics tendencies of these special language phenomena are most often a far cry from its shallow literal meaning, therefore, it increases the difficulty of text sentiment analysis in social media. In view of this, this paper studies ironically linguistic recognition tasks in Chinese social media, constructs a corpus contains irony and satire, mining ironically language characteristics according to the corpus. In this paper, via comparing some effective field features with each other, it demonstrates the importance of structural and semantic features of irony and satire in text recognition. This paper also presents an efficient multi-feature hybrid neural network model, the model fuses the Convolutional Neural Network and LSTM sequential models, it can learn more deep layer's structural and semantic features and gained the better recognition accuracy, superior to the traditional neural network models and BOW (bag-of-words) model.

**Key words:** Irony; Neural Networks; Multi-feature Fusion; Sentiment Analysis

### **1 引言**

反讽和讽刺是特殊的修辞表达方式, 反讽是通过说“反语”的方式来表达一种情感, 通常为负面。讽刺是通过一些修辞的表达方式, 对一种现象或事物加以批判, 通常以积极的字面意思来表达消极的情绪。两个存在细微的差别, 因此两者之间的识别更是一项富有挑战性的工作。由于批判性、讽刺性的言论多见于微博、博客等社交媒体中, 讽刺和反讽语言严重影响了社交媒体情感分析的识别精度, 因此针对反讽和讽刺的研究, 在情感分析和人机会话等 NLP 问题中, 具有重要意义。

反讽和讽刺识别的研究目前主要存在以下难点: (1) 至今没有一个完整的、权威的中文语料库用于研究; (2) 单纯的文本信息, 缺乏说话人的语音、语调以及说话人的状态信息, 难以判断是否是讽刺或者反讽; (3) 由于社交媒体或者会话中, 以短文本多见, 文本长度制约了对上下文信息的获取; (4) 反讽和讽刺的表达方式与字面意思无关, 需要获取深层的语义信息, 才能准确的识别反讽和讽刺; (5) 由于反讽和讽刺与语言习惯有关, 不同语言的语言结构对其有着重大影响。

针对反讽和讽刺的识别,目前已经引起了部分研究人员的关注,但是主要是针对英文中反讽和讽刺的研究,面向中文的反讽和讽刺的识别研究较少。中文的语言结构复杂,常常通过一些谐音词、引用经典语句、网络流行词汇等方式来表达反讽和讽刺,因此中文的反讽和讽刺研究相比外文的研究更有难度。鉴于此,本文在参考针对外文的相关研究基础上,通过对比不同特征对讽刺语用识别的影响,分析讽刺语用现象的主要特点,基于此,做出一些基础的研究工作。

此外,为了解决反讽和讽刺的识别问题,本文在结合了反讽和讽刺表达的语言结构、深层语义、上下文关系特点的基础上,提出了一种面向社交媒体中反讽和讽刺识别的混合神经网络模型,该模型能够在各项指标中,优于其他模型。

## 2 相关工作

讽刺语用识别属于文本情感分析中的重要组成部分,因此众多学者对此展开了研究。近些年,针对讽刺语用的研究多为英文语料,中文由于其语言的复杂性,研究较少。Konstantin Buschmeier<sup>[1]</sup>等人在针对产品评论的“反讽”研究中,通过不同特征和模型的组合实验,发现在“反讽”这类特定任务的分类中,特定特征对实验的分类准确率提升明显,但是降低了召回率,研究表明这些特定特征和词袋模型(BOW)的结合能够很好解决召回率较低的问题,使用逻辑回归获得了最高F值为0.74。Edwin Lunando<sup>[2]</sup>在针对印尼社交媒体中讽刺检测的研究,使用消极信息和感叹词数量等特征对文本分别进行分级分类和直接分类,发现直接分类的效果更优,在Max E模型上获得了最高的实验准确率78.4%。David Bamman<sup>[3]</sup>等人在针对Twitter的讽刺检测时,发现Twitter的上下文特征对讽刺文本的分类起着至关重要的作用,比如同一作者的其他Twitter、其好友回复作者的内容以及作者评论别的帖子的内容,均对分类的精度有着重大影响。Aditya Joshi<sup>[4]</sup>等人在自动检测讽刺的文章中综述性地表述了当前该领域研究的现状和主流的模型和特征方法。Peng Liu<sup>[5]</sup>在不平衡语料的“反讽”检测中提出一种多策略的分类器,能够很好的解决“反讽”文本数据不平衡的问题,在多个语料上进行实验最终获得了最高为89.7%的准确率。Santosh Kumar Bharti<sup>[6]</sup>提出了两种讽刺的检测方式,一种是基于语义解析的词典生成算法,另一种是感叹词出现的次数的统计方法,前者获得了89%的准确率,后者获得了85%的准确率。Francesco Barbieri<sup>[7]</sup>中对修辞性的语言现象进行了研究,在Twitter上对多种修辞语言现象进行了实验,其中包括讽刺的识别,但是该类别的识别仅仅达到0.60的F值。以上的所有模型均使用传统的统计模型,未能将反讽的深层语义信息进行挖掘,因此识别精度不高。

