

# 面向社交媒体的幽默计算理论与实践

林鸿飞<sup>1</sup> 吐妮可·吐尔逊<sup>1</sup> 张冬瑜<sup>2</sup>

(1 大连理工大学 计算机科学与技术学院, 辽宁 大连 116024; 2 大连理工大学 外国语学院, 辽宁 大连 116024)

**摘要:**幽默是人类独有的语言表达方式,具有幽默感的人可以构建和谐的交际氛围。如何识别社交媒体上的海量幽默文本和评价用户幽默感成为计算语言学领域面临的重大挑战。因此,我们从认知语言学的视角,提出从资源建设、幽默识别到幽默感评价的社交媒体幽默计算框架。首先,参照幽默的语义脚本理论,构建面向社交媒体的中文幽默语料库。然后,利用情感倾向性、语义不一致性及模糊性特征,对文本进行幽默识别。最后,从创作和欣赏两个维度进行用户的幽默感评价。幽默计算将有助于提升语言智能的水平,有助于提高阅读理解、人机对话、机器翻译等领域的应用性能。

**关键词:**幽默计算;社交媒体处理;语料库构建

中图分类号:H087 文献标志码:A 文章编号:1674-6414(2024)04-0045-15

## 0 引言

幽默普遍存在于语言之中,是沟通交流的重要组成部分,也是人类智慧与创造力的结晶。“幽默”一词起源于拉丁语的医学术语,对于幽默的定义存在着不同的视角。西格蒙特·弗洛伊德(Sigmund Freud)(1989)认为幽默是一种强烈冲动或欲望的宣泄方式,霍布斯·托马斯(Hobbes Thomas)(1840)把幽默看作优越感的一种表现形式;《牛津辞典》把幽默解释为令人发笑的品质或者具有发笑的能力。林语堂(1923)最早将英文的“humor”翻译成“幽默”,指出了幽默“戏谑”的特征,并对幽默的文学价值加以提倡。由此可见,“有趣、可笑”可以作为判断幽默的广义标准。然而,对于幽默统一的界定并非易事,其内涵和外延涉及语境、认知、文化等诸多问题。

社交媒体深刻地改变着人们的沟通交流方式,其开放性使得用户得以充分地、随意地

收稿日期:2024-01-15

基金项目:国家自然科学基金项目“面向社交媒体的中文幽默计算研究”(62076046)、“基于认知视角和语义表示的隐喻识别与应用研究”(62076051)、国家语委重点项目“汉语自然口语语料库建设及应用研究”(ZD1145-80)的阶段性成果

作者简介:林鸿飞,男,大连理工大学计算机科学与技术学院教授,博士生导师,主要从事自然语言处理、情感计算、信息检索和文本挖掘研究。

吐妮可·吐尔逊,女,大连理工大学计算机科学与技术学院硕士研究生,主要从事幽默计算研究。

张冬瑜,女,大连理工大学外国语学院教授,博士生导师,主要从事隐喻计算、幽默计算、情感计算研究。

引用格式:林鸿飞,吐妮可·吐尔逊,张冬瑜.面向社交媒体的幽默计算理论与实践[J].外国语文,2024(4):45-59.

表达观点和情感并逐渐形成了一些独特的带有幽默性质的表达方式,它们往往通过多样化的语言表达方式造成诙谐有趣的含义。这些海量的用户生成信息为幽默的分析与研究提供了便利的平台,海量的幽默文本迫切需要能够自动分析和挖掘。因此,如何让计算机识别幽默和理解幽默具有重要的理论和应用价值,它将极大地提升目前的语言智能水平,使之更加接近人类思维和语言习惯,并有助于提升人机交互的性能。

面向社交媒体的幽默计算研究旨在赋予计算机识别和理解幽默的能力,使之能够基于社交媒体的基本特征以及传播属性进行文本的幽默识别和用户的幽默感评价。它涉及认知语言学、计算语言学和人工智能等多个学科的交叉,在机器翻译、人机对话、问答系统、阅读理解等相关领域都具有重要的理论和应用价值。

## 1 国内外研究现状

面向社交媒体的幽默计算目前仍处于起步阶段,它的处理对象是用户和文本,主要任务可以分成三个部分:一是构建幽默语料库,从文本和用户两个维度为幽默计算提供数据支持,标注的语料可以为后续的幽默识别提供训练语料。二是以语义表示为核心,对于社交媒体的文本进行幽默识别,获取文本的幽默属性。三是利用社交媒体中用户与信息的相关特性,获取用户发布的文本的幽默属性以及反馈行为数据,进行用户的幽默感评价。

### 1.1 幽默语料库构建研究

幽默语料库在幽默研究中扮演着关键的角色,优质的语料库可以提高识别的准确性。

早期,幽默数据集主要以文本形式为载体。米哈尔恰等人(Mihalcea et al., 2005)构建了一个包含16 000个单句笑话和16 002个具有相似特征的负例的数据集“16000 One-Liners”。杨等人(Yang et al., 2015)则从双关语网站收集数据,创建了包含2 423个幽默正例和2 403个负例的双关语识别数据集“Pun of The Day”。近期的工作显示,幽默数据集的构建呈现出多样化趋势。哈桑等人(Hasan et al., 2019)从TED演讲中收集数据,创建了多模态幽默语料库UR-FUNNY,从视频中选择了8 257个片段作为幽默正例,并从同一视频中选择了与正例数量相同的负例。受到UR-FUNNY的启发,吴等人(Wu et al., 2021)构建了第一个中文多模态幽默数据集MUMOR。

