



DOI:10.13364/j.issn.1672-6510.20230021

## 融合文本内容和标签的中文商品评论情感分类

张路, 史艳翠

(天津科技大学人工智能学院, 天津 300457)

**摘要:** 为了解决目前基于深度学习的情感分析方法对语言知识和标签信息利用不足的问题, 本文对基于深度学习的情感分析方法加以改进, 提出一种融合文本特征和标签信息的中文文本情感分析方法。该方法首先对文本进行预处理, 然后分别使用句法规则和深度学习方法来提取文本的特征信息。在句法规则部分, 通过构建句法规则提取文本中情感倾向更加明确的信息, 并将其与标签描述信息进一步融合, 突出情感信息; 在深度学习部分, 通过结合注意力机制和门控循环神经网络提取文本的高阶特征。最后, 将上面两部分提取的特征信息进行融合并使用分类器对其进行分类, 得到文本的情感倾向。实验结果表明, 与传统情感分析方法相比, 使用本文方法进行情感分析的效果更好。

**关键词:** 情感分析; 深度学习; 句法规则; 标签信息

中图分类号: TP391 文献标志码: A 文章编号: 1672-6510(2024)01-0065-08

## Affective Classification of Chinese Product Reviews by Combining Text Content and Labels

ZHANG Lu, SHI Yancui

(College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** In order to solve the problem of insufficient utilization of language knowledge and label information in the emotion analysis methods based on deep learning, in this article we improve the emotion analysis method based on deep learning, and propose a Chinese text emotion analysis method that combines text features and label information. In our improved method we first preprocess the text, and then use syntax rules and deep learning methods to extract the feature information of the text respectively. In the part of syntax rules, we extract the more explicit information of emotion tendency in the text by constructing syntax rules, and combine it with the label description information to further highlight the emotion information. In the part of deep learning, we extract the high-level features of the text by combining the attention mechanism and gated recurrent neural network. Finally, we combine the features extracted from the above two parts and use the classifier to classify them to get the emotional orientation of the text. The experiment showed that the method proposed in this article was better than the traditional emotion analysis methods.

**Key words:** emotion analysis; deep learning; syntactic rules; label information

### 引文格式:

张路, 史艳翠. 融合文本内容和标签的中文商品评论情感分类[J]. 天津科技大学学报, 2024, 39(1): 65-72.  
ZHANG L, SHI Y C. Affective classification of Chinese product reviews by combining text content and labels[J]. Journal of Tianjin university of science & technology, 2024, 39(1): 65-72.

随着科学技术的发展, 互联网与人们的日常生活越来越密切, 越来越多的用户开始在互联网上发表个人见解、表达自己的情感和意见。利用情感分析技术

对微博评论或淘宝评价等文本进行分析, 可以从中发现用户对舆论的关注程度以及对商品的购买喜好, 从而为用户推荐更合适的信息与商品<sup>[1]</sup>。

收稿日期: 2023-02-07; 修回日期: 2023-05-26

基金项目: 天津市教委理工类基本科研业务费项目(2018KJ105); 天津市科技计划项目(22KJXMRC00210)

作者简介: 张路(1998—), 男, 安徽阜阳人, 硕士研究生; 通信作者: 史艳翠, 副教授, syc@tust.edu.cn

情感分析是自然语言处理的一个分支,是指通过制定相关的规则和搭建合适的模型对微博评论、淘宝商品评价等带有情感、意见、看法的文本进行分析处理,进而判断出这些文本中包含的用户对于某种事物、产品或服务所表达出的情感、意见、情绪<sup>[2]</sup>。目前的情感分析方法主要可以总结为4种:基于情感词典的方法、基于机器学习的方法、基于深度学习的方法和基于多策略融合的方法<sup>[3]</sup>。

传统的文本情感分析方法主要包括基于情感词典的方法和基于机器学习的方法。基于情感词典的方法通过对文本中的情感词赋值,然后计算文本整体的情感得分,判断文本整体的情感极性<sup>[4-6]</sup>,但该方法需要大量的人工参与。基于机器学习的方法则是通过使用机器学习算法抽取文本的特征,学习文本的一般性规律,然后使用这些规律对未知文本进行预测,判断出文本的情感倾向<sup>[7-10]</sup>。与基于情感词典的方法相比,使用基于机器学习的方法进行情感分析的效率更高,但同样需要人工参与。