## 3 特征说明

### 3.1 词统计特征

在很多传统经典的NLP问题,比如文本分类,情感分析中,统计特征往往起着重要作用,比如TFIDF、卡方、n-gram(一元、二元、三元)等。本文为了实验的可靠性,对比了上述统计特征在SVM分类器中的最佳分类效果,最终选择了TFIDF+2-gram作为最佳的特征组合。在解决反讽、讽刺问题上,基于词的统计特征能够很好的获取到一些有利于反讽、讽刺识别的关键词。

### 3.2 句法结构特征

#### (1) 搭配规则

由于反讽属于一种特殊的修辞手法,句式较普通陈述句不同,存在一些固定词语的同现规则,为了不失研究的一般性,避免人为的主观判断对实验结果产生影响,本文采用一种自动化的方式构建反讽、讽刺常见搭配规则。文献<sup>[8]</sup>重点介绍了,针对反讽问题的一些手工收集的搭配规则。由于反讽的搭配规则多为skip-n的二次组合,本文采用一种skip-n元词的方

案,对文本进行解析,自动化构造出 *skip-n* 元词,其中涵盖多数文献<sup>[8]</sup>中所提出的手工模板。

例: |很好|今天|下了|倾盆大雨|, |我|又|没|带|伞|

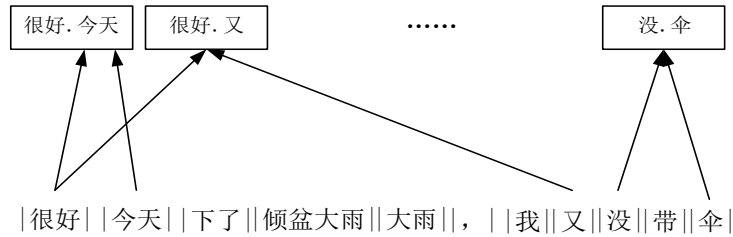


图 1 *skip-n* 元词解析示意图

一个分词长度为  $n$  的一个句子,可以构造出  $c_n^2$  个组合词语,称之为“*skip-n* 元词”,对“*skip-n* 元词”进行词频统计,获得同现统计值较高的“*skip-n* 元词”作为反讽、讽刺的搭配规则,计算过程如公式(1)所示。

$$\Phi(P) = \lambda(P, C) \quad (1)$$

其中  $\lambda(P, C)$  表示了 *skip-n* 元词  $P$  在  $C$  类文本中的卡方值。

部分解析计算结果如下表所示:

表 1 *skip-n* 元词统计评分实例

编号	规则	评分
1	可以...一点	2.52531189081
2	可以...再	2.70140289786
3	再...一点	2.79237467606
4	很好...annoy	3.91160625193
5	很好...又	4.22745920135

## (2) 情感语义特征

根据 Tang<sup>[8]</sup>对反讽和讽刺的解读,反讽又称为说反话,一般会存在两个分句或者两个语义块。反讽中存在一个特殊的语义现象,那就是前后两分句的语义往往存在一个强烈的情感倾向的反差,这一特性可以用于实现对反讽和讽刺的识别。本文主要采用两种方法,对语义倾斜现象加以检测。

引入情感词典(基于大连理工情感词汇本体<sup>2</sup>的扩充),该词典中包含了该词语的情感倾向以及情感分数,通过检测句子中的情感词以及否定词对每个分句的语义情感,并计算两分句的情感差值,将此作为情感倾斜特征。

$$IS_1 = \max_{(w_1, w_2 \in \text{text})} \left( \frac{|(-1)^m Es(w_1) - (-1)^n Es(w_2)|}{\max Es - \min Es} \right) \quad (2)$$

为了克服情感词典的局限性,很多新词都难以识别其情感,因此引入 Word2Vec<sup>[9]</sup>(300w 条中文微博语料训练获得)来平衡这一缺陷,通过检测两个语义块中的情感词的词向量之间的余弦值,若是前后存在明显的语义偏差,则夹角大于 90 度,将余弦相似度这一特征引入本文中。

$$IS_2 = \frac{1}{10 \min \{ \text{Sim}[(-1)^m \text{vec}(w_1), (-1)^n \text{vec}(w_2)] \}} \quad (3)$$

