

# 融合情绪知识的案件微博评论情绪分类

郭贤伟 赖 华 余正涛 高盛祥 相 艳

(昆明理工大学信息工程与自动化学院 昆明 650500)

(昆明理工大学云南省人工智能实验室 昆明 650500)

**摘 要** 案件微博评论的情绪分类是一个特定领域的情感多分类任务,旨在快速有效地识别海量评论中的情绪,有助于相关部门及时评估舆情风险和制定相关政策.由于传统方法难以有效利用评论中常用的情绪词和表情符号等情绪知识,本文提出一种融合情绪知识的案件微博评论情绪分类方法.首先,整合现有的情感计算资源构建了一个包含案件微博情绪词典、表情符号、网络用语、否定词及程度副词等的情绪知识库.其次,考虑情绪知识库和词性的作用定义了15种情绪知识,通过提出的连续向量表示方法构建评论的情绪知识表示.然后将评论的语义表示和情绪知识表示分别输入一个语义初始滤波器的卷积神经网络(INIT-CNN)和一个结合注意力机制的全连接网络中,得到深层的语义特征向量和情绪知识表示的注意力特征向量.最后,将两个特征向量进行拼接以融合语义特征和情绪知识特征,训练一个情绪分类模型,称为EK-INIT-CNN(Emotional knowledge enhanced INIT-CNN).在案件微博评论数据集上的实验表明,相比INIT-CNN,EK-INIT-CNN的 *Macro\_Precision*、*Macro\_Recall* 和 *Macro\_F1* 指标分别提升了1.87%、1.95%和1.88%,EK-INIT-CNN在NLPCC中文微博情绪分析评测数据集上的性能则超过了目前已知文献中的最好结果.实验证明,该方法能有效地融入外部情绪知识,且相对传统方法在情绪分类任务上具有明显的优势.

**关键词** 案件微博评论;情绪知识表示;卷积神经网络;注意力机制;情绪分类

中图法分类号 TP18 DOI号 10.11897/SP.J.1016.2021.00564

## Emotion Classification of Case-Related Microblog Comments Integrating Emotional Knowledge

GUO Xian-Wei LAI Hua YU Zheng-Tao GAO Sheng-Xiang XIANG Yan

(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500)

(Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500)

**Abstract** Currently, social media platforms represented by Weibo allow users to express their opinions and emotions anytime and anywhere due to their openness and convenience. The case-related microblog public opinion is a kind of Internet public opinion related to the case, which has the characteristics of fast transmission speed and high sensitivity. Emotion classification of case-related microblog comments is a multi-classification task of sentiment in a specific field, which aims to quickly and effectively identify the emotions in a large number of comments. This task helps relevant departments to timely evaluate public opinion risks and formulate relevant policies.

收稿日期:2020-03-13;在线发布日期:2020-09-22. 本课题得到国家重点研发计划(2018YFC0830105)、云南省高新技术产业专项(201606)、国家自然科学基金(61972186)资助. 郭贤伟,硕士研究生,主要研究方向为自然语言处理、情感分析. E-mail: xwguo111@163.com. 赖 华,硕士,副教授,中国计算机学会(CCF)会员,主要研究方向为智能信息处理、机器学习. 余正涛(通信作者),博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为自然语言处理、信息检索和机器翻译. E-mail: ztyu@hotmail.com. 高盛祥,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为自然语言处理、信息检索和机器翻译. 相 艳,博士研究生,副教授,中国计算机学会(CCF)会员,主要研究方向为文本挖掘、情感分析.

Emotion classification usually uses the following methods: methods based on emotion dictionary, methods based on traditional machine learning, and methods based on deep learning. Emotion dictionary-based methods, which rely heavily on the emotion dictionary, and the Internet is full of new words to express emotions, it is impractical to improve the emotion dictionary in real-time. Traditional machine learning-based methods often use supervised learning methods, which rely on large amounts of labeled data and complex feature engineering. Deep learning-based methods generally encode text as a whole and lack the effective use of existing emotion computing resources. Since it is difficult for traditional methods to effectively use emotional knowledge such as the emotion words and emoticons commonly used in comments, this paper proposes an emotion classification method of case-related microblog comments that integrates emotional knowledge. This method uses a convolutional neural network with a semantic initialization filter and a fully connected network combined with an attention mechanism to fuse the semantic features and emotional knowledge features of comments to achieve emotion classification. Firstly, it integrates the existing sentiment computing resources to construct an emotional knowledge base that includes the case microblog emotion dictionary, emoticons, network buzzwords, negative words, and degree adverb words. Secondly, considering the role of the emotional knowledge base and part of speech, 15 kinds of emotional knowledge are defined, and the emotional knowledge representation of comments is constructed by the continuous vector representation method proposed in this paper. Then, the semantic representation and emotional knowledge representation of comments are input into a convolutional neural network with semantic initialization filters (INIT-CNN) and a fully connected network with attention mechanism, to obtain a deep semantic feature vector and an attention feature vector. Finally, the two feature vectors are concatenated to fuse semantic and emotional knowledge features to train an emotion classification model, called EK-INIT-CNN (Emotional knowledge enhanced INIT-CNN). In order to prove the effectiveness of the model in this paper, experiments were conducted on the case-related microblog comment dataset and the NLPCC Chinese microblog emotion analysis evaluation dataset. Experiments on the case-related microblog comment dataset show that compared to INIT-CNN, the *Macro\_Precision*, *Macro\_Recall*, and *Macro\_F1* indicators of EK-INIT-CNN have increased by 1.87%, 1.95%, and 1.88%, respectively. The performance of EK-INIT-CNN on NLPCC Chinese microblog emotion analysis evaluation dataset exceeds the best results in the known literature. Experiments show that this method can effectively integrate external emotion knowledge, and has obvious advantages in emotion classification tasks compared to traditional methods.

**Keywords** case-related microblog comments; emotional knowledge representation; convolutional neural network; attention mechanism; emotion classification

## 1 引言

微博作为社交媒体平台的代表,以其开放性、实时性、便捷性的特点使得用户可以随时随地发表自己的观点,表达自己的情感<sup>[1]</sup>. 案件微博舆情是一种

与案件相关的互联网舆情,具有传播速度快、关注度高、敏感度高等特点. 案件微博评论的情绪分类可以看作是特定领域的情感多分类任务,旨在快速有效地识别海量评论中的各种情绪,及时了解和掌握案件微博舆情,以便相关部门迅速做出决策.

微博情感分析通常分为 3 个任务:观点句识别、

情感分类和情感要素抽取<sup>[2]</sup>. 情绪分类是情感分类任务的一个子任务, 要求识别文本中细粒度的情感, 例如: 乐(happiness)、好(like)、怒(anger)、哀(sadness)、惧(fear)、恶(disgust)和惊(surprise). 这与一般的情感分类任务将文本情感分为正负极性或者主客观有很大差别, 难度也更大. 微博文本的情绪分类通常采用以下方法: 基于情绪词典的方法、基于传统机器学习的方法和基于深度学习的方法. 基于情绪词典的方法主要通过情绪词典对文本中的情绪词进行一定的组合计算, 实现文本的情绪分类<sup>[3-4]</sup>. 这种方法对情绪词典的依赖大, 且互联网中表达情绪的新词层出不穷, 实时地完善情绪词典又不现实. 基于传统机器学习的情绪分类常用有监督学习的方法, 需要依赖于大量的标注数据和复杂的特征工程. 现有的情绪分类研究大多是基于深度学习的方法. 虽然基于深度学习的方法在一定程度上有效地避免了基于情绪词典和基于传统机器学习的方法存在的不足, 但是大多数深度学习的方法只是将文本作为一个整体进行编码, 缺乏对现有情感计算资源的有效利用, 也不能很好地体现表情符号和网络用语等对情绪分类的作用.

微博文本具有以下特点: 口语化、富含表情符号和网络用语、常用否定规则和程度副词规则等. 相比一般微博文本, 案件微博文本还具有领域特性, 涵盖许多领域情绪新词. 因此, 针对案件微博评论情绪分类任务本身, 本文还重点关注了领域情绪新词带来的影响. 这些特点给案件微博评论的情绪分类带来很大的挑战. 例如, 针对“西安奔驰女车主维权案”的评论句: “奔驰吃相难看[吐]”和“企业为所欲为要么官商勾结要么背后有黑恶势力”, 其中的表情符号“[吐]”和“吃相难看”、“官商勾结”等领域情绪新词都直接表达了对奔驰公司的厌恶情绪. 如果挖掘了这类情绪词并加入词典的厌恶情绪类别中, 就可以用来辅助识别这类评论句的正确情绪类别. 又如“你积极维权, 不丢人”和“小姐姐加油啊, 哭腔听得我难受, 太委屈了”, 这两条评论中的“不丢人”以及“太委屈”分别使用了否定规则和程度副词规则, 对句子的情绪分别起到反转和加强的作用.