随着深度学习的兴起,研究人员开始尝试将深度学习技术引入文本情感分析领域中,深度学习技术在文本情感分析任务中逐渐占据主导地位<sup>[11-13]</sup>。刘龙飞等<sup>[14]</sup>使用卷积神经网络(convolutional neural networks, CNN)分别对文本的字级别特征和词级别特征进行处理,并判断文本的情感极性,两部分的实验皆取得了良好的成绩,表明CNN可以用来处理文本信息,适用于情感分类。滕飞等<sup>[15]</sup>通过改进的长短期记忆网络(long short-term memory, LSTM)模型对微博评论进行分析,分类准确率达到96.5%。为了突出文本中部分关键特征的重要性,研究人员提出了注意力机制并将其和神经网络进行融合。Zhai等<sup>[16]</sup>将注意力机制和双向长短期记忆网络(bidirectional long short-term memory, Bi-LSTM)结合,用于双语的文本情感分析,并取得了很好的效果。与传统的文本情感分析方法相比,以上基于深度学习模型的方法具有更强的适应性和扩展性,但这些方法忽略了句法规则等语言知识在中文文本中的作用。

为了体现语言知识对中文情感分类的重要性,有研究人员将情感词典融入深度学习模型中,进而提高情感分析的准确率<sup>[17-20]</sup>。邱宁佳等<sup>[21]</sup>将根据电商评论数据集构建的3种情感词典和神经网络模型进行融合,提出一种融合了语法规则的中文情感分类模型,但该模型并没有将CNN和Bi-LSTM很好地结合起来。张仰森等<sup>[22]</sup>通过整理现有的情感词典构建

了微博情感符号库,并采用Bi-LSTM模型对文本进行编码,获取文本的语义信息,实验表明该模型在情感分析上表现良好,但该方法所构建的情感符号库过于复杂,有一定的难度,且只使用了Bi-LSTM编码文本语义信息,模型泛化能力不足。

现有的研究工作表明文本标签所携带的信息对情感分析的结果有积极作用。Sheng等<sup>[23]</sup>定义了一组情感词典,并将它作为文本的情感标签。王嫒等<sup>[24]</sup>利用注意力机制将标签信息和文本表示进行融合,取得了不错的结果。然而,一些数据集中只有简单的标签类别而没有给出具体的标签信息,导致模型不能准确理解标签代表的具体含义。

相对于只使用单一的机器学习模型或深度学习模型进行情感分析,将情感词典融入机器学习或深度学习中进行情感分析更具有优势,取得的效果更好。本文通过融合句法规则和深度学习构建了情感分析模型,分别使用句法规则和深度学习提取文本的特征。在深度学习部分通过结合注意力机制和门控循环单元(gate recurrent unit, GRU)模型提取文本的高阶特征;在句法规则部分通过构建句法规则降低文本复杂度,然后将句法规则模块生成的信息与生成的标签描述信息进行融合,进一步突出情感信息;再将两部分提取的特征进行融合,最后使用分类器对文本进行情感分类。

本文的主要贡献如下:

(1)提出了一种融合句法规则和深度学习的情感分析模型。该模型通过使用句法规则降低文本复杂度,结合深度学习在提取文本特征上的优势,在情感分类任务上取得了较好的结果。

(2)在模型中使用了改进的GRU模型,在提取文本上下文信息的同时捕获了文本中的重点信息,提高了模型的特征表达能力。

(3)构建领域标签描述信息,在模型中引入标签信息,使用标签信息为模型带来全局分类信息,在降低文本复杂度的同时降低文本中噪声对训练结果的影响。

## 1 算法描述

### 1.1 问题描述

中文文本情感分析是情感分析任务中的一个子任务,通过对数据集进行去停用词、分词等预处理操作将数据转为结构化信息,接着将数据表示成向量后

输入到构建的模型中,得到和文本对应的标签。在本文中,数据集  $D = \{(x_i, y_j)\}$  是由  $N$  个文本  $x_i$  以及各个文本相对应的标签  $y_j$  组成。其中,  $i \in (1, N), j \in (1, D), N$  表示数据集中文本的数量,  $D$  表示标签类别的数量。每个文本的词向量初始化表示为  $x_i = (\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n)$ , 其中  $\bar{e}_n \in R^k$ ,  $n$  为文本的长度,  $k$  为词向量的维度。经过句法规则处理的文本的词向量初始化表示为  $x'_i = (\bar{e}'_1, \bar{e}'_2, \dots, \bar{e}'_n)$ , 其中  $\bar{e}'_n \in R^k$ 。情感分析的目的是训练一个模型,使其可以对一个新的文本  $x$  判断其情感倾向  $y$ 。