<sup>2</sup> <http://www.datatang.com/data/45448>

综合上述两种针对语义情感不平衡分数（Imbalance Score, IS）计算的方法，本文在实验中综合了二者的优点，表达如公式（4）所示：

$$IS(\text{text}) = \alpha IS_1 + (1 - \alpha) IS_2, (\alpha \in (0, 1)) \quad (4)$$

其中， $m$ ， $n$  分别为情感词  $w_1$ ， $w_2$  前修饰否定词的数量， $Es(w)$  表示情感词  $w$  的情感得分， $vec(w)$  表示情感词的词向量。

### 3.3 词向量(Word2Vec)

词向量的表示方法有多种，Mikolov<sup>[9]</sup> 提出了两种新的模型，分别是 Continuous Bag-of-Words Model（CBOW）和 Continuous Skip-gram Model (Skip-gram)。本文采用 Skip-gram 算法计算词向量，Skip-gram 算法的目标是使得如下条件概率最大：

$$\max_{\theta} \prod_{(w,c) \in D} p(c | w; \theta) \quad (5)$$

其中  $D$  为单词  $w$  和他的上下文词汇构成的集合， $w$  为输入的词汇， $c$  为上下文词汇，寻找使得该概率乘积最大化的参数集合  $\theta$ 。在下文给出了词向量训练过程的详细参数设置以及训练过程。

### 3.4 词序与随机词向量(Word\_Index)

对所有进行词序编码，并将词序映射为随机词向量，进而由词序构成句子向量。本文的语料通过文本文档保存，通过读取训练集和测试集文档（Doc）中的所有句子（S），构建词库（Voc），并对词库中的词语进行随机编序，针对词序进行词向量的映射，从而构建句子特征矩阵（Mat）。

$Doc = (S_1, S_2, S_3, \dots, S_m)$ ， $Voc = (v_1, v_2, v_3, \dots, v_n)$ ，其中  $S_m$  表示一条语料， $v_n$  表示一个词语。由此可以将一条语料表示为  $S_m = (v_p, v_{p+r}, \dots, v_q)$ ，其中  $v_{p+r}$  可映射为  $dim$  维的向量，该过程采用随机初始化的方式实现， $V_{p+r} = (e_1, e_2, e_3, \dots, e_{dim})$ ，进而构成句子矩阵，实现对文本的细粒度序列特征表示。

## 4 混合神经网络模型

为了更好的能够更好地提取反讽、讽刺文本的特征，结合上文所述反讽、讽刺等语言现象的特殊语义结构特征，本文提出一种基于神经网络的输出层融合的混合模型。该模型主要由两部分组成，选择 CNN<sup>[10]</sup> 作为语义信息的特征提取器，将时间序列模型 LSTM<sup>[11]</sup> 作为上下文语义结构的映射器，协同训练双通道模型，并融合这两者的输出特征进行分类。

该模型将每一个输入文本  $S_i$  在  $\Phi$  的作用下特征转换，转换为长度一致的  $S_i' \in \mathbb{R}^d$ ，针对等长的  $S_i'$  以两种映射方式作为模型的输入特征，一种是引入 3.3 的词向量 Word2Vec<sup>[9]</sup> 构成特征矩阵  $S_{ct}' \in \mathbb{R}^{d \times L}$  作为 CNN 通道的特征输入。引入 3.4 的词序作为文本的另一种特征表示  $S_{Lt}' = (w_1, w_2 \dots w_L)$ ，将该特征表示方法作为 LSTM 通道的输入特征。