中文幽默研究刚刚起步,幽默语料库构建研究还相对匮乏。鉴于社交媒体具有更强的时效性和互动性,构建面向社交媒体的幽默语料库势在必行。

### 1.2 幽默识别研究

幽默识别就是赋予计算机能够处理和理解幽默文本的能力,能够识别出幽默文本。幽默常常依赖于多样的语境,有时需要借助外部知识才能全面理解。

基于特征工程的机器学习方法被广泛应用于幽默识别领域,研究主要集中在文本幽默识别上。早期的方法中,人工特征设计是被广泛采用的策略。杨等人(Yang et al., 2015)构造了包括语音、歧义、不一致性和情感特征等在内的多种特征,利用社交媒体上的数据进行幽默识别。樊小超(2021)基于多维潜在语义分析,总结了幽默表达中的不一致、模糊、情感、语音和句法特征,针对不同类别特性设计了多种子特征,用于幽默识别任务。随着技术的发展,深度学习在幽默识别领域的应用逐渐崭露头角。多纳休等人(Donahue et al., 2017)从不同维度提取了数据中潜在的语义和语音特征,并采用多种深度学习融合的方法取得了 SemEval 2016 幽默评测的最优性能。张童越等人(2023)利用卷积神经网络结合标签转移关系提出多任务学习模型识别幽默。

综上,传统机器学习方法主要涉及幽默的表层特征,深度学习方法往往忽视了幽默的认知语言学特性。因此,引入常识作为基础,将幽默的认知语言学特性融入神经网络之中是今后值得深入研究的问题。

### 1.3 幽默感评价研究

人的幽默感与幽默有着密切的联系,是一种特殊的情绪体现。具有幽默感的人说话总是富有趣味,使人发笑。因此,幽默感的强弱是衡量个体素质的一个重要指标。

克里什等人(Kirsh et al., 2003)将幽默感分为三个维度:粗鄙淫秽幽默、社交技巧幽默和矫饰贬抑幽默,包含积极和消极两个方面,因此幽默并不一定只会带来积极影响。马丁等人(Martin et al., 2018)根据应用对象和应用方式两个维度,将幽默分为四种风格,分别为自强型幽默、亲和型幽默、嘲讽型幽默和自贬型幽默。徐海波等人(2013)制定幽默感量表进行分析,发现在幽默感各维度中以幽默来应对心理症状具有显著的负向预测作用。结论表明男生幽默感高于女生,幽默感越高的大学生心理健康水平也越高。

然而,利用人工智能技术进行幽默感的评价目前还处于空白,已有的工作都是在心理量表的基础上进行。因此,基于社交媒体开展用户的幽默感评价是一个崭新的研究视角。

## 2 面向社交媒体的幽默计算研究框架

面向社交媒体的中文幽默计算旨在基于社交媒体平台将用户与发布信息相关联的特性,完成文本的幽默识别以及用户的幽默感评价,即对于社交媒体的两大要素用户和文本进行幽默分析,主要任务包括两个部分,即识别文本的幽默属性和评价用户的幽默感等级。

因此构建的研究体系呈“2-3-2”结构,研究与社交媒体上幽默表达的两个重要因素:文本和用户;对于文本的研究重点是判断文本是否幽默,主要从情感特征、语义不一致性和模糊性等三个方面开展研究。对于用户的研究重点是幽默感评价,用户的幽默感从创作和

欣赏的两个维度进行综合评价。同时,语料库是所有研究基础,它服务于上述两个任务,即幽默识别的语料库和幽默感评价的语料库。

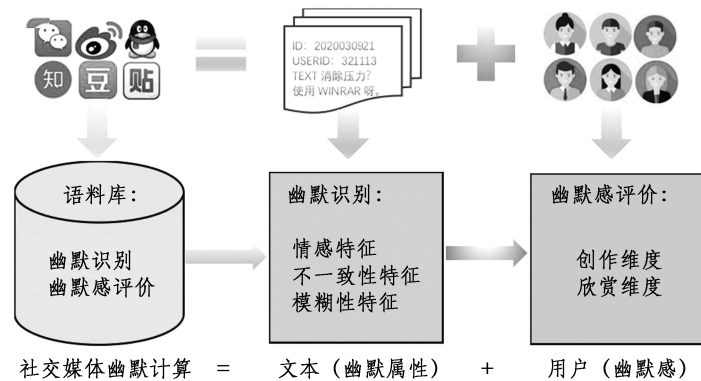


图1 面向社交媒体的幽默计算研究框架

面向社交媒体的幽默计算的主要研究内容如下:

(1) 幽默语料库构建

根据幽默理论和语料库构建原则构建社交媒体幽默表达特征,对幽默语料进行整合、修改、删减等处理,进而建立幽默语料库的标注体系,包括标题、场景、角色、强度、分类、情感和场景关键词等信息。确定幽默的识别依据,制定有效的语料库质量控制机制,保证语料标注的准确性和一致性,供下一步的幽默识别和幽默感评价使用。

(2) 幽默识别

幽默识别对于文本给出是否为幽默的判定,主要是基于文本内容的判定。幽默的产生方式主要在于语义的模糊性和不一致性。因此,幽默识别需要充分利用幽默的表现形式,包括情感、语义不一致性、模糊性特征等,捕获语义的差异程度,获取更丰富的上下文语境信息,使它们在面向社交媒体的中文幽默识别中发挥必要的作用,这是最重要的环节。

(3) 幽默感评价

目前幽默感评价都是基于心理量表的测量方式,这种方式需要人工填写,耗时耗力,无法承担自动评价的任务。因此,依据心理学的幽默感理论,从创作和欣赏两个维度进行幽默感评价。首先对于用户发布的文本和转发的文本以及相应的反馈行为数据进行深入分析,然后从幽默“创作”和“欣赏”两个维度进行综合评价,完成用户幽默感自动评价。

