

# 基于深度学习的情感分类模型构建与在社交文本中的应用

刘涛

西安思源学院，陕西省西安市，710038；

**摘要：**社交媒体上的文章的迅速增多也带来内容十分丰富、涵盖面十分广泛的多种不同情感表达。为了加强对这些非规范词汇和情感信息的表达识别，提出运用 BERT 和卷积神经网络 (CNN) 构建的集成式分类模型，运用注意力机制与焦点损失机制来强调关注相对较为重要的情感词信息，并且修正各类别样本数据分布不平衡现象，在各种社交文本数据中做了大量的实验结果表明，这种分类方法不仅分类精度较高且稳定可靠，具有显著的应用价值，有一定的通用性。

**关键词：**深度学习；情感分类；社交文本；预训练语言模型

DOI:10.69979/3041-0673.25.02.098

## 引言

情感分析作为 NLP 的重要分支之一，被广泛应用于客户服务指数统计、智能客服、建议反馈等应用场景。随着社交媒体的泛用以及 emoji 表情的出现，文本的表达愈加简洁、随意和模糊，在这种情况下，传统的文本情感分析算法无法有效处理。基于深度学习的情感分析算法模型，基于 BERT 的预训练模型，FocalLoss 算法，多属性融合方法都大大提高了对情感极性差的分类效率，在非结构化数据和真实数据集合上均有着出众的分类精度，具有高实践价值和高商业推广价值。

## 1 情感分类模型构建的理论概述

情感分析是自然语言处理的重要研究方向之一，已经广泛应用于计算客户满意度指标、智能客服以及提供建议反馈等。随着社交网络的普及与 Emoji 表情的应用，文本的表达也越来越简洁、自由和模糊，语义的结构也在逐渐变得松散，与此同时，人们对于情绪的表达方式也在增多。在这种情况下，传统的手动的文本情感识别技术很难应对多变、跳跃和不规范的文本环境，这导致其分类准确率较低，且难以适应现实世界中丰富多样、不具规则性的感情表达及其非线性特点。

## 2 情感分类模型在社交文本中面临的问题

### 2.1 文本碎片化与表达非规范

微博文本作为文章，在语式上具有碎片化特点，其以简洁、跳跃的方式表达，大量出现缩写字、拼音写法、网络缩略词及表情符号等非规范格式，如“笑晕了”“没关系”，部分文章缺乏规范的语法结构，其意义更多依

赖上下文理解或语境感知，现有情感分析算法难以明确其情感倾向。此外，部分词语在上下文不同环境中有截然相反的含义，也加剧了区分复杂性。这种多元化表达以及非规范格式对于情感分类任务的语义构建能力都提出了更高的要求，这也是导致分类准确率下降的关键要素。

### 2.2 情绪反转与语义讽刺难以识别

社交文字中常常出现情绪转变、意义讥讽的情况，如“你们的服务态度太好了，但是却没有人来帮助我”。这句词语义正面的评价却蕴藏着负向的语气情绪，传统的机器学习方法因为缺乏对语义上下文的考察能力，由于片面关注词本身的语法而不可以识别整句话所蕴含的情绪，而导致分类结果的低精准率。对于带隐藏意义、包含有比较关系的讥讽口语则需要对语境进行深层认知能力以及较强的语义相关性识别能力。

### 2.3 类别不均衡造成模型偏倚

在进行情感分析的任务中，常常会出现样本不平衡，即正向情感样本通常大于负向或者中性情感样本的数量。因此机器更有可能关注在主流情感的提取上，而忽略了辅助情感的提取，导致情感分析结果片面化。该模型的宏观精度得到了良好体现，但在 F1 值和召回率等方面出现了明显的下降，难以对各情绪类型进行有效的捕捉，降低了实际使用效果。

### 2.4 模型缺乏上下文建模能力

传统的情绪识别方法如基于 bag-of-words 或普通 CNN，仅依赖词或短语本身的信息，难以捕捉长距离依

赖与上下文关系。在文本存在情绪反转、语义扩展或多段落结构时，模型易受局部词语影响，出现误判。例如，首句含否定表达时，可能导致整体情感方向判断错误。这类方法在真实复杂语境下表现不稳定，难以准确还原文本的完整情绪意图。

### 3 基于深度学习的情感分类模型构建策略与应用

#### 3.1 BERT 嵌入处理非规范文本表达

社交媒体中的文本往往含有大量的非规范化用语、简体字、表情符号、网络热词、拼写错误等。这种特点是对传统情感计算方式的巨大挑战，特别是只使用静态词汇向量的模型无法意识到细微差别，可能无法正确捕捉上下文关系，从而造成情感倾向性判断的错误。BERT 成功地利用双向 Transformer 的结构解决了这一问题，与其它形式相比，它可以充分获取整个上下文信息，并且根据其提供的词向量更能满足语境感知的需要。其输入形式统一为：