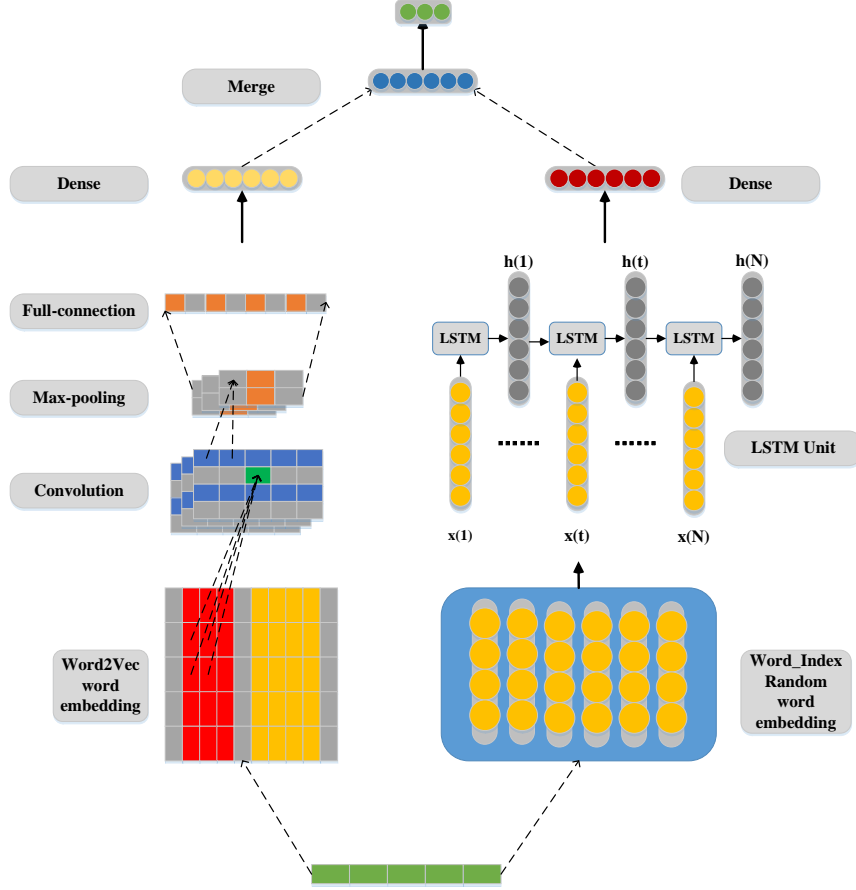


图 2 多特征混合神经网络模型图

#### 4.1 卷积通道 (Convolutional Channels, CCs)

卷积通道 CCs 接收 3.3 所述词向量特征  $S_{ct}$  作为输入特征，在卷积层的作用下进行卷积运算获得输入文本的语义信息和上下文信息，给定输入文本特征表示为：

$X_{m,n} \in S_{ct}, (0 \leq m < d, 0 \leq n < L)$ ，卷积运算过程如下所示：

$$X_{m,n} = f\left(\sum_{i=1}^u \sum_{j=1}^v W_{i,j} \cdot X_{m-i+u, n-j+v} + b\right) \quad (6)$$

本文为了增强卷积层的能力，选择了 32 组卷积核获得 32 组卷积输出。以下表示第  $L$  层的第  $K$  组的特征映射：

$$X^{(l,k)} = f\left(\sum_{p=1}^{n_l-1} (W^{(l,k,p)} \otimes X^{(l-1,p)}) + b^{(l,k)}\right) \quad (7)$$

由上述表达可知训练过程 CCs 的权值  $w$  和  $b$  是共享的，因此可以降低计算的复杂度，提升训练速度。

经过卷积层  $C$  特征提取后能够获得多组卷积输出，将该输出作为采样层  $S$  的输入，采样层能够大大降低卷积输出的高维向量的维度。采样层的计算过程可表示如公式(8)：

$$X^{(l+1)} = f(w^{(l+1)} \cdot pool(X^{(l)}) + b^{(l+1)}) \quad (8)$$

其中该模型选择最大值池化函数作为采样函数：

$$pool(R_k) = \max_{i \in R_k} R_{ki} \quad (9)$$

采样层  $S$  的输出作为带有 Dropout 的全连接层  $F$  的输入，并将全连接层的输出作为卷积通道的输出。

## 4.2 长短期记忆通道 (Long short term memory Channels, LSTMCs)

长短期记忆通道 LSTMCs 以随机词序作为输入，LSTMCs 作为特征映射器。LSTMCs 接收等长的词序向量  $S_i' \in \mathbb{R}^d$  作为文本表示，在序列模型的  $W$  参数的作用下将输入  $X_t$  和前一个时刻的隐含层输出值  $h_{t-1}$  映射成该时刻的输出  $Z_t$  并更新该时刻的隐含层输出值  $h_t$ 。因此可以根据上一时间节点的信息计算出当前时间节点的预测信息  $h_t$ ，本文使用<sup>[11]</sup>中所描述的 LSTM 单元，使用 sigmoid 函数作为激活函数， $\sigma(x) = (1 + e^{-x})^{-1}$ ，sigmoid 能够将输入值映射到  $[0,1]$  之间。使用双曲线正弦非线性函数 (10)，将输入值映射到  $[-1,1]$  之间。