3 面向社交媒体的幽默计算技术实现

3.1 幽默语料库构建

本文构建的幽默语料库包括两个方面,幽默识别语料库和幽默感评价语料库。语料库的构建分为三个阶段:语料采集、语料标注、语料分析,如图2所示。

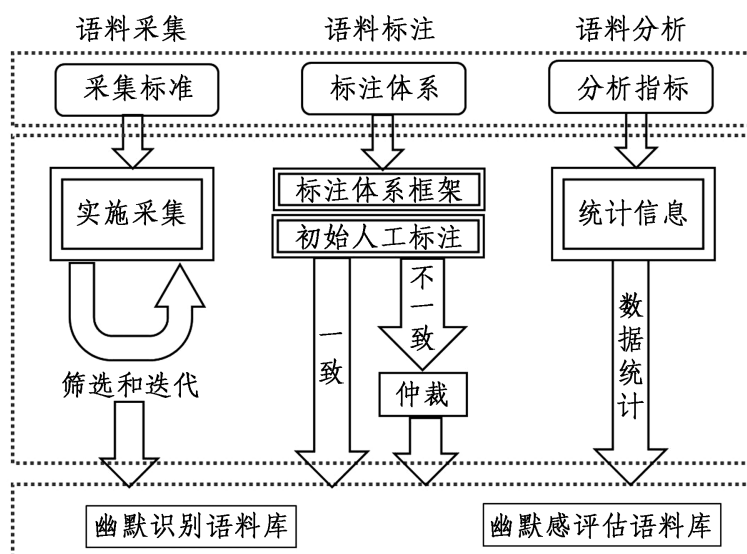


图2 幽默语料库构建流程图

### 3.1.1 语料采集

对于幽默识别,关注的重点是文本是否带有幽默内容;对于幽默感评价,关注的重点是用户的幽默感。所以在数据爬取的过程中,不仅关注幽默文本,还要保留用户信息。本文以常用的社交媒体平台为数据源,从用户节点出发进行数据采集。采集时主要处理方法:

① 采集时使用关键词在社交媒体中进行搜索,例如“笑话”“段子”“搞笑”和“幽默博主”等,获取一系列幽默博主。同时,需要额外筛选一些幽默含量较低的用户,如演员、企业家和数码博主,构成后续训练的负例。

② 需要对爬取的微博进行粗粒度筛选,包括文本长度是否合适,是否有过多的非法字符以及是否为空微博等。

制定以下标准筛选数据,并迭代采集数据,将数据进行过滤和规范化后保存。

① 微博过滤:过滤掉评论少于10条的微博。

② 核心用户过滤:保留发布50条以上符合条件微博的用户。

③ 转发和评论过滤:保留对上述微博进行评论或转发的微博。

④ 普通用户过滤,需要至少评论或转发五个不同用户的微博,这既可以保证用户有一定的参与度,又减少了对特定用户刷活跃度的可能。

### 3.1.2 语料标注

本文构建的幽默识别语料库,旨在应用于社交媒体的幽默识别任务,因此标注对象为微博内容,标注体系内容如表1所示,基本框架如下:

$$\text{HumorText} = (\text{TextID}, [\text{UserID}], \text{Type}, \text{Mark}, \text{Level}, \text{Category}, \text{Class}, \text{Sentiment})$$

在幽默识别语料库的基础上,本文构建了幽默感评价数据集,旨在对文本发布用户进行幽默感评价。通过统计分析样本数据的类别和幽默标签即可得到用户的幽默感评分。

表1 幽默识别语料库

标注项	含义	标注值域
TextID	文本标识符	博文 ID
UserID	用户标识符	用户 ID
Text	文本内容	原始文本
Type	微博类别	评论、转发、原创、个人描述
Mark	是否为幽默文本	0、1
Level	幽默强度	1、3、5
Category	幽默类别	古代、校园、家庭、名人、职场、其他
Class	幽默方式	谐音、谐义、反语、类比、委婉、讽刺、夸张、反转、其他
Sentiment	情感倾向性	-1、0、1

幽默感的评价指标包括幽默欣赏和幽默创作两个维度。对于个体欣赏幽默的能力,本文使用评论微博和转发微博的幽默程度进行评价;对于个体创造幽默的能力,本文使用个人描述、原创微博的幽默程度和得到的好评程度进行评价。

对于用户的幽默感评价主要思想是依据内容幽默程度和受到外界喜欢程度来综合决定,鉴于幽默识别工作的研究进展,对于内容的幽默判定得到基本解决。因此,我们对于行为数据(转发、评论、点赞)进行分析,辅助人工进行用户幽默感的判断。具体统计指标见表2所示,参考这些行为数据的统计指标,再由人工浏览用户的微博,最终标注用户的幽默感,幽默感分为三个级别:低幽默感、一般幽默感、高幽默感。幽默感标注基本框架:

$$HumorSense = (UserID, [Text], HumorSenseLevel)$$

表2 幽默感评价语料库

类目	含义	值域
UserID	用户标识符	
Description	个人描述	文本
DescriptionHumor	个人描述是否幽默	0、1
OriginalNumber	原创微博数量	数值
OriginalHumorNumber	原创幽默数量	数值
OriginalLikeNumber	原创微博点赞数量	数智
ForwardNumber	转发微博数量	数值
ForwardHumorNumber	转发幽默微博数量	数值
ForwardLikeNumber	转发微博点赞数量	数智
CommentNumber	评论微博数量	数值
CommentHumorNumber	评论幽默微博数量	数值
CommentLikeNumber	评论微博点赞数量	数值
HumorsenseLevel	用户幽默感强度	0、1、2

### 3.1.3 语料分析

表3列出了幽默感识别语料库在文章级别上的统计数据,包含了标准训练集、开发集和测试集的分类分布和字符长度信息。可以看出,各集合的数据类别分布相近,属于幽默类别的文章在数量与非幽默类别有一定的差距,并且每条数据的长度略大于非幽默类别。