Input = [CLS]Token1Token2...Token<sub>n</sub>[SEP]

BERT 通过特定的 CLS 标记进行句子级的分类学习，并通过 SEP 来进行句子间的分隔。该模型使用了 WordPiece 的方法对诸如“zqsg”等网络词汇进行了拆分，能够更好地捕获非正规的表达形式。基于自注意力机制，BERT 为各个词语分配不同的权重系数，以便对词句情绪变化以及细微表情捕捉的更细颗粒度。例如，对于“我好想笑哭哈哈[流泪]”这样情绪丰富、复杂的词句表达，也能正确地对“笑哭”“哈哈”这样的情绪进行融合识别。经过训练后，BERT 在对情绪强度敏感性和分辨率的表达上更加精准，这是 BERT 应对网络语言复杂性灵活、稳定、广适的基础语义支撑，为后续的情感识别模型奠定坚实基础。

#### 3.2 注意力机制增强反讽识别

情感表达在社交文本中往往不是直接呈现出来的，而是通过讽刺、反语等方式，通过文字表面含义及其真实含义间极大的差距来隐喻表达出真实的情感。例如，“真的很佩服他天天加”，从语义上看这句是对他的勤奋工作的赞赏，但从本意上看是无奈的吐槽。这种情感突变现象一般需借助上下文语言上的反常矛盾、标点符号或者固定搭配才能隐藏起来，这就导致了传统文本情感分析模型难以对其做出明确的准确判断。对于缺乏对

上下文建模能力的模型，它们对于关键单词进行判断，假如只是抓住“羡慕”这个字眼，很容易将其划分为积极评价类，而忽视了“天天加班”的消极含义。上述现象的发生是因为这些模型不能理解句子中的语义张力，在社交网络中的文本由于口语的简洁，经常省略标点符号，也会伴随语气的变化，造成文本建模的难度上升。使用注意力机制可以进行有效弥补，基于输入的不同部分之间对输入的关注程度来引导模型只关注与表达关键意思相对应的词汇或短语，从而实现语义建模的精细化。其基本运算方式为：

$$\alpha_i = \frac{\exp(score(h_i, q))}{\sum_j^n \exp(score(h_j, q))}$$

利用此模型，模型可以更加重视“加班”和“羡慕”之间存在一个较大的意义反差，它挖掘出了二者暗含的对立关系，并且正确地鉴别出真实情感为消极。在实际应用中，如句子其中  $h_i$  为第  $i$  个词的隐藏状态， $q$  为当前查询向量， $\alpha_i$  为该词的重要性权重。“你真好，你真是把项目毁了！”通过注意力机制加强对于“毁了”的关注程度和减少“你真好”这个褒义词的影响，使得模型根据整句话的语义场景理解该短句表达的意思是一个嘲讽。这样使得模型更加敏感语义反转和讽刺词语，显著提升了复杂情绪语料的识别准确率。

#### 3.3 引入 Focal Loss 缓解类别不平衡

在社交媒体中，为了提高情绪分类系统对稀有情绪类别的分类能力，可以考虑通过在训练中的运用 Focal Loss 进行改进。该方法依据相应的条件改变样本损失权重，使系统更加关注难以区别的样本，进而大幅度提升了对边界情绪类别的处理能力。其中，易被识别为主类别的样本对损失造成的影响减小，帮助模型更加学懂稀有类别，提高了总的分级能力。focalloss 对实际文本的适用性在于其训练过程中通过控制焦点的分布方式优化了模型处理多元情绪表达和语料风格的社交文本的能力，实现了增强其识别稳定性和精准度的目标，为开发情绪识别系统提出可行方案。动态调整损失响应：

$$FL(p_t) = -\alpha_t (1-p_t)^\gamma \log(p_t)$$

其中， $p_t$  表示模型对真实类别的预测概率， $\alpha_t$  为平衡因子， $\gamma$  为聚焦参数，用以增强模型对低置信度样本的响应。当模型对主类预测非常确定时，即  $p_t \rightarrow 1$ ，该项损失趋近于 0，若不对损失函数进行优化调整，模型在训练过程中对小类样本的关注度较低，难以有效学