## 1.2 模型框架

本文提出的模型整体框架如图1所示,模型的整体结构由句法规则构建、特征提取、交互模块和特征融合组成。模型结构自底向上划分为5个部分:输入层、嵌入层、神经网络模型层、特征融合层以及输出层。

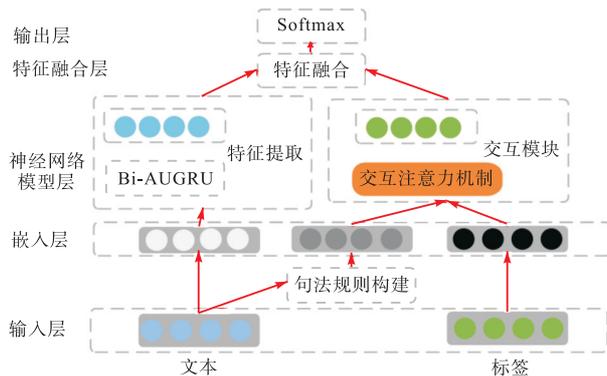


图1 模型整体结构图

Fig. 1 Overall framework of the model

### 1.2.1 句法规则构建

通过对中文文本的研究可以发现,文本中部分句子的情感倾向代表了文本整体的情感倾向。当文本中出现情感倾向明显的情感词时,该情感词所在分句的情感倾向往往代表了文本整体的情感倾向。当文本中含有总结词时,总结词所引导的句子总会表明文本的中心思想。当文本中含有转折词时,一般分为两种情况:当文本中出现的是甲类转折词时,该转折词会改变文本的情感倾向;当文本中出现的是乙类转折词时,该转折词起到过渡作用,不会对文本的情感倾向产生影响。

邱宁佳等<sup>[21]</sup>通过对数据集的情感词、转折词和总结词进行提炼构建3种词典:情感倾向词典、转折词词典、总结词词典,然后根据这3种词典设计句法规则降低中文文本复杂度,突出文本情感倾向。本

文对上述方法进行优化,在大连理工大学的中文情感词本体库的基础上结合本文数据集的特点构建了3类词典,并根据这些词典设计句法规则对数据集进行处理,保留文本中情感倾向突出的部分,降低文本中噪声对训练结果的影响。令  $x$  表示中文文本,  $x'$  表示经过句法规则处理过的文本。本文设计的句法规则如下:

**规则1** 判断文本中是否有情感词典中的情感词,若文本中所有情感词的情感倾向全部相同,则直接提取情感词所在的分句。

**规则2** 若文本中存在情感极性相反的情感词,则保留原文本内容。

**规则3** 若文本中不存在情感词,则判断文本中是否存在总结词,若存在,则保留总结词所引导的文本。

**规则4** 当前3个规则都不满足时,若文本中出现转折词,则先判定转折词的类别,再根据转折词的类别处理文本。当转折词是甲类转折词时,提取转折词之后的文本;当转折词是乙类转折词时,提取该转折词所在句子之外的所有文本。

**规则5** 若文本均不符合以上4条规则,则直接保留原文本内容。

经过句法规则处理后的文本  $x'$  相较于原文本  $x$  来说通常比较简短且结构简单,情感倾向突出,有利于提升中文文本情感分析的效果。

### 1.2.2 标签描述信息的生成

对于情感分析任务,标签对模型的最终表现起着重要的作用。然而,一些数据集中只有简单的标签类别而没有给出具体的标签信息,导致模型不能理解标签代表的具体含义。另外,一些数据集同时包含多个领域内的数据集,每个领域关注的内容不同。例如同样是“苹果”一词,在手机领域更关注手机内存大小、屏幕好坏,而在水果领域则更关注水果的味道、新鲜度等。故而即便两个领域间数据的标签相同,相应的标签描述信息也应有所区别。