表3 文章级别统计信息

统计项目	数据集	训练集	验证集	测试集
文章数量	5 291	3 703	794	794
非幽默文章数量	3 281	2 287	499	495
幽默文章数量	2 010	1 416	295	299
字符总数	224 037	157 606	33 337	14 787
非重复字符总数	4 245	3 888	2 536	1727
平均字符数	42.34	42.56	41.99	41.68
非幽默文章平均字符数	41.48	41.65	41.31	40.86
幽默文章平均字符数	43.75	44.04	43.12	43.03

表4列出了幽默感识别语料库的用户级别统计数据。从中可以发现,低幽默感用户相对较多,而一般幽默感和高幽默感用户数量相近。

表4 用户级别统计信息

统计项目	数据集	训练集	验证集	测试集
拥有标签的用户数量	2 151	1 504	322	325
低幽默感用户数量	895	626	134	135
一般幽默感用户数量	594	415	89	90
高幽默感用户数量	662	463	99	100

## 3.2 幽默识别

幽默作为一种具有创造性的表达技巧,识别的难度较大。幽默的特征包括情感、语义不一致性和表达的模糊性等,这些特征将有助于利用算法去理解幽默和识别幽默。

### 3.2.1 幽默文本特征提取

#### 3.2.1.1 情感特征提取

幽默与情感紧密相连。通过运用充满感情色彩的词汇,可以有效地激发读者的情感共鸣,从而实现幽默的效果。然而,幽默内在的情感表达通常较为含蓄,这使得捕捉文本中的情感特征变得困难。为了解决这一问题,本文引入了预训练的常识推理模块 COMET (Bosselut et al., 2019),其基础架构建立在 Transformer 上,并通过在 ATOMIC (Sap et al., 2019)数据集上训练得到,本文主要采用“[xReact]”这一关系,基于上下文信息推断出句子中主语的内在情感状态,并将其以文本形式进行输出。

通过 COMET 对文本序列进行处理,来推断说话者可能的内心情感,我们选取概率最高的情感结果形成情感候选集。将原始的文本序列与情感候选集进行拼接,组成一个显式情感增强的幽默文本序列。随后,使用预训练语言模型(BERT)对显式情感增强的文本序列进行上下文编码。在这一过程中,BERT 不仅能够捕捉上下文信息,还可在语义层面上将幽默文本中的单词与情感候选集中的情感单词关联起来,有效地帮助我们捕捉文本内部的情感特征。为了进一步加深对语义信息的学习,我们采用双向长短时记忆神经网络(Bi-LSTM)和注意力机制构成的编码器  $M_f$ ,更加全面地理解说话者的内心情感状态,使模型能够在文本中更精准地捕捉情感特征,获取潜在的情感特征  $z_e$ 。

### 3.2.1.2 语义不一致性特征提取

已有的研究表明幽默的本质在于表现出两种不一致的思想或概念的冲突,勒夫考特(Lefcourt, 2012)指出:“幽默通常源自两种或更多情况之间的不一致、不协调或是强烈对比。这种表现往往在复杂的上下文中难以得到合理的解释,或者是由于特定的组合方式产生了相互矛盾。”

本文通过 GloVe 词向量对于输入的文本序列进行初始化。对于文本序列中的每个词语,获取其词义特征。将词义特征与语义特征进行拼接,形成更丰富的文本表示。为了更精确地捕捉文本中词语的语义特征,并进一步分析词语级别的语义不一致性,本文引入了一个语义编码器,该编码器由全连接神经网络组成。通过这个语义编码器,我们对拼接后的文本表示进行压缩,得到更紧凑的表示。在压缩后的两个语义信息向量上进行点积运算,生成不一致性矩阵,经过池化后的向量  $z_s$ ,用以描述幽默文本的词语级别的不一致性特征。

### 3.2.1.3 模糊性特征提取

歧义在幽默中属于普遍现象,是有意产生的交流效果。幽默的产生源于单词的多重含义,通过创造语义和语境的歧义,使得句子产生多样的理解。模糊性是幽默中一个至关重要的要素,是语言表达中常见的现象。拉斯金(Raskin, 1979)总结指出:幽默的表达通常伴随着语义的模棱两可,正是这种多义性和模糊性赋予了幽默的表达形式。

本文借助 WordNet 获得同义词集,通过评估每个单词的相对歧义特征来进行模糊性处理。为了更好地了解具有多个含义的单词对幽默识别的影响,我们根据同义词集的数量将单词分为不同的等级类别,并将模糊性类别序列进行 one-hot 表示。为了使模型更好地学习具有多个含义的单词对幽默识别的影响,将同义词集类别表示与文本表示以拼接的方式进行融合。通过引入由双向长短时记忆神经网络和注意力机制实现的模糊特征编码器  $M_f$ ,深入挖掘模糊性特征,并将其融合到文本表示中,最终得到模糊性特征表示  $z_f$ 。

## 3.2.2 幽默识别模型

本文提出的幽默识别模型构造如图 3 所示。该模型分为四部分:(1)情感特征提取单



元,主要是为了获取幽默文本中的隐含情感特征;(2)语义不一致性特征提取单元,主要是为了获取幽默中的上下文语义信息;(3)模糊性特征提取单元,主要是获取含有模糊特征的文本表示;(4)幽默标签预测,结合上述特征,利用深度神经网络进行识别,即给出判定结果,该文本是否属于幽默文本。