综上, 在案件微博评论中常用表情符号、领域情绪新词、否定规则以及程度副词规则等情绪知识, 传统的情绪分类方法难以有效利用这些情绪知识. 因此, 提出一种融合情绪知识的案件微博评论情绪分

类方法, 主要贡献总结如下:

(1) 通过整合现有的情感计算资源, 对微博常用的表情符号和网络用语进行收集和分类, 构建了一个微博情绪知识库;

(2) 提出了一种基于连续向量的情绪知识表示方法, 表示每个词包含的情绪知识, 使文本在输入时就携带外部的情绪知识;

(3) 提出了一种融合情绪知识的案件微博评论情绪分类方法, 该方法通过一个语义初始化滤波器的卷积神经网络和一个结合注意力机制的全连接网络对评论的语义特征和情绪知识特征进行融合, 实现了情绪分类;

(4) 实验结果表明, 该方法能改善案件微博评论情绪分类任务的效果, 在多个分类实验中取得最好的结果.

## 2 相关工作

微博情感分类研究一直以来都是国内外的研究热点. 传统的情感分类方法通常将微博情感分为正向或负向<sup>[5-6]</sup>, 或者识别细粒度的情感, 例如: 乐(happiness)、好(like)、怒(anger)、哀(sadness)、惧(fear)、恶(disgust)和惊(surprise)<sup>[7-8]</sup>. 这种细粒度的情感分类通常也称为情绪分类.

传统的情绪分类方法常用基于词典和规则的方法, 以及基于机器学习的方法. 张晶等人<sup>[9]</sup>利用情绪词、情绪短语、特殊的情绪表达形式以及结合标点符号和表情符号在情绪分类中的作用构建了情绪词典和情绪规则库, 通过对情绪词典和情绪规则的匹配和计算, 实现对微博的情绪分类. 蒋盛益等人<sup>[10]</sup>利用 word2vec 模型<sup>[11-12]</sup>对微博语料以增量式学习的方法扩展了人工构建的基准情绪词典, 然后结合 HowNet 词典匹配和人工筛选生成最终的情绪词典, 最后分别采用基于情绪词典和基于 SVM 的方法实现了微博的情绪分类. 张仰森等人<sup>[13]</sup>基于情感词典和新浪表情符号词典, 并结合朴素贝叶斯分类器, 提出了一种级联式的微博情绪分类模型, 在 NLPCC2014 中文微博情绪分析评测数据集上证明了所提方法的有效性. 虽然这些研究都取得了不错的效果, 但是由于传统的方法过于依赖人工构建的词典、规则以及复杂的特征工程, 距离实际应用还存在很大的差距.

近年来,越来越多的学者开始使用深度学习的方法进行微博的情绪分类研究. Jiang 等人<sup>[14]</sup>提出的 ESM(Emoticon Space Model),首先用 word2vec 模型预训练中文微博语料的词向量,再利用表情符号的词向量构建表情符号空间,并利用此空间完成所有词向量的重映射,之后使用 SVM 完成分类,在 NLPCC 微博情绪分析评测数据集上取得了当时的最好结果. 何炎祥等人<sup>[15]</sup>为常用的表情符号构建情感空间的特征表示矩阵,通过表情符号的特征表示矩阵与词向量矩阵的乘积运算完成词义到情感空间的映射,使用多通道卷积神经网络模型<sup>[16]</sup>进行微博的情绪分类,效果上实现了对文献[14]的超越. 张仰森等人<sup>[17]</sup>提出的 DAM(Double Attention Model),先利用双向长短时记忆网络和全连接网络,分别对微博文本和文本中包含的情感符号进行编码;接着采用注意力模型分别构建微博文本和情感符号的语义表示,并将两者的语义表示进行融合以提高微博情绪分类的效果;该方法在 NLPCC2013 和 NLPCC2014 中文微博情绪分析评测数据集上,均取得目前已知的最好结果. 除了上述情绪分类方法,一般的情感分类方法也适用于情绪分类中. 关鹏飞等人<sup>[18]</sup>提出一种注意力增强的双向 LSTM 模型,分别使用双向 LSTM 和注意力机制学习文本的语义信息以及每个词对句子情感倾向的权重分布,最后通过并行融合的方式提升微博情感分类的效果.

文献[14-15]仅仅考虑了表情符号对于微博情绪分类的作用,忽略了词语本身的词性及情绪词等情绪知识的作用;文献[17]虽然考虑了对现有情绪知识的利用,通过提取文本中的情感符号以及结合注意力机制来加强微博文本的情感语义,但是该方法在表示情感符号时只是简单地将其映射为词向量,这种词向量表示方式本身并不携带情感信息;文献[18]通过注意力机制衡量每个词对文本情感倾向的重要程度,却完全忽略表情符号、否定词等外部知识的作用. 文献[17-18]的方法虽有不足,但是都证明了注意力机制可以有效地关注表达情感或情绪的重要片段.

综上,大多数基于深度学习的情绪分类方法都缺乏对情绪知识的有效利用. 因此,本文提出一种融合情绪知识的案件微博评论情绪分类方法. 首先,提出一种连续向量表示方法构建评论的情绪知识表

示,该方法可以有效地表示每个词所包含的情绪知识,使文本在输入时就携带外部的情绪知识;然后通过卷积神经网络和结合注意力机制的全连接网络将评论的语义表示和情绪知识表示进行融合,提取深层的语义合成特征用于最终的情绪分类. 实验证明,该方法能有效地融入外部情绪知识,且相对传统方法在情绪分类任务上具有明显的优势.

### 3 融合情绪知识的案件微博评论情绪分类

#### 3.1 微博情绪知识库

本文整理了现有的情感计算资源,收集了微博常用的表情符号和网络用语并分类,采用人工整理与自动扩充相结合的方法构建了一个微博情绪知识库,如表 1 所示.

表 1 微博情绪知识库

情绪知识	数量/个	举例
DUTSD	27466	高兴、愤怒、厌恶
案件微博情绪词典	30836	为民服务、拍桌子、滥用职权
表情符号	226	[哈哈]、[怒]、[吐]
网络用语	210	稀饭、TMD、坑爹
否定词	71	非、不是、绝不
程度副词	219	很、非常、十分

表 1 中, DUTSD 表示大连理工大学的情感词汇本体<sup>[19]</sup>;案件微博情绪词典以 DUTSD 为基础,沿用其 7 大类情绪进行扩充得到,具体构建流程见 3.2 节;表情符号和网络用语从新浪微博开放平台上获得,从中筛选保留了使用频率较高的、带情绪色彩的部分,按照 DUTSD 的 7 大类情绪进行标注分类得到;否定词通过互联网收集现有资源得到;程度副词以 HowNet 中的程度副词为基础进行构建.

#### 3.2 案件微博情绪词典的构建

由于 DUTSD 无法涵盖案件微博评论语料中可能出现的所有领域情绪新词,且这些情绪新词对于案件微博评论的情绪分类十分重要. 因此,基于 SO-PMI 算法和 word2vec 词向量余弦相似度计算方法,从案件微博评论语料库中挖掘领域情绪新词对 DUTSD 进行扩充,构建了案件微博情绪词典,构建流程如图 1 所示.

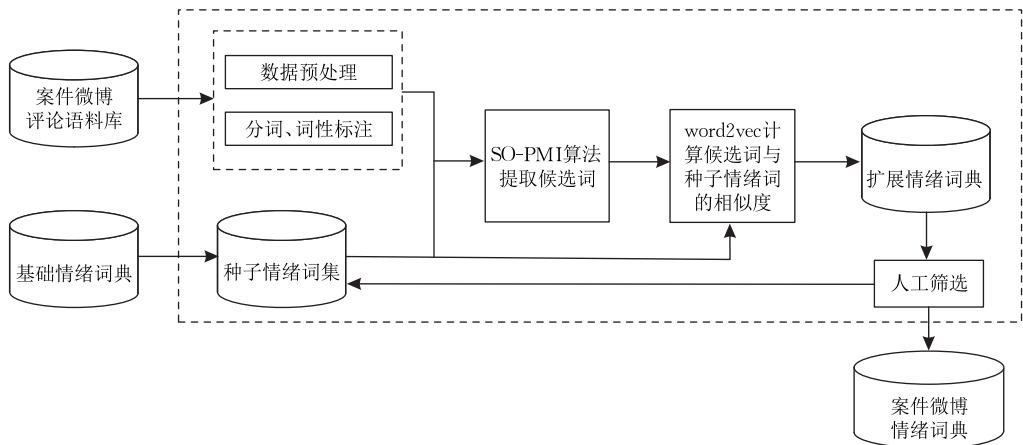


图 1 案件微博情绪词典构建流程

### 3.2.1 基础情绪词典的准备

考虑到表情符号和网络用语对于情绪表达的重要作用,对微博常用的带情绪色彩的表情符号和网络用语进行收集整理,按照乐(happiness)、好(like)、怒(anger)、哀(sadness)、惧(fear)、恶(disgust)和惊(surprise)7类情绪进行分类,加入 DUTSD 中作为基础情绪词典。