针对这些问题,本文提出了一种特定领域的标签描述信息的生成方法,为每个领域的标签生成特定的描述信息。将标签描述信息引入交互模块中,对标签信息和文本信息进行融合,进而为模型带来全局分类信息线索。

合理的标签描述应能准确描述该标签类别下文本的特点且与其他标签的描述信息有较大区别。对于语料库  $D$  包含的所有领域  $\{D_1, D_2, \dots, D_a\}$ , 以其

中一个领域语料库  $D_a$  中的标签  $y_j$  为例, 构建标签  $y_j$  的标签描述信息。为了使标签  $y_j$  的标签描述信息和标签  $y_j$  具有较高的相关性, 本文使用词频-逆向文件频率 (term frequency-inverse document frequency, TF-IDF) 算法计算单词  $w$  和标签  $y_j$  的相关性, TF-IDF 算法可以通过统计词语在文本中出现的次数判断该词语在文本中的重要程度。计算公式为

$$T_{w,y_j} = \sum_{d \in D_a^{y_j}} f_{w,d} \times \ln \left( \frac{|D_a^{y_j}|}{f_{w,D_a^{y_j}}} \right) \quad (1)$$

其中:  $T_{w,y_j}$  表示单词  $w$  在语料库  $D_a^{y_j}$  中的重要程度,  $|D_a^{y_j}|$  表示领域语料库  $D_a$  中标签为  $y_j$  的文本的数量,  $D_a^{y_j}$  由领域语料库  $D_a$  中标签为  $y_j$  的所有文本组成,  $w$  为语料库  $D_a^{y_j}$  中的一个单词,  $d$  为  $D_a^{y_j}$  中的一个评论文本,  $f_{w,d}$  表示单词  $w$  在  $d$  中出现的次数,  $f_{w,D_a^{y_j}}$  表示领域语料库  $D_a^{y_j}$  中出现单词  $w$  的文本数量。

为了让标签描述信息对其他标签具有区分度, 需要评估单词  $w$  和所有情感标签的相关性, 单词  $w$  和所有标签的相关性越高, 说明单词  $w$  对所有标签的区分度越小。计算公式为

$$L_{w,y_j} = \sum_{d_L \in D_L} f_{w,d_L} \times \ln \left( \frac{|D_L|}{f_{w,D_L} + 1} \right) \quad (2)$$

其中:  $L_{w,y_j}$  表示单词  $w$  在语料库  $D_L$  中的重要程度,  $|D_L|$  表示语料库  $D_L$  中文本的数量,  $D_L$  由领域语料库中标签不为  $y_j$  的所有数据组成,  $d_L$  为  $D_L$  中的一个评论文本,  $f_{w,d_L}$  表示单词  $w$  在  $d_L$  中出现的次数,  $f_{w,D_L}$  表示语料库  $D_L$  中出现单词  $w$  的文本数量。通过上面两个公式, 即可得到单词  $w$  相对于标签  $y_j$  的相关性分数, 为

$$r_{w,y_j} = \frac{T_{w,y_j}}{L_{w,y_j}} = \frac{\sum_{d \in D_a^{y_j}} f_{w,d} \times \ln \left( \frac{|D_a^{y_j}|}{f_{w,D_a^{y_j}}} \right)}{\sum_{d_L \in D_L} f_{w,d_L} \times \ln \left( \frac{|D_L|}{f_{w,D_L}} \right)} \quad (3)$$

从式(3)可以看出, 当单词  $w$  和标签  $y_j$  的相关性越高时, 单词  $w$  相对于标签  $y_j$  的相关性分数越大, 越适合成为标签  $y_j$  的描述信息; 当单词  $w$  和其他标签的相关性越高时, 单词  $w$  相对于标签  $y_j$  的相关性分数越小, 越不适合成为标签  $y_j$  的描述信息。

选取  $r_{w,y_j}$  得分最高的  $M$  个词  $[w_{1,y_j}, w_{2,y_j}, \dots, w_{M,y_j}]$  作为标签  $y_j$  的描述, 然后使用词向量矩阵将标签描述的每个词嵌入词向量进而得到领域语料库  $D_a$  内标签  $y_j$  的描述信息的向量表示  $\mathbf{q}_{a,y_j}$ 。