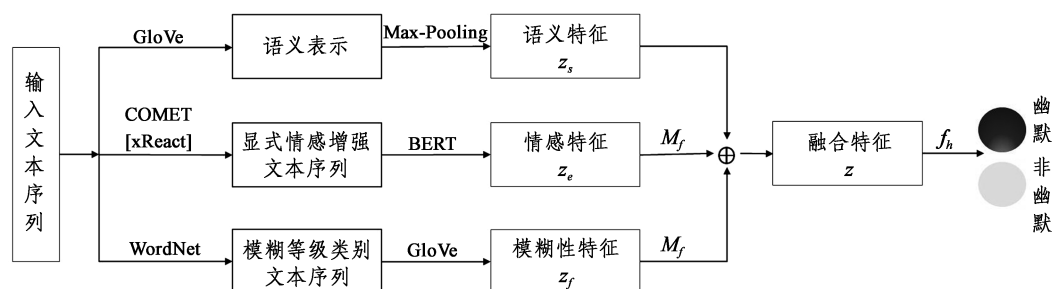


图3 模型框架图

在获得幽默文本的情感特征 $z_e$ 、语义不一致性特征 $z_s$ 以及模糊性特征 $z_f$ 后,将三种特征拼接进行融合。利用注意力机制加强对三种特征内在关系的学习,随后将这些学习到的关联信息输入到以全连接层构建的幽默分类器 $f_h$ 中,获得文本序列的幽默标签预测。同时,采用交叉熵(Cross Entropy)作为损失函数。

### 3.2.3 数据集及实验结果分析

本文采取的数据集有:

Pun of The Day (Yang et al., 2015):该数据集目前广泛用于幽默识别中,其中包括各种类型的幽默,如双关语、笑话、俏皮话等。

SemEval 2021 Task7-1a (Meaney et al., 2021):该任务来自国际语义技术评测 SemEval,本文利用 Task7-1a 的数据进行幽默识别。

ColBERT (Annamoradnejad et al., 2024):该数据集包含 20 万个英文幽默文本,其中从 Reddit 平台收集了 10 万正样本,另外从新闻头条获得 10 万负样本。

实验在 python 3.7 环境下进行,采用 12 层的 BERT-base-cased<sup>①</sup> 作为预训练语言模型编码;语义特征和模糊性特征提取采用 GloVe,维度为 100;使用 WordNet 获取同义词集合;Bi-LSTM 的神经元数量为 128;Dropout 为 0.3;Batch 大小为 64;使用 Adam 优化器。选取准确度 (Accuracy)、精确率 (Precision)、召回率 (Recall) 和 F1 值 (F1-Score) 作为实验评价指标。

本文采用如下基线模型进行对比:

Bi-LSTM (Graves et al., 2013):采用捕捉双向语义依赖关系的 Bi-LSTM 模型进行幽默识别。

<sup>①</sup> <https://huggingface.co/bert-base-cased>.

CNN(Kim, 2014):采用 CNN 获取文本潜在语义及模糊性特征进行幽默识别。

BERT(Devlin et al., 2019):采用 BERT 模型进行幽默识别。

IEANN(Fan et al., 2020):使用情感词典结合内部及外部注意力神经网络获取幽默特征。

ABML(Ren et al., 2021):结合幽默双关语检测的多任务学习模型进行幽默识别。

ANPLS(Ren et al., 2021):融合发音、词汇和句法特征的注意力网络,提取幽默特征进行幽默识别。

表 5 对比实验结果

Dataset	PUN OF THE DAY				SemEval 2021 Task 7-1a				ColBERT			
	ACC	P	R	F1	ACC	P	R	F1	ACC	P	R	F1
Bi-LSTM	86.11	85.13	85.87	85.50	84.90	87.79	87.64	87.71	94.07	94.80	93.19	93.99
CNN	86.42	83.18	91.56	87.17	86.15	87.20	90.32	88.73	94.40	93.18	95.81	94.45
BERT	90.50	88.75	91.80	90.25	91.78	91.66	92.62	92.14	95.55	95.57	95.47	95.52
IEANN	92.24	91.14	92.25	91.69	91.03	91.32	92.10	91.71	94.92	95.33	93.87	94.59
ABML	93.18	92.45	92.07	92.26	92.20	91.92	92.77	92.34	94.42	94.39	94.16	94.27
ANPLS	92.94	93.00	92.55	92.79	92.06	92.38	93.07	92.72	94.39	94.82	95.07	94.94
Ours	94.56	93.47	92.61	93.04	92.15	92.67	93.40	93.03	96.23	95.98	96.40	96.19

实验结果如表 5 所示,对实验结果进行具体分析,得到如下结论:

本文提出的方法在三个数据集上均取得了最好的结果,F1 值相比于现存的最优结果分别提升了 0.25%、0.31%、0.67%,证明了从情感、不一致性以及模糊性三个维度构建幽默语义并应用于幽默识别是有效的。相比于基于 CNN 的幽默识别方法,基于 Transformer 的方法在四项评价指标上有明显提升,这说明其能够通过全局注意力机制更好地捕捉幽默文本的上下文信息。本文设计幽默识别模型能够通过深度神经网络结构,在外部知识驱动下,自动构建幽默特征,相比于 BERT 取得了明显的提升(在 Pun of The Day 数据集上 F1 值提升 2.79%),这也验证了深度学习模型结合外部知识能够在幽默理论约束下学习到幽默相关特征。在与使用情感字典提取文本情感信息的 IEANN 相比,本文所提出模型在三个数据集上 F1 值分别提升 1.35%、1.32%、1.6%,这也表明利用外部常识知识能够更准确推断文本内部情感。

### 3.3 幽默感评价

#### 3.3.1 幽默感的评价维度

幽默感的评价在心理学领域有着较多的研究成果,幽默感的度量理论也有许多。本文采用双维度评价方式,从幽默的创作能力和对于幽默的欣赏能力入手。因此,根据社交媒体的特点,幽默的创作能力主要通过用户发布文本的幽默程度加以度量,而幽默的欣赏能

力则从用户对于其他人创作的幽默文本的态度来加以度量,例如可以考虑被用户转发、评论和点赞文本的幽默程度。综合考虑社交媒体上用户与文本的关系,对幽默感的评价将从“创作”和“欣赏”两个维度来衡量。