### 3.2.2 种子情绪词集的获取

首先对案件微博评论语料进行去重、删除“//”、“@”和 URL 等数据预处理操作;然后使用 jieba 分词工具的添加用户词典功能,以基础情绪词典作为用户词典对评论语料进行分词和词性标注,得到语料库词表。这样做的好处在于:一是可以确保在语料库中出现的种子情绪词可以被独立切分;二是可以确保表情符例如“[哈哈]”的特殊格式存在,并将表情符号的词性标记为“emoji”。这样不仅可以防止分词器将表情符号划分为几个不同的单元,也便于表情符号的识别,对微博的情绪分类有重要的影响。最后将基础情绪词典在语料库词表中出现的所有词作为种子情绪词,构成种子情绪词集。

### 3.2.3 微博领域情绪新词的识别与扩充

文献[20]指出:情感词不仅是指形容词带有感情色彩,还包含有动词、名词、副词等词性,另外,情感词还会与领域有着极强的相关性。因此,考虑词性以及领域相关性会对情感词典的构建有很大的帮助。此外,因为基础情绪词典无法涵盖语料库中所有可能出现的情绪词,所以基于构建的案件微博评论语料库,将 Turney 等人[21]提出的 SO-PMI 算法改进为一种 7 类别的 SO-PMI 算法,用于自动挖掘语料库的候选情绪词。具体实现方法为:通过筛选出语料库词表中词性为:形容词、动词、名词、副词以及表情符

号词性“emoji”的所有词,然后计算每个词与种子情绪词集中各个情绪类别所有情绪词相应的 SO-PMI 值,保留 SO-PMI 值大于零的词作为对应类别的候选情绪词。词的 SO-PMI 值越大,表示该词与当前情绪类别越相关。具体计算公式如式(1)和(2)所示。

$$PMI(word1, word2) = \log_2 \frac{p(word1 \& word2)}{p(word1)p(word2)} \quad (1)$$

$$SO-PMI(word1) = \sum_{word2 \in S_{some-kind}} PMI(word1, word2) - \sum_{word2 \in S_{others}} PMI(word1, word2) \quad (2)$$

式(1)和(2)中,  $word1$  表示在语料库中出现的词,  $word2$  表示种子情绪词;  $p(word1 \& word2)$  表示  $word1$  与  $word2$  在语料库中共同出现的概率,  $p(word1)$  表示  $word1$  在语料库中出现的概率,  $p(word2)$  表示  $word2$  在语料库中出现的概率;  $S_{some-kind}$  表示某一情绪类别种子情绪词集合,  $S_{others}$  表示其它 6 类种子情绪词集合。

通过上述 SO-PMI 算法就能得到 7 个情绪类别对应的候选情绪词。由于 SO-PMI 算法是基于词共现概率来挖掘候选情绪词的,导致部分高频非情绪词也被挖掘出来。为了过滤候选词中的高频非情绪词,采用基于 word2vec 的词向量相似度计算方法<sup>[10,22]</sup>来计算候选词与种子情绪词的相似度,保留高质量候选词作为最终的领域情绪新词。具体实施时,采用 python 第三方模块 gensim<sup>①</sup> 中的 word2vec 词向量余弦相似度计算函数,计算每个类别的候选情绪词与其对应类别的种子情绪词的平均相似度,保留平均相似度大于 0.5 的候选词作为其对应类别的领域情绪新词。计算公式如下:

① <https://radimrehurek.com/gensim/apiref.html>

$$\cos(v_i, v_j) = \frac{v_i \cdot v_j}{\|v_i\| \cdot \|v_j\|} \quad (3)$$

$$aver\_cos_i = \left( \sum_{j=1}^n \frac{v_i \cdot v_j}{\|v_i\| \cdot \|v_j\|} \right) / n, i=1, 2, \dots, m \quad (4)$$

式(3)和(4)中,  $v_i$ 表示候选词的词向量,  $m$ 表示当前类别候选词总数;  $v_j$ 表示种子情绪词的词向量,  $n$ 表示当前类别种子情绪词总数。

通过上述方法就可以得到一个扩展情绪词典, 然后通过人工筛选高质量的扩展情绪词加入种子情绪词集, 进行增量式的迭代进而挖掘更多领域情绪新词. 最后, 在算法挖掘不到情绪新词时停止迭代, 得到最终的扩展情绪词典. 本文选择将扩展情绪词典中表情符号和网络用语单独筛选出来, 方便后续的实验分析. 筛选完后, 将扩展情绪词典和 DUTSD 整合成一个相对完善的情绪词典, 称为案件微博情绪词典. 扩展前后的词典对比如表 2 所示。

表 2 扩展前后的词典对比

词典名称	数量/个	举例
DUTSD(扩展前)	27466	惊吓、没皮没脸、狼狽为奸
案件微博情绪词典(扩展后)	30836	吓成、吃相难看、官商勾结

### 3.3 EK-INIT-CNN 模型

为了同时表示文本自身的语义特征和情绪知识特征, 引入情绪知识库为词语定义了词性和情绪标签等 15 种情绪知识, 提出一种基于连续向量的情绪知识表示方法. 然后, 将案件微博评论的语义表示和情绪知识表示分别输入一个滤波器语义初始化的卷积网络和一个结合注意力机制的全连接网络, 提取深层的语义特征和情绪知识特征, 帮助模型更好地学习文本的情绪. 由此, 构建了一个融合情绪知识的神经网络模型 EK-INIT-CNN, 模型结构如图 2 所示。

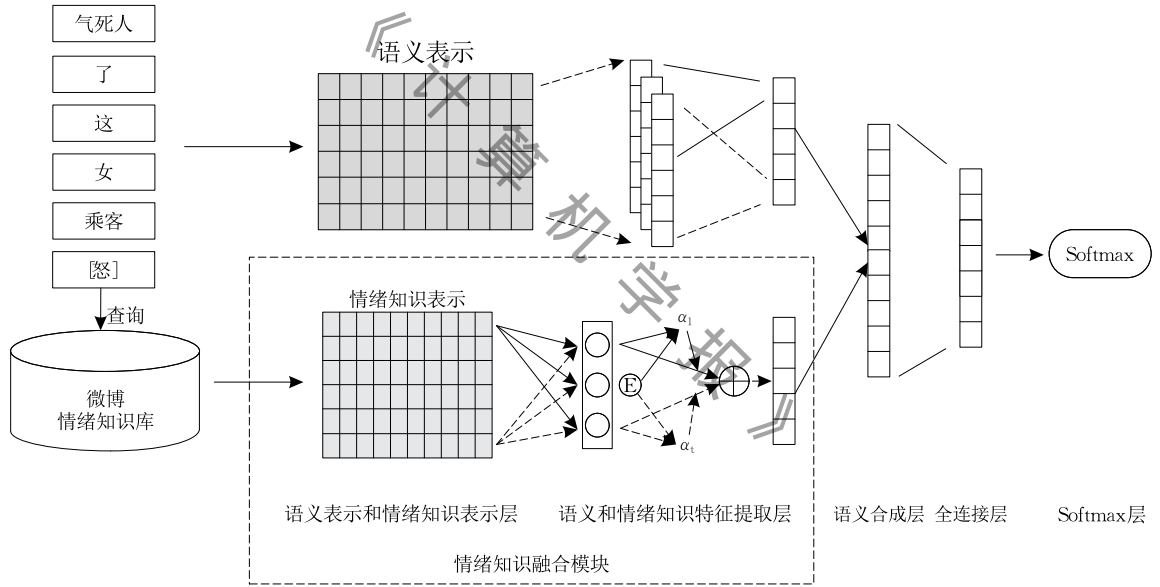


图 2 EK-INIT-CNN 模型结构

#### 3.3.1 语义表示和情绪知识表示层

EK-INIT-CNN 模型的输入分为案件微博评论的语义表示和情绪知识表示. 案件微博评论的语义表示是一个对评论句进行分词, 再通过加载预训练词向量词表  $W^{N \times d}$  对每个词进行查询并分配词向量的过程. 其中,  $N$  表示词表的词汇数,  $d$  表示词向量维度. 假设一个包含  $n$  个词的评论文本序列  $T = \{w_1, w_2, \dots, w_n\}$ , 对于  $T$  中的每个词  $w_i$  都能通过词表  $W^{N \times d}$  查询到一个词向量  $v_i$ , 用  $\oplus$  运算符进行向量方向的拼接操作, 则序列  $T$  的语义表示矩阵  $M_T$  为

$$M_T = v_1 \oplus v_2 \oplus \dots \oplus v_n \quad (5)$$

案件微博评论的情绪知识表示是一个基于连续向量表示方法构建情绪知识表示矩阵的过程. 首先,

根据微博情绪知识库, 为每个词定义了词性及情绪标签等 15 种情绪知识, 见表 3; 然后, 给定包含  $n$  个词的评论句  $T = \{w_1, w_2, \dots, w_n\}$ , 对于每个词  $w_i$ ,  $i=1, 2, \dots, n$ , 按照表 3 的 15 种情绪知识顺序依次进行如下操作: 若  $w_i$  包含第一种知识, 则为其随机初始化一个  $m$  维的非零连续向量, 否则为其初始化一个  $m$  维的全零向量, 本文取  $m=10$ . 以此类推, 每个  $w_i$  都将得到 15 个  $m$  维向量, 将这 15 个向量拼接成一个  $(15 \times m)$  维的向量  $e_i$ , 称为情绪知识向量. 遍历完  $T$  中的所有词后将得到的  $n$  个情绪知识向量按行拼接, 得到一个  $n \times (15 \times m)$  维的情绪知识表示矩阵  $M_E$ :

$$M_E = e_1 \oplus e_2 \oplus \dots \oplus e_n \quad (6)$$

表3 词语可能包含的情绪知识

情绪知识	是否为:happiness 标签、like 标签、anger 标签、sadness 标签、fear 标签、disgust 标签、surprise 标签、否定词、程度副词、名词、动词、形容词、副词、网络用语、表情符号
$m$ 维向量	连续向量(是)/全零向量(否)