习其特征，导致识别能力不足。例如，在微表情识别任务中，“真羡慕你天天加班”这类带有讽刺意味的句子应被归类为讽刺情绪，但该类别在整体语料中占比极低，错误分类所产生的梯度不足以驱动模型调整。引入 Focal Loss 后，此类小类错分会被赋予更高的训练权重，显著增强模型对少数类的学习能力，从而逐步掌握讽刺类表达的特征。该方法不仅提升了模型在整体样本上的 F1 指标，也增强了对不均衡数据的泛化能力，有效改善了复杂语境中微妙情绪表达的识别表现。

### 3.4 CNN 模块加强局部情绪表达建模

传统的情感分析技术在处理文本帖子时，因缺乏对上下文以及局部情绪元素的考虑，在理解过程中经常出现认知偏差的现象。尤其是在一些简短的段落中或者是语义十分简约的文字中，这种类型的系统难以从中识别关键情感元素，因此更有可能采取仅依据字词自身这一规则而非词语的上下文和搭配去理解文字。例如“真得无话可说这个桥段”中，由于“真是”这个词被错误地解读为褒义词，然而“无话可说”才是关键情绪元素，这种场景加大了机器学习系统在建立过程中的挑战。因此借助 CNN 对于文本中理解上的问题（即上下文缺少的问题），以及在提取局部情绪元素过程中面临困难，CNN 在图像识别的过程中通过提取局部特点识别区域这种功能也可以被用在文本的 ngram 情绪元素的识别上。模型是通过滑动窗口对输入的词向量序列做卷积运算。公式如下：

$$c_i = f(W \cdot x_{i:i+k-1} + b)$$

其中， $x_{i:i+k-1}$  表示从位置  $i$  开始的连续  $k$  个词向量组成的局部片段， $W$  为卷积核， $b$  为偏置项， $f$  为非线性激活函数，如 ReLU。每个  $c_i$  表示该窗口下的局部情绪特征。再以最大池化选取最大情绪信号做分类处理。“我现在真的爆炸了”是网络用语，表明一种强烈的情感表现，如愤怒、兴奋等情绪，通过 CNN 提取的关键词，如“爆炸了”，可与其他词语构成语义更丰富、情绪更强烈的表达，例如“真的”用于加强语气，“现在”指示时间背景。相较于均值或全局表示，CNN 能够保留更细致的局部信息，有助于捕捉情绪突变中所蕴含的复杂语义。结合已训练好的语言模型（如 BERT）提供的上下文

嵌入，再在输出端引入 CNN 模块，使模型在处理情感变化表达时更具精度与敏感度。实验表明，该结构在面对大量内容短小、情绪密集的社交媒体文本时表现优异，有助于更深入理解人类情绪表达特征，提升情感识别的准确性与决策支持能力。

### 4 结语

本文旨在研究如何将深度学习模型应用到情感分类的任务中，提出基于 BERT Embedding、Attention、Focal Loss 和 CNN 模块的复合式模型，此策略有效地缓解了社交媒体文本中的非标准化表达、情感表达方向和标注不平衡的问题。实验证明本模型在精度、稳定性和通用性方面得到了较大提高，具有很高的实用价值和重大影响。未来工作可以引入更多的信息（例如图像、音频）和人的行为特征，以辅助本模型更深刻地感知人的主观状态，使模型能够应用于舆论监测、智能推荐和公众安全告警等各种复杂场景的智能化程度和通用性。

### 参考文献

- [1] 贺加贝, 周菊香, 甘健侯, 等. 基于多任务学习的课堂表情分类模型[J]. 应用科学学报, 2024, 42(6): 947-961.
- [2] 张宇, 吴静. 基于深度学习的商品评论情感分析[J]. 智能计算机与应用, 2024, 14(8): 54-58.
- [3] 王浩畅, 王宇坤, Marius Gabriel Petrescu. 融合情感词典与深度学习的文本情感分析研究[J]. 计算机与数字工程, 2024, 52(2): 451-455.
- [4] 刘保旗, 林丽. 基于残差网络的汽车前脸情感化图像分类模型[J]. 智能计算机与应用, 2024, 14(3): 192-198.
- [5] 季玉文, 陈哲. 基于 BERT 的金融文本情感分析与应用[J]. 软件工程, 2023, 26(11): 33-38.
- [6] 刘星, 杨波, 郁云. 一种基于 LSTM 和 ResNet 网络的情感极性分析方法[J]. 电子器件, 2023, 46(6): 1629-1633.

作者简介：刘涛（2004.3.7-），男，汉族，北京市，本科，职称无，研究方向暂无。西安思源学院，陕西省西安市，710038