### (1) 幽默创作维度

创作是指对文学作品的创造,幽默创作能力就是指用户创作幽默文本的能力。因此,幽默创作能力可以通过用户原创文本中幽默文本比率来度量,幽默文本占比高,得到的转发、评论和点赞数量多,则幽默创作能力较强。因此,对“创作”维度的衡量指标从内容和行为两个角度出发,内容特征是指用户原创的幽默文本占比、幽默文本的笑点、强度、情感倾向性等。行为特征是指原创文本的被转发数,正向评论与负向评论的比率以及点赞数量等。

### (2) 幽默欣赏维度

欣赏是指享受美好的事物,而幽默欣赏能力是用户对于他人创作的幽默作品的欣赏能力。因此,幽默欣赏能力的度量来自用户转发、评论或者点赞的(其他作者的)文本以及得到他人的评价行为。显然,如果转发、评论和点赞的文本中幽默文本占比较高,则幽默的欣赏能力较强。对“欣赏”维度的衡量也是从内容和行为两个角度出发,内容特征是指用户转发、评论或者点赞的(其他作者的)幽默文本占比数、幽默文本的幽默强度、幽默的情感倾向性等。行为特征指转发幽默文本而得到的被转发数、正负向评论的比率以及点赞数量。

## 3.3.2 幽默感评价模型

根据上述关于幽默感的维度定义和特征描述,从社交媒体的内容和行为数据中提取各类特征,本文提出了一种基于融合多维特征的集成分类模型对用户幽默感的两个维度进行等级划分,具体模型如图4所示。

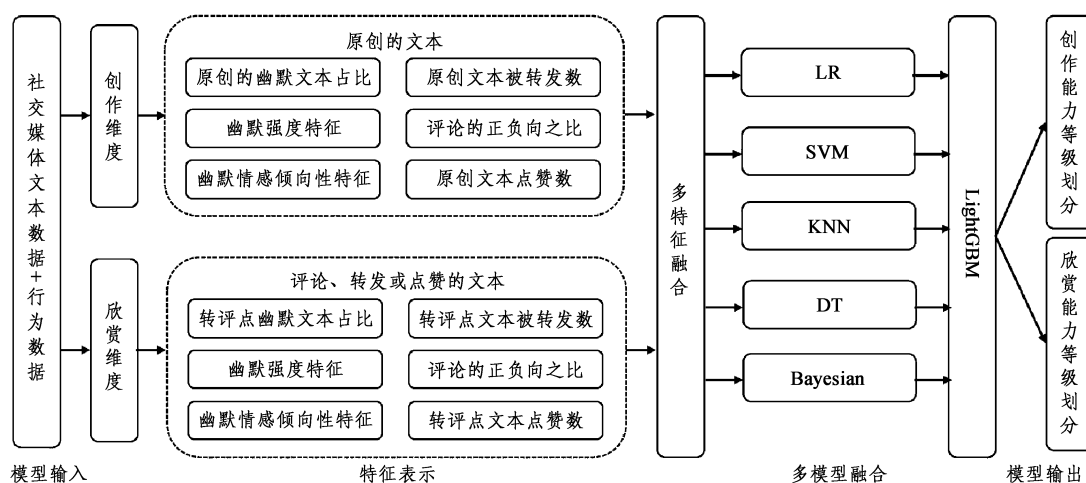


图4 用户幽默创作能力和欣赏能力评价模型

首先获取用户的全部文本信息(包括文字内容、评论、转发和点赞),分别获取基于原创文本的幽默文本占比、幽默强度特征、幽默情感倾向性特征和原创文本被转发数、评论的情感正负向之比以及被点赞数作为创作维度的输入特征;获取基于转发、评论或点赞文本的幽默文本占比、幽默强度特征、幽默情感倾向性特征和转发、评论或点赞文本被转发数、评论的情感正负向之比以及被点赞数作为欣赏维度的输入特征。其次,将六类特征融合;然后,采用线性逻辑回归(LR)、支持向量机(SVM)、K最近邻(KNN)、决策树(DT)、贝叶斯模型(Bayesian)等多个分类器;最后,将上述多个分类器的分类结果作为特征,利用LightGBM(Light Gradient Boosting Machine)做二次分类,能够进一步融合多个基学习器以提升模型的准确性与泛化能力。最终根据分类结果,实现对用户的创作能力和欣赏能力两方面的评估。

结合对用户的创作能力和欣赏能力两方面的评估结果,同时重点考虑幽默文本内容对幽默感评价的重要性,构建幽默感评价模型。根据特征权重不同,对影响幽默感的因素进行不同粒度的划分,因此,采用多粒度注意力幽默感综合评价模型方法对用户幽默感进行评价。首先,将幽默感评级因素分为三个层次:一是目标层,为用户的幽默感评价;二是特征层,分为幽默文本特征和幽默能力特征两个方面的主要因素;三是指标层,是上述两个方面因素的具体指标;建立层次的指标体系,具体的模型如图5所示。