### 3.3.2 语义和情绪知识特征提取层

EK-INIT-CNN 模型的特征提取层分为语义特征提取层和情绪知识特征提取层. 语义特征提取层由一个  $N$ -gram 卷积层和池化层组成. 微博文本中存在许多完整的  $n$ -gram 知识可作为情绪分类的重要依据, 采用了一种卷积滤波器语义初始化技术<sup>[23]</sup>: 首先使用一种新颖的朴素贝叶斯加权技术从语料中挑出一系列  $n$ -gram 特征; 接着采用  $K$ -Means 聚类对这些  $n$ -gram 特征的嵌入向量进行聚类, 得到 top- $k$  个聚类中心; 最后, 在模型的卷积层中将它们的聚类中心向量输入滤波器的第一个  $n$  位置. 举个例子: 如果  $n=2$  (本文取值), 滤波器窗口大小为 3, 那么就在滤波器的第 1、2 行放入一个 2-gram 聚类中心向量, 第 3 行则采用随机初始化的策略. 与传统的卷积神经网络不同, 采用这样的策略能使模型在训练一开始就集中学习重要的  $n$ -gram 特征.

假设卷积滤波器  $\mathbf{W}^f \in R^{k \times d}$ ,  $k$  为滤波器窗口大小,  $d$  表示词向量维度. 对于输入的从  $i$  行到  $i+k-1$  行的局部特征矩阵  $\mathbf{X}_{i:i+k-1}$ , 卷积滤波器所提取的特征图的第  $i$  个特征值表示为

$$z_i = f(\mathbf{W}^f \otimes \mathbf{X}_{i:i+k-1} + b) \quad (7)$$

其中,  $f(\cdot)$  表示非线性激活函数, 可以是  $\tanh$  或  $\text{relu}$  等.  $\otimes$  表示卷积运算符,  $b$  表示偏置值.

对于文本块序列  $\mathbf{X} = \{\mathbf{X}_{1:k}, \mathbf{X}_{2:k+1}, \dots, \mathbf{X}_{n-k+1:n}\}$ , 经过一个卷积层就能得到其局部特征图  $\mathbf{Z}$ :

$$\mathbf{Z} = [z_1, z_2, \dots, z_{n-k+1}] \quad (8)$$

随后, 将经卷积层得到的特征图  $\mathbf{Z}$  输入到池化层, 采用的是 max-pooling 池化, 得到局部最优特征  $\hat{z}$ :

$$\hat{z} = \max\{\mathbf{Z}\} \quad (9)$$

此外, 卷积层采取了  $j$  种不同窗口大小的卷积滤波器, 每种窗口大小的滤波器各  $m$  个, 以此来提取更丰富的语义特征. 因此, 经过卷积层和池化层后得到的局部最优特征向量表示为

$$\mathbf{V} = [\hat{z}_{1,k_1}, \dots, \hat{z}_{m,k_1}, \dots, \hat{z}_{1,k_j}, \dots, \hat{z}_{h,k_j}, \dots] \quad (10)$$

其中,  $\hat{z}_{h,k_j}$  表示窗口大小为  $k$  的第  $j$  种类型滤波器提取的第  $h$  个局部最优特征.

模型的情绪知识特征提取层由一个全连接层和注意力层组成. 由于每个词映射得到的情绪知识向

量之间并不存在较强的语义依赖关系, 因此采用全连接网络对情绪知识表示矩阵进行非线性变换, 进行最初的编码, 编码后的情绪知识表示的维度保持不变. 并且, 每个情绪知识向量对于文本情绪表达的重要程度各不相同, 所以进一步将经全连接层编码后的情绪知识表示输入注意力层中, 得到情绪知识表示的注意力特征向量. 注意力层采用 Luong 等人<sup>[24]</sup>提出的全局注意力机制. 注意力层对输入的每个状态  $h_i$  进行加权求和, 从而得到情绪知识表示的注意力特征向量  $\mathbf{v}_{att}$ :

$$\mathbf{v}_{att} = \sum_{i=1}^T \alpha_i h_i \quad (11)$$

其中,  $T$  表示输入状态的数目,  $\alpha_i$  表示输入状态  $h_i$  的注意力权重, 通过如下公式计算:

$$\alpha_i = \frac{\exp(\mathbf{e}_i^T \mathbf{E})}{\sum_{i=1}^T \exp(\mathbf{e}_i^T \mathbf{E})} \quad (12)$$

$$\mathbf{e}_i = \tanh(\mathbf{W}h_i + b) \quad (13)$$

式(12)和(13)中,  $\mathbf{W}$  与  $b$  分别为注意力层的权重和偏置,  $\mathbf{E}$  表示随机初始化的上下文特征向量. 三个参数都参与模型的训练, 通过不断学习得到.  $\mathbf{e}_i$  表示计算注意力权重时的中间状态,  $\mathbf{e}_i^T$  是其转置. 通过上述注意力层, 衡量每个词映射得到的情绪知识向量的重要程度, 最终为每个评论构建一个  $(15 \times m)$  维的情绪知识注意力特征向量.

### 3.3.3 语义合成层和全连接层

案件微博评论的语义表示矩阵和情绪知识表示矩阵输入特征提取层后, 分别得到语义特征向量  $\mathbf{V}_{\text{semantic}}$  和情绪知识注意力特征向量  $\mathbf{V}_{\text{knowledge}}$ . 将两个特征向量直接进行拼接, 得到文本的语义合成向量  $\hat{\mathbf{V}}$ :

$$\hat{\mathbf{V}} = \mathbf{V}_{\text{semantic}} \oplus \mathbf{V}_{\text{knowledge}} \quad (14)$$

然后, 语义合成向量  $\hat{\mathbf{V}}$  经全连接层进行变换后得到案件微博评论的情绪得分向量  $\mathbf{s}$ . 全连接层的隐藏单元个数等于分类类别数.

### 3.3.4 Softmax 层(情绪得分计算层)

最后, 将案件微博评论的情绪得分向量  $\mathbf{s}$  输入到 Softmax 层, 使其转化为一个条件概率分布:

$$P_i(\mathbf{s}) = \frac{\exp(\mathbf{s}_i)}{\sum_{j=1}^C \exp(\mathbf{s}_j)}, \quad i=1, 2, \dots, C \quad (15)$$

其中,  $C$  表示情绪标签的数目. 采用交叉熵损失函数来衡量情绪标签的预测概率分布与真实概率分布之间的差距, 并通过反向传播算法来训练和更新模型的参数. 损失函数表示为

$$loss = - \sum_{x \in T} \sum_{i=1}^C P_i^p(x) \log_2(P_i^p(x)) \quad (16)$$

其中,  $T$  为训练集,  $x$  为训练集中的每一个样本,  $P_i^p(x)$  表示样本  $x$  的情绪标签的真实概率分布,  $P_i^p(x)$  表示样本  $x$  的情绪标签的预测概率分布。

## 4 实验与分析

### 4.1 实验数据集

为了验证情绪知识融合方法在公共数据集上的有效性, 在 NLPCC2013 和 NLPCC2014 中文微博情绪分析评测数据集上设置了多组对比实验。为了体现情绪知识融合方法在案件微博评论情绪分类任务上的优势, 在构建的案件微博评论数据集<sup>①</sup>上设置了多组对比实验。

NLPCC2013 和 NLPCC2014 中文微博情绪分析评测数据集来自新浪微博, 每个数据集分为训练集和测试集, 分别记为 2013\_train<sup>②</sup>、2013\_test<sup>③</sup>、2014\_train<sup>④</sup> 和 2014\_test<sup>⑤</sup>。每条微博分为有无情绪两大类, 无情绪的标注为 none, 有情绪的按照 DUTSD 的乐 (happiness)、好 (like)、怒 (anger)、哀 (sadness)、惧 (fear)、恶 (disgust) 和惊 (surprise) 7 大类情绪标注, 与本文的标注体系一致。其中, 在 NLPCC2014 数据集的训练集中, happiness 和 sadness 类各有一条微博内容为空, 选择将其删除。评测数据集可用于有无情绪的主客观分类和有情绪的情绪分类。本文研究的是情绪分类任务, 最终未用到 none 标签数据, 两个数据集的具体分布如表 4 所示。