$$\phi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (10)$$

LSTM 单元结构如图 2 所示：

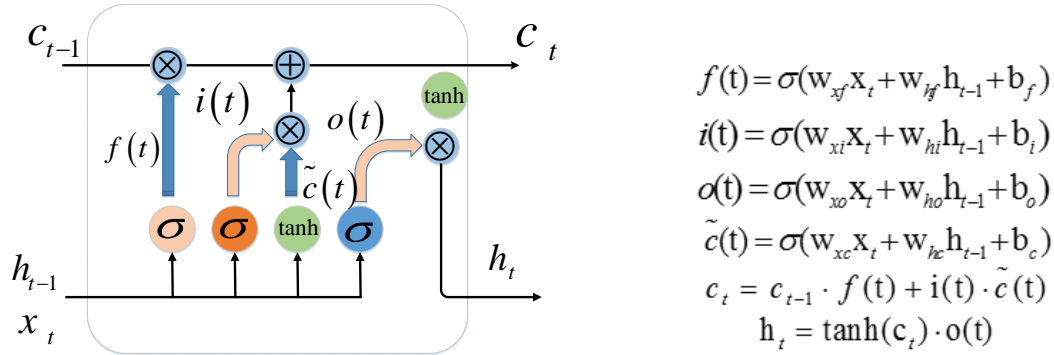


图 3 LSTM 单元结构图

其中  $f(x)$  表示对输入信息的激活值称之为“遗忘门”， $i(t)$  表示对  $t$  时刻的输入信息的激活称之为“输入门”， $o(t)$  表示  $t$  时刻的输出值称之为“输出门”， $c(t)$  表示细胞的状态值，LSTM 每个时间节点对应于一个 LSTM 单元，会相应地输出一个值作为对此前所有时间节点对下一时刻的预测，如公式(11)所示：

$$y_t = f(W_o \cdot h_t) \quad (11)$$

本文的 LSTM 通道中，以最后一个时间节点的输出信息作为 LSTM 的输出，将 LSTM 作为文本另一个特征提取器，将输入向量经过映射和激活函数的处理，获得一组输出，并将该输出进行全连接，最为 LSTMCs 的最终输出。

该模型将上述两个通道的中间层输出向量进行几何连接运算，作为最终该模型的特征融合输出，模型结构如图 2 所示。该输出特征能够将两部分的特征加以融合，综合了两种特征并协同训练。最后一步是运用 Softmax 分类器计算每个输出向量在类别空间中的置信度分布如公式(12)所示。该模型根据输出结果计算残差并修正 CCs 和 LSTMCs 的参数，残差公式如公式(13)所示。

$$P(y_t = c) = \frac{\exp(W_{zc} \cdot z_{t,c} + b_c)}{\sum_{c \in C} \exp(W_{zc} \cdot z_{t,c} + b_c)} \quad (12)$$

$$Loss = -\frac{1}{n} \cdot \sum_{i=1}^n y_i \cdot \log(f(x_i)) + \lambda \sum_{k=1}^L (\|W_k\|^2) \quad (13)$$

其中， $P(y_t = c)$  表示  $y_t$  分为  $c$  类别的置信度概率； $Loss$  表示通过交叉熵计算损

失函数， $y_i$  表示正确标签， $f(x_i)$  表示预测标签。

## 5 数据和实验

### 5.1 数据集

中文针对反讽和讽刺的研究很少，更少有一份完整的训练数据，为了填补这一空缺，方便科研工作者深入对反讽和讽刺的研究，本文采用半自动化的方式对微博、博客中的反讽和讽刺的文本进行收集。本文中所用的数据集主要来自三个方面：

(1) 台湾大学<sup>[8]</sup>针对反讽收集了 950 条实验语料，本文在此基础上加入了一些未涵盖在其中的特殊反讽文本，以扩充语料。

(2) 通过爬虫技术在新浪微博平台抓取了 300W 余条微博<sup>[12]</sup>，Sun<sup>[13][14]</sup>在此语料上进行多次情感分析任务，验证了语料的可靠性。本文通过人工收集的方式从中挑选出讽刺语料 500 条和正常语料 500，使用卷积神经网络 (CNN) + Word2Vec 对讽刺语料进行建模，训练模型并将剩下的微博作为测试集作预测，最终获得预测讽刺语料 10W 条 (本文称为“准讽刺文本”)，为确保语料的可靠性，构建平衡语料，避免对后续实验的干扰，本文加入了人工筛选的过程，对“准讽刺文本”进行多人投票挑选，选择出讽刺语料 1030 条。

(3) 为了方便实验开展，平衡语料数量，在微博、博客中随机获取了 1000 条非反讽、非讽刺文本作为负类，其中包括正向、负向和中立情感的句子，以去除对实验的干扰。

根据以上构建了一个涵盖反讽、讽刺和正常三个类别的实验语料，共计 3030 条 (其中反讽：1000，讽刺：1030，正常：1000)，数据集部分语料示例如表 2 所示。