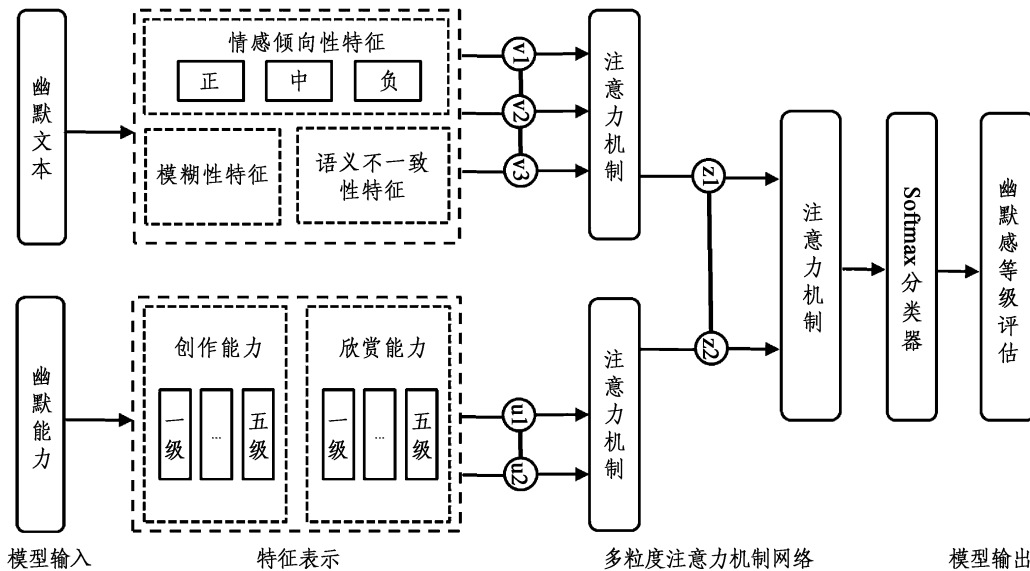


图5 幽默感综合评价模型

利用幽默文本的情感倾向特征、模糊性特征和语义不一致性特征  $v_1$ 、 $v_2$ 、 $v_3$ ;同时将幽默能力特征  $u_1$ 、 $u_2$  转为 One-Hot 形式,与幽默句子特征共同作为模型的输入。然后通过第一层注意力(attention)机制,学习得到特征层内部各指标的权重,获得基于不同指标权重表示的因素表示  $z_1$ 、 $z_2$ ;然后再通过第二层注意力机制,动态地为每个因素学习各自的权重,

将各个因素按权重进行融合,作为最终目标的幽默感等级评估的依据。

### 3.3.3 数据集和实验结果分析

选择在前面构建的幽默感评价语料库作为数据集,标注时将幽默感划分为三个等级,所以幽默感评价归结为三分类任务。由于目前国内外有关幽默感评价的研究近乎空白,因此基线模型选择了适合于分类任务的主流模型,包括机器学习算法 SVM(Noble 2006)、深度学习算法 CNN(Kim, 2014)、Bi-LSTM(Graves et al., 2013)以及预训练模型 BERT(Devlin et al., 2019),其中 SVM 使用用户的非文本特征和文本特征作为分类器输入,其中文本内容为用户的个人描述信息,使用 GloVe(Pennington et al., 2014)提取特征;BERT 使用用户的个人描述信息作为输入对 BERT 进行微调。评价指标选择了精确率(Precision)、召回率(Recall)和 F1 值(F1-Score),实验结果如表 6 所示。

表 6 幽默感评价模型的性能对比

模型	P	R	F1
SVM	47.69	33.70	39.49
CNN	49.27	37.49	42.58
Bi-LSTM	50.46	38.43	43.63
BERT	52.62	40.64	45.86
Ours	60.63	46.04	52.34

实验结果显示我们的方法优于目前主流的一些分类算法。基于用户描述利用 SVM 或者 BERT 进行幽默感评价效果不理想,说明用户的个人描述不能很好地反映一个人是否具有幽默感。CNN 与 Bi-LSTM 相对于 SVM 能够学习到更多的特征,但是提升的范围有限。借助基于社交媒体的文本以及反馈行为数据,利用幽默文本的情感特征,从创作和欣赏两个维度进行幽默感的综合评价,取得了较好的效果,达到了对用户的幽默感进行自动化评价的目的。

## 4 结语

本文旨在基于认知语言学理论构建大规模中文幽默语料库,并以此为基础开展幽默计算的相关研究。首先,参照幽默的语义脚本理论,构建面向社交媒体的中文幽默语料库。然后,在语义表示基础上进一步结合情感倾向性、语义不一致性及模糊性特征,对于文本进行幽默识别。最后,在获取文本的幽默属性以及行为数据基础上,从创作和欣赏两个维度进行用户的幽默感评价。这些来自社交媒体的特征和属性为幽默计算提供了丰富的信息,便于从认知语言学的角度更加深入地挖掘幽默表达的方式和评价人的幽默感。因此,面向社交媒体的幽默计算是可行的、高效的、且具有较高的理论和应用价值。

多媒体技术的发展使得意义建构越来越依赖于非语言模态的参与,因此多模态幽默的理解成为一项富有挑战性的任务。同时,随着大模型的发展,语言生成成为常态。关于大模型生成内容的幽默识别以及与人类价值观对齐研究将成为研究的重点,这些研究将有助于提升计算机理解人类的深层语义的能力,创造更加和谐的人机环境。