表 4 NLPCC2013 和 NLPCC2014 评测数据集 (单位: 条)

情绪标签	2013_train	2013_test	2014_train	2014_test
none	1828	4873	6591	3603
happiness	370	1116	1459	441
like	595	1558	2204	1042
anger	235	436	669	128
sadness	385	744	1173	189
fear	49	102	148	46
disgust	425	935	1392	389
surprise	113	236	362	162
总数	4000	10000	13998	6000

通过新浪微博平台采集了近年来备受关注的 7 个案件微博的评论数据, 构建了案件微博评论数据集。其中, 涞源反杀案共 10 812 条; 江歌案共 13 624 条; 赵宇见义勇为案共 17 491 条; 章莹颖遇害案共 17 875 条; 重庆公交车坠江案共 33 774 条; 河南玛莎拉蒂肇事逃逸案共 37 162 条; 以及西安奔驰

女车主维权案共 58 626 条, 总共 189 364 条。从中随机采样了 30 000 条评论, 沿用 DUTSD 的 7 类情绪标签进行人工标注, 标注工作由两个人完成。最后, 保留标注一致的、包含情绪的 11 593 条评论作为案件微博评论实验数据集, 按照近似 8:1:1 的比例将实验数据集分为训练集、验证集和测试集, 数据集分布如表 5 所示。

表 5 案件微博评论实验数据集 (单位: 条)

情绪标签	训练集	验证集	测试集
happiness	1371	152	129
like	1348	174	152
anger	1348	152	152
sadness	1355	162	145
fear	1367	169	133
disgust	1342	144	158
surprise	1338	140	162
总数	9469	1093	1031

### 4.2 预训练词向量的使用

本文分别在计算候选词与种子情绪词的余弦相似度和为实验中的所有神经网络模型构建输入的词嵌入时, 使用文献[25]公开的预训练词向量<sup>⑥</sup>, 为每个词分配一个相应的词向量。该词向量使用 NLP Lab 公开的大规模微博语料库进行预训练, 预训练模型采用的是 word2vec<sup>⑦</sup> 开源工具中的 Skip-gram 模型, 词向量维度为 300 维, 词表大小为 195 197。此外, 在神经网络模型的训练过程中, 使预训练词向量不断学习更新。

### 4.3 实验对比模型

MNB 模型 (Multinomial Naïve Bayes)<sup>[26]</sup>。在许情感多分类任务中都取得优秀的结果, 模型沿用文献[15]的文本特征选取方法。

SVM 模型 (Support Vector Machines)。被广泛用于情感分类任务中并取得最好的结果, 例如文献[27]和文献[28]。用于 SVM 模型的文本特征提取和表示方法与文献[15]一致。

BLSTM 模型 (Bidirectional Long Short-Term Memory)<sup>[29]</sup>。该模型通过双向的 LSTM 对文本表示进行编码, 得到上下文相关的文本特征用于情感分类。模型的参数设置与文献[17]保持一致。

Att-BLSTM 模型 (Attention-Based Bidirectional Long Short-Term Memory)<sup>[30]</sup>。在 BLSTM 模型的

① <https://github.com/guo-huojian/dataset/archive/master.zip>

② <http://tcci.ccf.org.cn/conference/2013/dldoc/evsam02.zip>

③ <http://tcci.ccf.org.cn/conference/2013/dldoc/evsam02.zip>

④ <http://tcci.ccf.org.cn/conference/2014/dldoc/evtestdata1.zip>

⑤ <http://tcci.ccf.org.cn/conference/2014/dldoc/evsam01.zip>

⑥ <https://github.com/Embedding/Chinese-Word-Vectors>

⑦ <https://code.google.com/p/word2vec/>



基础上引入注意力机制,文本词嵌入经 BLSTM 编码后输入到注意力层中,学习一个包含注意力信息的上下文向量,使得模型能够学到更重要的语义信息.模型参数与文献[30]保持一致.

Capsule-B 模型(CNN Based Capsule)<sup>[31]</sup>.该模型基于 Sabour 等人提出的胶囊网络<sup>[32]</sup>进行改进得到,在情感分类、新闻分类等多个文本分类任务中取得最好的结果.模型沿用文献[31]的参数设置.

HAN模型(Hierarchical Attention Networks)<sup>[33]</sup>.一个基于双向 GRU 的层次注意力模型.模型通过词级的双向 GRU 编码后输入词级注意力层,得到文本的句子级语义表示;再将其输入一个句子级的双向 GRU 层和注意力层,得到整个文本的语义表示用于分类.该模型能有效关注文本中重要的词和句子,在包括情感分类在内的多个文本分类任务中取得最好的结果.模型的参数设置与文献[33]保持一致.

CNN 模型(Convolutional Neural Networks)<sup>[16]</sup>.Kim 提出的卷积神经网络,通过将文本的语义表示输入卷积神经网络中,提取深层的语义特征实现文本的情感分类.该模型有多种变体,实验采用的是 CNN-non-static,表示使用预训练词向量并在训练过程中动态更新词向量.

INIT-CNN 模型(CNN with Initializing Convolutional Filters)<sup>[23]</sup>.该模型基于上述的 CNN 模型,通过 K-Means 算法对一系列重要的  $n$ -gram 嵌入向量进行聚类,然后用得到的聚类中心来初始化卷积滤波器,实现微博文本的情感分类.因为 INIT-CNN 具有性能优秀、可并行处理和模型复杂度低的优点,所以将其选做基础模型进行改进.

#### 4.4 模型参数设置

对于机器学习模型,MNB 模型的平滑因子  $\alpha$  设置为 1.0;SVM 模型的正则化常数  $C$  设置为 1.0,核函数选择线性核函数 linear.对于神经网络模型,除了在 4.3 节所列模型中提及了采用相关文献的参数设置外,CNN-non-static、INIT-CNN 以及 EK-INIT-CNN 的可调参数设置如表 6 所示.

表 6 可调参数设置

可调参数	值
优化器	Adam
batch_size	64
dropout	0.5
learning_rate	0.001
卷积滤波器窗口大小 $k$	2, 3, 4, 5
每种窗口大小的滤波器个数 $m$	100
注意力层隐藏单元个数	150

#### 4.5 评价指标

本文使用 NLPCC2013 中文微博情绪分析评测<sup>①</sup>中的:精确率(Precision)、召回率(Recall)、F1 值(F1)、宏平均精确率(Macro\_Precision)、宏平均召回率(Macro\_Recall)、宏平均 F1 值(Macro\_F1)以及微平均 F1 值(Micro\_F1)作为评价指标.

#### 4.6 实验结果分析

##### 4.6.1 NLPCC 中文微博情绪分析评测

为了验证本文的情绪知识融合方法在公共数据集上的有效性,在 NLPCC 微博情绪分析评测数据集上进行了对比实验,结果见表 7 和表 8.表中,E-ESM、BLSTM、MCNN、EMCNN 以及 DAM 的实验结果摘自文献[15]和文献[17].EK-INIT-CNN(base)表示利用微博情绪知识库中的 DUTSD、否定词、程度副词、网络用语和表情符号来融入情绪知识.因为 NLPCC 评测数据集不具有案件微博领域特性,所以在公共数据集上未使用案件微博情绪词典,而是使用 DUTSD.

文献[14]提出的表情符号空间模型(ESM)通过两种不同的求和策略得到两个版本,E-ESM 是其中表现更好的版本,在 NLPCC 情绪分析评测数据集上取得当时最好的结果.BLSTM<sup>[29]</sup>通过双向的 LSTM 提取上下文相关的文本特征用于情感分类,是情感分类常用模型之一.MCNN 是文献[15]中 EMCNN 模型的基础模型,是经典的卷积神经网络模型.EMCNN<sup>[15]</sup>为常用的表情符号构建情感空间的特征表示矩阵,并完成词义到情感空间的映射,然后通过 MCNN 模型训练一个分类器,性能超越了 ESM.DAM 由文献[17]提出,在基于注意力机制的 BLSTM 模型的基础上增加了一个通道,将文本的情感符号集合的语义表示输入全连接层和注意力层获取附加的语义特征,增强文本的情感语义;该模型在 NLPCC2013 和 NLPCC2014 情绪分析评测数据集上均取得当前已知的最好结果.