表 2 数据集部分语料示例

类别	示例
反讽	很好，基金就快跌破 40 大关了。 雨可以再大一点，我真的一点都不介意。 网络可以再慢一点！
讽刺	股票赔的只能扮超人出去打劫了。 国足不是病，踢起来真要命。 不要怕被人利用，人家利用你，说明你还有利用的价值。

部分训练集中样本 2.2 所述不平衡分数 (IS) 分布图：

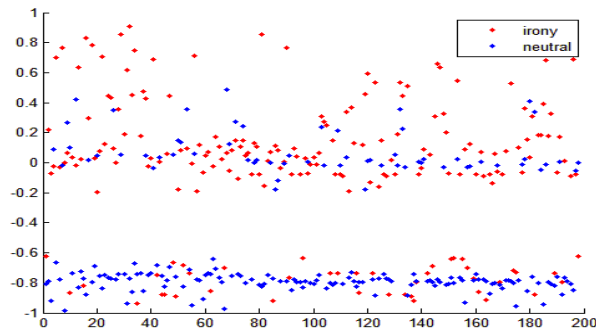


图 4 部分训练集样本不平衡分数分布图

图 4 也可以证实，本文所述语义结构特征和语义不平衡特征的有效性。

### 5.2 实验设置

#### (1) 模型与参数设置

本文主要使用词序随机词向量、Word2Vec<sup>[9]</sup>以及随机词向量的形式作为神经网络模型的初始化特征。其中词向量的训练过程中涉及的算法和参数如表 3 所示：



表 3 词向量训练参数设置

可调参数	参数值
维度 (d)	300
滑动窗口 (window)	10
采样值 (sample)	1e-3
算法选择	Skip-gram

本文对 LSTM 和 CNN 分别进行了实验,并与本文所提出的混合神经网络(LSTMCNN)模型作对比实验,通过多次试验获得最佳的训练参数以及卷积设置如表 4 所示:

表 4 神经网络模型参数设置

	文本维 度	LSTM 隐 含层	CNN 卷积 核	CNN 卷积 窗口	CNN 采样 窗口	optimizer	迭代 次数
CNN	300	-	100	3*3	2*2	SGD	20
LSTM	300	128	-	-	-	Adadelta	20
LSTMCNN	300	128	32	3*3	3*3	Adam	20

## (2) 对比试验

针对反讽、讽刺问题,在上文根据其语言特征加以分析并作出假设,反讽和讽刺现象与其语言结构有关,与句法结构有关,与其深度语义相关。因此设置对比实验,验证假设的正确性。本文选择实验效果最佳的二元词和一元词的作为文本特征,使用 libSVM<sup>3</sup>进行实验。除了浅层的词频特征,本文还提出了基于自动学习的领域特征——具有讽刺、反讽意味的特征模板。考虑到反讽、讽刺现象的识别与语义结构的强烈相关性,本文引入句法结构特征(SF)和语义不平衡特征(IS)。综合上述,本文开展如下八组实验进行三分类,如表 5 所示:

表 5 七组实验设置说明

编号	模型特征	实验说明
1	SVM+BOW	提取二元词和一元词卡方特征, SVM 分类
2	SVM+BOW+SF	实验 1 基础上加入结构特征
3	SVM+BOW+SF+IS	实验 2 基础上加入不平衡分数特征
4	CNN+RandomW2V	CNN 模型引入随机词向量训练分类
5	CNN+Word2Vec	CNN 模型引入 Word2Vec 训练分类
6	LSTM+Word_Index	LSTM 模型引入词序特征训练分类
7	LSTMCNN+W2V	LSTMCNN 双通道均使用 Word2Vec 特征
8	LSTMCNN	融合 5, 6 模型抽取的词序和词向量特征, 混合模型分类

## 6 结果分析

<sup>3</sup> <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

表 6 七组实验指标一览表

实验编号	模型特征	准确率	召回率	F-值
1	SVM+BOW	0.9110	0.9050	0.9070
2	SVM+BOW+SF	0.9347	0.9342	0.9342
3	SVM+BOW+SF+IS	<b>0.9390</b>	<b>0.9372</b>	<b>0.9365</b>
4	CNN+RandomW2V	0.9133	0.9133	0.91
5	CNN+Word2Vec	0.9175	0.9185	0.9182
6	LSTM+Word_Index	0.9340	0.9242	0.9266
7	LSTMCNN+W2V	0.9297	0.9282	0.9286
8	LSTMCNN	<b>0.9419</b>	<b>0.9374</b>	<b>0.9390</b>