#### 参考文献:

- Annamoradnejad, I. & G. Zoghi. 2024. ColBERT: Using BERT Sentence Embedding in Parallel Neural Networks for Computational Humor[J]. *Expert Systems with Applications* 249.
- Bosselut, A., Rashkin, H., Sap, M., Malaviya, C., Celikyilmaz, A. & Y. Choi. 2019. COMET: Commonsense Transformers for Automatic Knowledge Graph Construction[G]// A. Korhonen. *Proceedings of the 57<sup>th</sup> Annual Meeting of the Association for Computational Linguistics*. Florence: Association for Computational Linguistics, 4762-4779.
- Devlin, J., Chang, M. W., Lee, K. & K. Toutanova. 2019. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]. *Proceedings of NAACL-HLT*. USA: Association for Computational Linguistics, 4171-4186.
- Donahue, D., Romanov, A. & A. Rumshisky. 2017. HumorHawk at SemEval-2017 Task 6: Mixing Meaning and Sound for Humor Recognition[C]. *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. Vancouver, Canada: Association for Computational Linguistics, 98-102.
- Fan, X., Lin, H. et al. 2020. Humor Detection via an Internal and External Neural Network [J]. *Neurocomputing* 394: 105-111.
- Graves, A., Mohamed, A. & G. Hinton. 2013. Speech Recognition with Deep Recurrent Neural Networks[G]// 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing*. Canada: IEEE, 6645-6649.
- Hasan, M. K., Rahman, W. & Z. Bagher. et al. 2019. UR-FUNNY: A Multimodal Language Dataset for Understanding Humor [G]// K. Inui & J. Jiang. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9<sup>th</sup> International Joint Conference on Natural Language Processing*. Hong Kong, China: Association for Computational Linguistics, 2046-2056.
- Hobbes, T. 1840. *Human Nature in English Works*[M]. Molesworth.
- Kim, Y. 2014. Convolutional Neural Networks for Sentence Classification[C]. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar:ACL, 1746-1751.
- Kirsh, G. A. & N. Kuiper. A. 2003. Positive and Negative Aspects of Sense of Humor: Associations with the Constructs of Individualism and Relatedness[J]. *Humor -International Journal of Humor Research* (1):33-62.
- Lefcourt, H. M. 2012. *Humor: The Psychology of Living Buoyantly*[M]. Berlin: Springer Science & Business Media.
- Li, J., Zhao, M. et al. 2023. This Joke is [MASK]: Recognizing Humor and Offense with Prompting[J]. *Transfer Learning for Natural Language Processing Workshop*. PMLR (2023):1-9.
- Martin, R. A. & T. Ford. 2018. *The Psychology of Humor: An Integrative Approach* [M]. London & San Diego: Academic press.
- Meaney, J. A., Wilson, S., Chiruzzo, L., Lopez, A. & W. Magdy. 2021 SemEval 2021Task 7: HaHackathon, Detecting and Rating Humor and Offense [G]// A. Palmer. *Proceedings of the 15<sup>th</sup> International Workshop on Semantic Evaluation (SemEval-2021)*. Thailand: Association for Computational Linguistics, 105-119.
- Mihalcea, R., Strapparava, C. et al. 2010. Computational Models for Incongruity Detection in Humour[J]. *Int Scholarly Works* (4):364-374.
- Noble, W. S. 2006. WhatIs a Support Vector Machine? [J]. *Nature biotechnology* (12):1565-1567.
- Pennington, J., Socher, R. & C. Manning. 2014. Glove: Global Vectors for Word Representation [G]// A. Moschitti.

- Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, 1532-1543.
- Raskin, V. 1979. *Semantic Mechanisms of Humor* [M]. Dordrecht: Springer Science & Business Media.
- Ren, L., Xu, B. et al. 2021. ABML: Attention-based Multi-task Learning for Jointly Humor Recognition and Pun Detection [J]. *Soft Computing* (22): 14109-14118.
- Ren, L., Xu, B. et al. 2021. An Attention Network via Pronunciation, Lexicon and Syntax for Humor Recognition [J]. *Applied Intelligence* (3):2690-2702.
- Sap, M., Le Bras, R., Allaway, E., Bhagavatula, C., Lourie, N., Rashkin, H. & Y. Choi. 2019. Atomic: An Atlas of Machine Commonsense for If-then Reasoning [C]. *Proceedings of the AAAI Conference on Artificial Intelligence*. Honolulu, Hawaii, USA: AAAI Press, 3027-3035.
- Wu, J. & Lin, H. et al. 2021. MUMOR: A Multimodal Dataset for Humor Detection in Conversations [J]. *Natural Language Processing and Chinese Computing* (1): 619-627.
- Yang, D., Lavie, A. & C. Dyer. 2015. Humor Recognition and Humor Anchor Extraction [C] // J. Su & D. Pighin. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal: The Association for Computational Linguistics, 2367.
- Zhang, R. & N. Liu. 2014. Recognizing Humor on Twitter [G] // J. Li. *Proceedings of the 23<sup>rd</sup> ACM International Conference on Conference on Information and Knowledge Management*. Shanghai, China: ACM, 889-898.
- 樊小超, 杨亮, 林鸿飞等. 2021. 基于多维潜在语义特征的幽默识别[J]. 中文信息学报(8):38-46.
- 林语堂. 1923. 征译散文并提倡幽默[N]. 北京晨报副刊, 1923-05-23.
- 西格蒙特·弗洛伊德. 1989. 机智及其与无意识的关系[M]. 张增武, 闰广林, 译. 上海: 上海社会科学院出版社.
- 徐海波, 张莹瑞, 刘红梅. 2013. 大学生的幽默感与心理健康的关系[J]. 中国健康心理学杂志(1):150-152.
- 张童越, 张绍武, 林鸿飞, 等. 2023. 结合标签转移关系的多任务笑点识别方法[J]. 中文信息学报(11):142-150.

## Theory and Practice of Humor Computing for Social Media

LIN Hongfei TUERXUN Tunike ZHANG Dongyu

**Abstract:** Humor is a unique language expression of human beings, and people with a sense of humor can build a harmonious communication atmosphere. How to identify massive humor texts on social media and evaluate humor sense of users has become a major challenge in the field of computational linguistics. Therefore, from the perspective of cognitive linguistics, we propose a social media humor computing framework that covers resource construction, humor recognition, and humor sense evaluation. Firstly, in light of the semantic script theory of humor, we construct a Chinese humor corpus for social media. Then, based on semantic representation, combined with sentimental tendencies, semantic inconsistency, and fuzziness features, we recognize the humor on the text. Finally, based on obtaining the humor attributes and behavioral data of social media, we evaluate the user's sense of humor from the dimensions of creation and appreciation. Humor computing will help improve the level of language intelligence and enhance the application performance in areas such as reading comprehension, human-machine dialogue, and machine translation.

**Key words:** humor computing, social media processing, corpus construction

责任编辑: 蒋勇军