表 7 NLPCC2013 评测数据集实验结果

Models	Macro_F1/%	Micro_F1/%
E-ESM	35.00	43.90
BLSTM	31.51	40.78
MCNN	34.62	43.65
EMCNN	35.17	44.22
DAM	36.56	45.48
INIT-CNN	33.32	45.80
EK-INIT-CNN(base)	40.96	51.88

① <http://tcci.ccf.org.cn/conference/2013/dldoc/ev02.pdf>

表 8 NLPCC2014 评测数据集实验结果

Models	Macro_F1/%	Micro_F1/%
E-ESM	37.86	44.25
BLSTM	33.85	44.56
MCNN	38.29	45.34
EMCNN	39.40	47.23
DAM	41.42	49.44
INIT-CNN	39.39	54.11
EK-INIT-CNN(base)	51.02	63.66

从表 7 和表 8 可以看出, EK-INIT-CNN(base) 在两个评测数据集上的性能均超过了所有对比模型. 相比目前已知的最好模型 DAM, 在 NLPCC2013 评测数据集上, EK-INIT-CNN(base) 的 *Macro\_F1* 和 *Micro\_F1* 分别提升了 4.40% 和 6.40%; 在 NLPCC2014 数据集上, *Macro\_F1* 和 *Micro\_F1* 分别提升了 9.60% 和 14.22%. 实验结果表明融合外

表 9 案件微博评论情绪分类实验结果

Models	Macro_Precision/%	Macro_Recall/%	Macro_F1/%	Micro_F1/%
MNB	75.88	75.80	75.50	75.46
SVM	85.54	85.15	85.19	84.77
BLSTM	88.65	88.24	88.29	88.07
Att-BLSTM	89.61	89.60	89.49	89.33
Capsule-B	88.94	88.61	88.70	88.38
HAN	90.84	90.88	90.82	90.69
CNN-non-static	88.23	88.28	88.24	88.07
INIT-CNN	90.00	89.95	89.93	89.82
EK-INIT-CNN	91.87	91.90	91.81	91.76

分析表 9 可知, 本文模型 EK-INIT-CNN 的四项指标均超过所有基准模型. 从表中可以看出, 传统的 MNB 模型和 SVM 模型都取得了不错的效果, 但是与各深度学习模型相比还存在一定的差距. 其中, 与机器学习常用的 SVM 分类器相比, EK-INIT-CNN 的 *Macro\_Precision*、*Macro\_Recall*、*Macro\_F1* 和 *Micro\_F1* 分别提升了 6.33%、6.75%、6.62% 和 6.99%; 与情感分类常用的 CNN-non-static 模型相比, EK-INIT-CNN 的 *Macro\_Precision*、*Macro\_Recall*、*Macro\_F1* 和 *Micro\_F1* 分别提升了 3.64%、3.62%、3.57% 和 3.69%; 与情感分类任务中表现优秀的 Att-BLSTM 相比, 分别提升了 2.26%、2.30%、2.32% 和 2.43%.

此外, 还与在多个情感分类任务中取得最好结果的模型进行了对比. 与 Capsule-B 相比, EK-INIT-CNN 的 *Macro\_F1* 和 *Micro\_F1* 分别提升了 3.11%、3.38%; 与 HAN 相比, 分别提升了 0.99% 和 1.07%. 实验结果表明, 本文融合情绪知识的方法在案件微博情绪分类任务中具有明显的优势.

另外, INIT-CNN 是在 CNN-non-static 的基础上使用了卷积滤波器语义初始化技术. 对比 CNN-

部情绪知识的方法在 NLPCC 评测任务中具有明显的优势, 证明了情绪知识融合方法的有效性. 同时, EK-INIT-CNN(base) 在 2014 年数据集上的性能提升要好于 2013 年的, 这是因为 2014 年数据集的训练集更大使得模型训练更充分. 这与文献[17]的结论是一致的. 此外, 由于 NLPCC 评测数据集属于类别不平衡数据集, 导致所有模型表现都较差. 但是, EK-INIT-CNN(base) 相比其它模型的表现已经取得了较大的改善.

#### 4.6.2 案件微博评论情绪分类任务

为了证明本文融合情绪知识的方法在案件微博评论情绪分类任务上具有优势, 将 EK-INIT-CNN 模型与现有 8 个基准模型进行了实验对比, 实验结果如表 9 所示.

non-static, 可以看出 INIT-CNN 的 *Macro\_F1* 和 *Micro\_F1* 分别提升了 1.69% 和 1.75%. 这表明本文使用的卷积滤波器语义初始化技术可以显著提升传统卷积神经网络的效果.

#### 4.6.3 各个情绪类别的分类性能分析

为了证明本文方法能有效改善模型对单个情绪类别的分类性能, 将 EK-INIT-CNN 与基础模型 INIT-CNN 进行了实验对比, 使用 *Precision*、*Recall* 和 *F1* 三个评价指标, 结果见图 3~图 5.

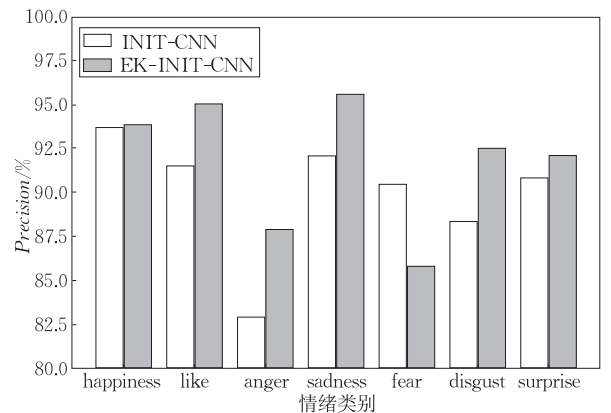


图 3 各个情绪类别的精确率对比结果

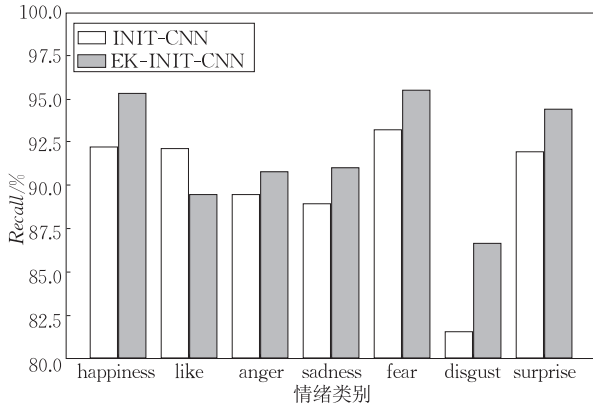


图 4 各个情绪类别的召回率对比结果

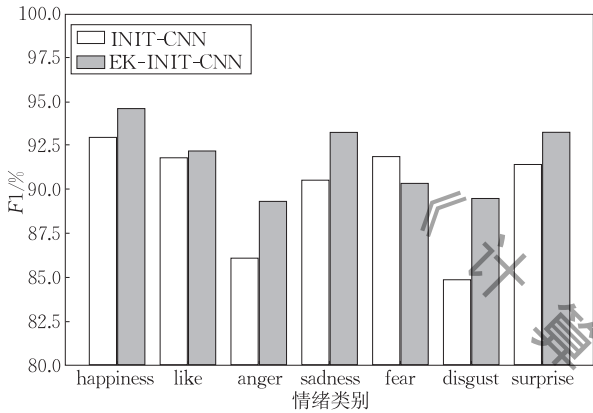


图 5 各个情绪类别的 F1 值对比结果

分析图 3~图 5 可知,本文模型 EK-INIT-CNN 除了 fear 类的 *Precision*、*F1* 指标以及 like 类的 *Recall* 指标比 INIT-CNN 的低之外,其它类别的三项指标都超过 INIT-CNN。

通过分析,我们发现 fear 类的测试数据有 133 条,其中被 EK-INIT-CNN 正确预测的有 121 条,而其它类别的数据被 EK-INIT-CNN 错误地预测成 fear 类的数据有 8 条。进一步分析发现这 8 条数据基本上存在两种情绪,但是在标注过程中只标注其中一种情绪,而 EK-INIT-CNN 把它们错误地预测为另外一种情绪。例如,“听到最后全车人的尖叫都觉得太恐怖了。[悲伤][悲伤][悲伤][悲伤][悲伤]”这一评论被人工标注为 sadness 类,由于 EK-INIT-CNN 更关注评论中“尖叫”和“太恐怖”所表达的情绪,因此将其错误地预测为 fear 类。like 类的 *Recall* 指标降低的原因则是 like 类测试数据中有些具有多种情绪的评论被 EK-INIT-CNN 误分类为其它情绪。例如,评论“我爱自己的国家,但是对人很失望,就这样”,被标注为 like 类, EK-INIT-CNN 将其误分类成 sadness 类。总结来说,测试集中具有多种情绪的评论有时对 EK-INIT-CNN 模型的学习机制不

利。这是导致 EK-INIT-CNN 在这两个类别的指标比 INIT-CNN 更低的主要原因。

此外,相比 INIT-CNN, EK-INIT-CNN 在识别 disgust 类的性能提升最明显。这是由于案件微博评论中表达 disgust 情绪的评论特征更加鲜明。例如,我们发现测试集中“恶心”、“厌恶”和“[吐]”等词或表情符号的使用更频繁,这对 EK-INIT-CNN 中额外融入情绪知识的模块更有利,因此它能更好地识别 disgust 类。总之,实验结果表明 EK-INIT-CNN 的总体性能优于 INIT-CNN,证明了本文方法能有效改善模型在单个情绪类别的分类性能。