本文是针对反讽、讽刺和正常语料进行的三分类实验，考虑试验指标的科学性，本文采用宏平均作为评价指标。分析上述 8 组实验结果，对比实验 1 和 2,3 可以得知，本文引入 2.2 所述的基于文本结构 (SF) 和语义结构(IS)的特征能够在实验准确率等指标上有一定幅度的提升，论证了本文所提出的反讽和讽刺现象与句子结构和语义结构的相关度较高的假设。对比实验 1 和 4、5 可知，CNN 和 LSTM 所提取的深度语义特征模型，相对于浅层特征的分类结果精度较高，表明 CNN 能够很好地学习 Word2Vec 的语义信息，LSTM 模型能够获得句子的更多上下文结构信息，从而实验精度有所提升。对比实验 4 和 6，LSTM 词序随机词向量模型相对 CNN 的随机词向量模型有一定提升，表明 CNN 在获取上下文信息的同时，由于卷积的操作使得词序信息遭到破坏，失去了部分重要的信息，因而实验 6 能够优于实验 4 的结果。对比实验 7 和实验 8，实验 7 使用单纯的 Word2Vec 作为输入特征，实验结果相对于实验 8 稍次，由于 LSTM 存在自编码和解码的机制，对于相对原始的词序信息性能更优，预训练的词向量会导致 LSTM 过多的学习词向量维度间的信息而忽略了词序的信息，从而降低了实验的精度。从实验结果可以清晰得知，本文所提出的混合神经网络模型 LSTMCNN 性能最优，获得 94.19% 的实验准确率，该模型相对于实验 6 能够获得更多卷积通道所提取 Word2Vec 的深度语义信息，相对实验 5 能够获得更多上下文句法结构的信息，并加以有效记忆。因此综合这两者的优势能够获得更好的实验性能。

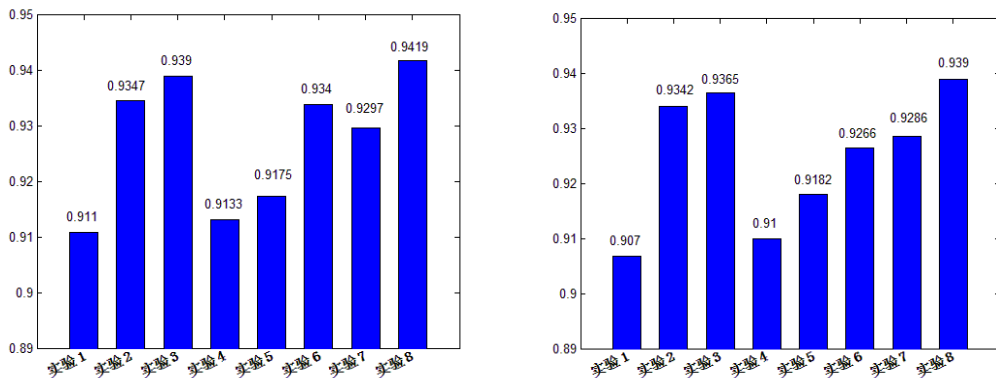


图 6 实验宏平均准确率 (左) 和 F 值 (右) 对比图

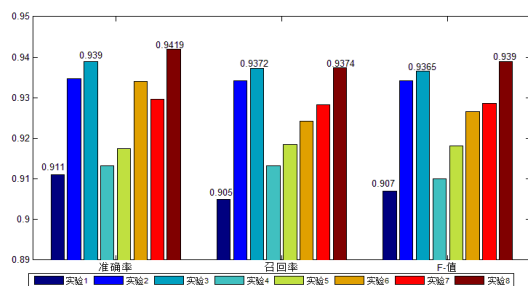


图 7 七组实验指标结果图

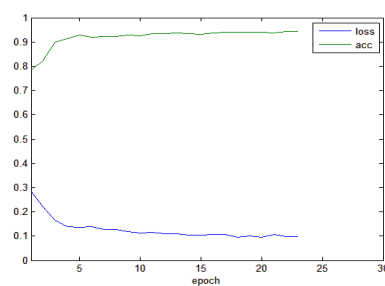


图 8 Loss 和 Acc 变化趋势

由图 6、7 可以直观地看到，本文所提出的模型在三项常规指标中，均表现最优。观察图 8（右）可知，该模型在前五次迭代中，迅速收敛，在第 20 次迭代后，实验准确率 acc 和 loss 基本趋于稳定，获得 94.19% 的实验准确率。