#### 4.6.4 情绪知识融合模块的有效性分析

为了证明本文提出的情绪知识融合模块(见图 2)与各基准模型相结合的有效性,设置了多组对比实验。首先分析了在基础模型 INIT-CNN 上增加情绪知识融合模块的有效性,对比模型和结果见图 6。其中,INIT-CNN 为本文的基础模型,未融入任何情绪知识;EK-INIT-CNN(base)表示利用微博情绪知识库中的 DUTSD、否定词、程度副词、网络用语和表情符号来融入情绪知识;EK-INIT-CNN 表示利用微博情绪知识库中的案件微博情绪词典、否定词、程度副词、网络用语和表情符号来融入情绪知识。

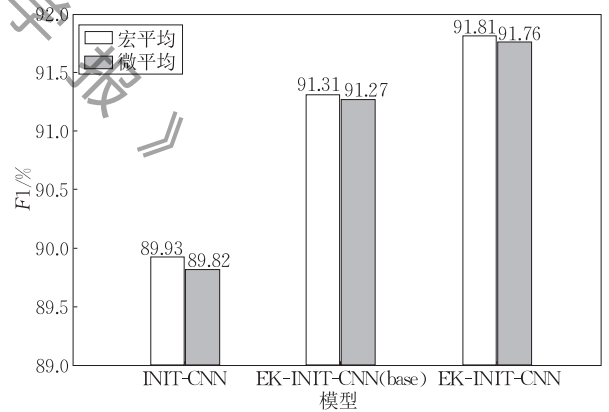


图 6 INIT-CNN 结合情绪知识融合模块的有效性验证

分析图 6 可知,相比 INIT-CNN 模型, EK-INIT-CNN(base)的宏平均 *F1* 和微平均 *F1* 分别提升了 1.38% 和 1.45%; EK-INIT-CNN 的宏平均 *F1* 和微平均 *F1* 分别提升了 1.88% 和 1.94%。这表明, INIT-CNN 与情绪知识融合模块结合,可以显著提升模型的性能。而相比 EK-INIT-CNN(base), EK-INIT-CNN 的宏平均 *F1* 和微平均 *F1* 分别提升了 0.50% 和 0.49%,大约占总提升(EK-INIT-CNN 相比 INIT-CNN 的提升)的 27% 和 25%。这表明,在 DUTSD 基础上扩充的案件微博情绪词典包含更

丰富的领域情绪新词,对于案件微博评论的情绪分类任务具有显著的作用,进一步的实例分析结果也证明了这一结论.例如,评论“为了钱官商勾结.苦的是老百姓”和“供不应求的行业都吃相难看[doge]”中的“官商勾结”及“吃相难看”是案件微博情绪词典中扩充的领域情绪新词.模型预测这两条评论时,利用案件微博情绪词典的 EK-INIT-CNN 将它们正确地预测为 *disgust* 类情绪,而利用 DUTSD 的

EK-INIT-CNN(base)都预测错误.

本文还将情绪知识融合模块分别与性能较优的 Att-BLSTM、Capsule-B 和 HAN 模型结合,新模型记为: Att-BLSTM + EK (Emotional knowledge)、Capsule-B + EK 和 HAN + EK. 然后将这三个模型与 EK-INIT-CNN 进行实验对比,以此说明本文将 INIT-CNN 模型作为基础模型与情绪知识融合模块结合的优势,实验结果见表 10.

表 10 情绪知识融合模块的有效性验证结果

Models	Macro_Precision/%	Macro_Recall/%	Macro_F1/%	Micro_F1/%
Att-BLSTM	89.61	89.60	89.49	89.33
Att-BLSTM+EK	91.45	91.52	91.44	91.31
Capsule-B	88.94	88.61	88.70	88.38
Capsule-B+EK	88.89	89.12	88.90	88.77
HAN	90.84	90.88	90.82	90.69
HAN+EK	91.14	91.36	91.14	91.08
INIT-CNN	90.00	89.95	89.93	89.82
EK-INIT-CNN	91.87	91.90	91.81	91.76

分析表 10 可知,将本文的情绪知识融合模块与 Att-BLSTM、Capsule-B 和 HAN 等模型相结合,能有效地改善模型的性能;其中,Att-BLSTM + EK 模型的性能提升最明显, *Macro\_Precision*、*Macro\_Recall*、*Macro\_F1* 和 *Micro\_F1* 分别提升了 1.84%、1.92%、1.95% 和 1.98%. 另外,由于 HAN 的分层注意力结构,本身已经有效地关注了文本中表达情绪的重要片段,导致 HAN 在结合情绪知识融合模块时性能提升不那么明显.

从表中还可以看出,在基础模型 INIT-CNN 上增加情绪知识融合模块,得到的本文模型 EK-INIT-CNN,取得了最好的性能. 相比性能第二的 Att-BLSTM + EK, EK-INIT-CNN 的 *Macro\_F1* 和 *Micro\_F1* 分别高了 0.37% 和 0.45%. 实验结果表明,结合本文的情绪知识融合模块能有效地改善基准模型的性能,且在 INIT-CNN 模型上进行结合的效果最好.

#### 4.6.5 各项情绪知识的性能分析

为了考察各项情绪知识对案件微博评论情绪分类性能的影响,将不同的情绪知识组合应用于 EK-INIT-CNN,设置了 6 组对比模型. 各模型记为: Model\_1(案件微博情绪词典)、Model\_2(表情符号)、Model\_3(网络用语)、Model\_4(表情符号 + 网络用语)、Model\_5(案件微博情绪词典 + 表情符号 + 网络用语)以及 Model\_6(案件微博情绪词典 + 否定词 + 程度副词). 因为否定词和程度副词只有与情绪词搭配才起作用,所以未单独应用于 EK-INIT-CNN. 6 组对比模型与 INIT-CNN 的实验结果如图 7 所示.

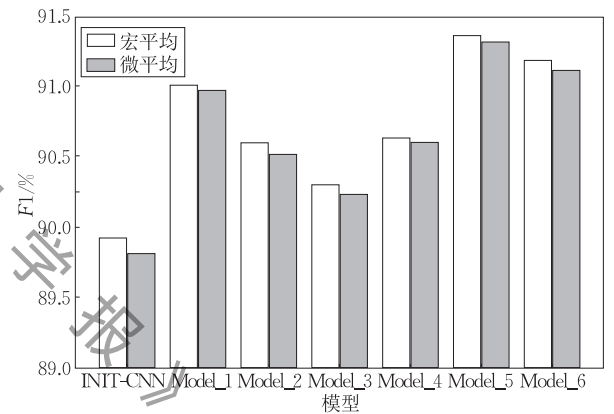


图 7 各项情绪知识性能对比

从图 7 可看出,相比基础模型 INIT-CNN,使用各项情绪知识的模型性能均有提升. 其中,单独使用某项情绪知识时,Model\_1(案件微博情绪词典)和 Model\_2(表情符号)的性能提升比较明显;Model\_3(网络用语)的性能提升较小. 组合使用时,Model\_5(案件微博情绪词典 + 表情符号 + 网络用语)的性能最优;而 Model\_6(案件微博情绪词典 + 否定词 + 程度副词)相比 Model\_1(案件微博情绪词典)的性能也有提升,这表明否定词和程度副词对情绪表达也是起作用的. 同时,否定词和程度副词对性能的作用不明显的原因主要是数据集中使用否定词和程度词的评论相对较少. 实验结果表明,本文构建的微博情绪知识库可以有效地与深度学习模型搭配使用,且融合情绪知识的方法还避免了传统深度学习模型缺乏利用现有情感计算资源的现象.

## 5 总结与展望

针对案件微博评论,本文提出一种融合情绪知识的情绪分类方法.该方法通过构建一个微博情绪知识库,并使用一种新颖的情绪知识表示方法将情绪知识融入神经网络模型中,提升了模型对文本情绪的识别能力.实验结果表明,采用本文的情绪知识融合方法,对于案件微博评论的情绪分类具有十分有效的指导作用,且该方法在公共数据集上也取得了当前已知的最好结果.当前,针对微博文本的多标签情绪分类任务是研究的热点和难点,在下一步的工作中,我们将针对案件微博评论的多标签情绪分类问题开展研究.

### 参 考 文 献

- [1] Ding Zhao-Yun, Jia Yan, Zhou Bin. Survey of data mining for microblogs. *Journal of Computer Research and Development*, 2014, 51(4): 691-706(in Chinese)  
(丁兆云, 贾焰, 周斌. 微博数据挖掘研究综述. *计算机研究与发展*, 2014, 51(4): 691-706)
- [2] Li Yong-Gan, Zhou Xue-Guang, Sun Yan, et al. Research and implementation of Chinese microblog sentiment classification. *Journal of Software*, 2017, 28(12): 3183-3205(in Chinese)  
(李勇敢, 周学广, 孙艳等. 中文微博情感分析研究与实现. *软件学报*, 2017, 28(12): 3183-3205)
- [3] Wiebe J, Wilson T, Cardie C. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 2005, 39(2): 165-210
- [4] Taboada M, Brooke J, Tofiloski M, et al. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 2011, 37(2): 267-307
- [5] Mohammad S, Kiritchenko S, Zhu X. NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets//*Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013)*. Atlanta, USA, 2013: 321-327
- [6] Zhang B, Xu X, Li X, et al. Sentiment analysis through critic learning for optimizing convolutional neural networks with rules. *Neurocomputing*, 2019, 356: 21-30
- [7] Wen S, Wan X. Emotion classification in microblog texts using class sequential rules//*Proceedings of the 28th Association for the Advance of Artificial Intelligence (AAAI 2014)*. Québec City, Canada, 2014: 187-193
- [8] Sintsova V, Pu P. Dystemo: Distant supervision method for multicategory emotion recognition in tweets. *ACM Transactions on Intelligent Systems and Technology*, 2016, 8(1): 13:1-13:22
- [9] Zhang Jing, Zhu Bo, Liang Lin-Lin, et al. Recognition and classification of emotions in the Chinese microblog based on emotional factor. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2014, 50(1): 79-84(in Chinese)  
(张晶, 朱波, 梁琳琳等. 基于情绪因子的中文微博情绪识别与分类. *北京大学学报(自然科学版)*, 2014, 50(1): 79-84)
- [10] Jiang Sheng-Yi, Huang Wei-Jian, Cai Mao-Li, et al. Building social emotional lexicons for emotional analysis on microblog. *Journal of Chinese Information Processing*, 2015, 29(6): 166-171(in Chinese)  
(蒋盛益, 黄卫坚, 蔡茂丽等. 面向微博的社会情绪词典构建及情绪分类方法研究. *中文信息学报*, 2015, 29(6): 166-171)
- [11] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality//*Proceedings of the Advances in Neural Information Processing Systems 26 (NIPS 2013)*. Lake Tahoe, USA, 2013: 3111-3119
- [12] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space. *Computing Research Repository*, arXiv preprint arXiv:1301.3781, 2013
- [13] Zhang Yang-Sen, Sun Kuang-Yi, Du Cui-Lan, et al. A cascaded construction of sentiment classifier for microblogs. *Journal of Chinese Information Processing*, 2017, 31(5): 178-184(in Chinese)  
(张仰森, 孙旷怡, 杜翠兰等. 一种级联式微博情感分类器的构建方法. *中文信息学报*, 2017, 31(5): 178-184)
- [14] Jiang F, Liu Y, Luan H, et al. Microblog sentiment analysis with emoticon space model. *Journal of Computer Science and Technology*, 2015, 30(5): 1120-1129
- [15] He Yan-Xiang, Sun Song-Tao, Niu Fei-Fei, et al. A deep learning model enhanced with emotion semantics for microblog sentiment analysis. *Chinese Journal of Computers*, 2017, 40(4): 773-790(in Chinese)  
(何炎祥, 孙松涛, 牛菲菲等. 用于微博情感分析的一种情感语义增强的深度学习模型. *计算机学报*, 2017, 40(4): 773-790)
- [16] Kim Y. Convolutional neural networks for sentence classification //*Proceedings of the 2014 Conference on Empirical Methods on Natural Language Processing (EMNLP 2014)*. Doha, Qatar, 2014: 1746-1751
- [17] Zhang Yang-Sen, Zheng Jia, Huang Gai-Juan, et al. Microblog sentiment analysis method based on a double attention model. *Journal of Tsinghua University (Science and Technology)*, 2018, 58(2): 122-130(in Chinese)  
(张仰森, 郑佳, 黄改娟等. 基于双重注意力模型的微博情感分析方法. *清华大学学报(自然科学版)*, 2018, 58(2): 122-130)
- [18] Guan Peng-Fei, Li Bao-An, Lv Xue-Qiang, et al. Attention enhanced bi-directional LSTM for sentiment analysis. *Journal of Chinese Information Processing*, 2019, 33(2): 105-111(in Chinese)  
(关鹏飞, 李宝安, 吕学强等. 注意力增强的双向 LSTM 情感分析. *中文信息学报*, 2019, 33(2): 105-111)
- [19] Xu Lin-Hong, Lin Hong-Fei, Pan Yu, et al. Constructing the affective lexicon ontology. *Journal of the China Society for Scientific and Technical Information*, 2008, 27(2): 180-185(in Chinese)

- (徐琳宏, 林鸿飞, 潘宇等. 情感词汇本体的构造. 情报学报, 2008, 27(2): 180-185)
- [20] Mei Li-Li, Huang He-Yan, Zhou Xin-Yu, et al. A survey on sentiment lexicon construction. *Journal of Chinese Information Processing*, 2016, 30(5): 19-27(in Chinese)  
(梅莉莉, 黄河燕, 周新宇等. 情感词典构建综述. 中文信息学报, 2016, 30(5): 19-27)
- [21] Turney P D, Littman M L. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 2003, 21(4): 315-346
- [22] Zhang W, Zhu Y, Wang J. An intelligent textual corpus big data computing approach for lexicons construction and sentiment classification of public emergency events. *Multimedia Tools and Applications*, 2019, 78(21): 30159-30174
- [23] Li S, Zhao Z, Liu T, et al. Initializing convolutional filters with semantic features for text classification//*Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*. Copenhagen, Denmark, 2017: 1884-1889
- [24] Luong M T, Pham H, Manning C D. Effective approaches to attention-based neural machine translation//*Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*. Lisbon, Portugal, 2015: 1412-1421
- [25] Li S, Zhao Z, Hu R, et al. Analogical reasoning on Chinese morphological and semantic relations//*Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)*. Melbourne, Australia, 2018: 138-143
- [26] Birmingham A, Smeaton A F. Classifying sentiment in microblogs: Is brevity an advantage?//*Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010)*. Toronto, Canada, 2010: 1833-1836
- [27] Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques//*Proceedings of the 2002 Conference on Empirical Methods on Natural Language Processing (EMNLP 2002)*. Philadelphia, USA, 2002: 79-86
- [28] Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. Stanford, California, USA, 2009
- [29] Graves A, Jaitly N, Mohamed A R. Hybrid speech recognition with deep bidirectional LSTM//*Proceedings of the 2013 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2013)*. Olomouc, Czech Republic, 2013: 273-278
- [30] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification//*Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)*. Berlin, Germany, 2016: 207-212
- [31] Zhao W, Ye J, Yang M, et al. Investigating capsule networks with dynamic routing for text classification//*Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*. Brussels, Belgium, 2018: 3110-3119
- [32] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules//*Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017)*. Long Beach, USA, 2017: 3859-3869
- [33] Yang Z, Yang D, Dyer C, et al. Hierarchical attention networks for document classification//*Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies (NAACL-HLT 2016)*. San Diego, USA, 2016: 1480-1489



**GUO Xian-Wei**, M. S. candidate. His research interests include natural language processing and sentiment analysis.

**LAI Hua**, M. S. , associate professor. His main research interests include intelligent information processing and machine learning.

**YU Zheng-Tao**, Ph. D. , professor. His main research interests include natural language processing, information retrieval and machine translation.

**GAO Sheng-Xiang**, Ph.D. , associate professor. Her main research interests include natural language processing, information retrieval and machine translation.

**XIANG Yan**, Ph.D. candidate, associate professor. Her main research interests include text mining and sentiment analysis.

## Background

Since the appearance of microblog, research on sentiment analysis of microblog at home and abroad has never stopped. Sentiment classification is a sub-task of text sentiment analysis. It is usually divided into subjective and objective classification, sentiment polarity classification, and emotion

classification. At present, the sentiment classification of microblog with Twitter as the main body has made very promising progress and results. However, most domestic and foreign sentiment classification researches on microblog mainly focus on subjective and objective and sentiment polarity

classification. As far as we know, there is relatively little research on emotion classification of Chinese microblog at home and abroad, and the existing research on emotion classification of Chinese microblog usually targets microblog in the general field. Existing researches are usually based on emotional dictionaries and traditional machine learning methods. The predecessors have accumulated a lot of sentiment computing resources, and some studies have mined attribute features such as emoticons and internet buzzwords for emotion classification of microblog. They are collectively called emotional knowledge. Although these methods have achieved good results, they rely too much on artificially constructed dictionaries and complex feature engineering, and there is still a large gap from practical applications. Since the emergence of distributed word vector representation methods, a large number of deep learning-based emotion classification methods have emerged. These methods benefit from neural networks that can automatically learn the feature representation of text. However, deep learning-based methods cannot effectively integrate the emotional knowledge commonly used in microblog, and the emotional knowledge has been proven to

be effective in traditional methods.

The research carried out in this article is aimed at the case-related microblog comments, which belongs to a specific field of tasks. Aiming at the above problems, this paper proposes a case-related microblog comments emotion classification method that integrates emotional knowledge. Compared with the existing methods, the method in this paper effectively uses emotional knowledge, helps the neural network model to better learn the emotion of the case-related microblog comments, and achieved the best results.

This work is supported by the National Key Research and Development Plan (No. 2018YFC0830105), the Yunnan High-tech Industry Special Project (No. 201606), and the National Natural Science Foundation of China (No. 61972186). Case public opinion monitoring is essential for courts to prevent public opinion risks. In order to address the problems in the current case public opinion monitoring, based on the needs of the courts, this work studies the emotion classification method of case-related microblog comments based on the characteristics of online public opinions involved in the case.