## 7 结论和展望

本文主要研究社交媒体中讽刺语用识别问题，挖掘反讽和讽刺的语言特点，并以问题特征为导向，通过实验探讨了混合神经网络模型对诸如反讽、讽刺此类较难解决的问题的可行性。通过实验发现，本文所提出的多特征融合的混合神经网络模型(LSTMCNN)对于解决反讽和讽刺等上下文信息高度相关的问题，各项指标优于传统的词袋模型和传统的 CNN、LSTM 等单一神经网络模型。

在下一步的工作中，将针对如下问题进行提升，微博、博客等社交媒体的文本中含有较多新词，应加强新词的识别。同时，原始数据中存在大量非中文词汇（部分乱码词，字母等）会干扰分词的精度，进而影响特征提取。此外，应当改进模型的泛化效果，将模型得以迁移，解决其他类似的修辞性语言现象。

## 参考文献

- [1] Konstantin Buschmeier, Philipp Cimiano, Roman Klinger. An impact analysis of features in a classification approach to irony detection in product reviews[C]//Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. 2014: 42-49.
- [2] Edwin Lunando, Ayu Purwarianti. Indonesian social media sentiment analysis with sarcasm detection [C]//Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference on. IEEE, 2013: 195-198.
- [3] David Bamman and Noah A Smith. Contextualized sarcasm detection on twitter[C]//Ninth International AAAI Conference on Web and Social Media. 2015.
- [4] Aditya Joshi, Pushpak Bhattacharyya, Mark James Carman. Automatic Sarcasm Detection: A Survey [J]. arXiv preprint arXiv:1602.03426, 2016.
- [5] Peng Liu, Wei Chen, Gaoyan Ou, et al. 2014.Sarcasm detection in social media based on imbalanced classification[C]//International Conference on Web-Age Information Management. Springer International Publishing, 2014: 459-471.
- [6] Santosh Kumar Bharti, Korra Sathya Babu, Sanjay Kumar Jena. Parsing-based sarcasm sentiment recognition in twitter data[C]//2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2015: 1373-1380.
- [7] Francesco Barbieri, Horacio Saggion, Francesco Ronzano. Modelling sarcasm in twitter, a novel approach[J]. ACL 2014, 2014: 50.
- [8] Yi-jie Tang, Hsin-Hsi Chen. Chinese Irony Corpus Construction and Ironic Structure Analysis[C]//The 25th International Conference on Computational Linguistics: Technical Papers. Dublin, Ireland, 2014: pages 1269-1278.
- [9] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space [J]. arXiv preprint arXiv:1301.3781, 2013.
- [10] Yoon Kim. Convolutional Neural Networks for Sentence Classification[J]. arXiv preprint arXiv:1408.5882, 2014.
- [11] Hochreiter, S., Schmidhuber, J. Long Short-Term Memory[J]. Neural computation, 1997, 9(8): 1735-1780.

- [12] 孙晓, 叶嘉麒, 唐陈意, 任福继. 基于多策略的新浪微博大数据抓取及应用[J].合肥工业大学学报, 2014, 39(10):1210-1215.
- [13] Xiao Sun, Chengcheng Li, Fuji Ren. Sentiment Analysis for Chinese Microblog based on Deep Neural Networks with Convolutional Extension Features[J]. Neurocomputing, 2016.
- [14] Xiao Sun, Jiaqi Ye, Fuji Ren. Detecting Influenza States based on Hybrid Model with Personal Emotional Factors from Social Networks[J]. Neurocomputing, 2016.
- [15] Graves, A., Schmidhuber, J. Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures[J]. Neural Networks, 2005, 18(5): 602-610.
- [16] M. Liwicki, A. Graves, S. Fernández, et al. A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks[C]//Proc. 9th Int. Conf. on Document Analysis and Recognition. 2007, 1: 367-371.
- [17] Andrew Trask, Phil Michalak, John Liu. sense2vec - A Fast and Accurate Method for Word Sense Disambiguation In Neural Word Embeddings[J]. arXiv preprint arXiv:1511.06388, 2015.
- [18] Aniruddha Ghosh, Guofu Li. Semeval-2015 task 11: Sentiment analysis of figurative language in twitter[C]//Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). 2015: 470-478.

## 作者简介



孙晓 (1980-), 男, 副教授, 主要研究领域为自然语言处理 (词法分析, 情感计算, 人机会话)  
情感机器人与人机交互  
Email:suntian@gmail.com



何家劲 (1992-), 男, 硕士, 主要研究方向为自然语言处理 (情感计算, 机器学习)。  
Email:hejudgin@mail.hfut.edu.cn



任福继 (1959-), 男, 教授, 主要研究领域为信号与信息处理, 情感计算  
Email:ren2fuji@gmail.com