同理, 通过上述方法得到其他领域内标签  $y_j$  的标签描述信息  $\{q_{1,y_j}, q_{2,y_j}, \dots, q_{g,y_j}\}$ , 其中  $g$  为语料库中领域的数量,  $\mathbf{q}_{g,y_j}$  为领域  $g$  中标签  $y_j$  的描述信息的词向量表示。

最后计算得到其他标签的表示, 得到标签表示矩阵  $\mathbf{C} \in \mathbf{R}^{g \times m \times k}$ , 其中  $m$  为标签类别的数量,  $g$  为语料库中领域的数量,  $k$  为词向量的维度。

$$\mathbf{C} = (q_{1,1}, q_{1,2}, \dots, q_{1,m}, q_{2,1}, \dots, q_{g,m}) \quad (4)$$

### 1.2.3 特征提取

为了提高模型提取文本特征的能力, 在特征提取模块中使用 Bi-AUGRU 模型<sup>[25]</sup>。Bi-AUGRU 网络模型是以门控循环神经网络为基础, 使用注意力分数对门控循环神经网络的更新门进行加权, 以此在保留门控循环神经网络提取文本上下文信息的能力的基础上, 提升了模型提取文本中重点信息的能力。Bi-AUGRU 模型的公式为

$$\mathbf{z}_t = \sigma(\mathbf{w}_z \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (5)$$

$$\mathbf{z}'_t = \mathbf{a}_t \cdot \mathbf{z}_t \quad (6)$$

$$\mathbf{r}_t = \sigma(\mathbf{w}_r \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (7)$$

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{w}_{\tilde{h}} \cdot [\mathbf{r}_t \cdot \mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (8)$$

$$\mathbf{h}'_t = (1 - \mathbf{z}'_t) \cdot \mathbf{h}_{t-1} + \mathbf{z}'_t \cdot \tilde{\mathbf{h}}_t \quad (9)$$

其中:  $\mathbf{x}_t$  是文本的向量化表示,  $\mathbf{z}_t$  是 AUGRU 的原始更新门,  $\mathbf{r}_t$  是 AUGRU 的重置门,  $\mathbf{z}'_t$  是 AUGRU 的注意力分数优化后的更新门,  $\mathbf{h}'_t$  和  $\mathbf{h}_{t-1}$  是 AUGRU 的隐藏状态,  $\tilde{\mathbf{h}}_t$  为候选隐藏状态,  $\mathbf{a}_t$  是将文本词向量输入到注意力机制中得到的注意力分数, 其计算公式为

$$\mathbf{a}_t = \text{Softmax}(s(\mathbf{h}_t, \mathbf{q})) = \frac{\exp(s(\mathbf{h}_t))}{\sum_{j=1}^n \exp(s(\mathbf{h}_j, \mathbf{h}_t))} \quad (10)$$

其中:  $\mathbf{h}_t$  为输入向量,  $s(\mathbf{h}_t, \mathbf{q})$  为打分函数。

为了进一步提高模型的准确度, 将 AUGRU 模型双向化后得到 Bi-AUGRU 模型, 计算公式为

$$\bar{\mathbf{h}}_t = \text{AUGRU}(\mathbf{x}_t) \quad (11)$$

$$\tilde{\mathbf{h}}_t = \text{AUGRU}(\mathbf{x}_t) \quad (12)$$

其中:  $\mathbf{x}_t$  为文本的向量化表示,  $\bar{\mathbf{h}}_t$  表示前向传播时文本的隐层状态,  $\tilde{\mathbf{h}}_t$  表示反向传播时文本的隐层状态。将两种状态进行拼接, 即可得到文本的双向语义信息  $\mathbf{h}_t = (\bar{\mathbf{h}}_t, \tilde{\mathbf{h}}_t)$ 。

### 1.2.4 交互模块

为了突出标签描述信息和情感信息对情感分类的作用, 使用注意力机制对 1.2.2 节得到的标签描述信息表示矩阵与情感文本隐藏状态表示矩阵进行交互处理, 进而得到经过标签信息增幅具有分类指导信

息的情感文本的新表示。具体计算公式为

$$h_q = \text{Softmax} \left( \frac{qk^T}{\sqrt{d_w}} \right) v \quad (13)$$

$$[q, k, v] = [CW_q, x'W_k, x'W_v] \quad (14)$$

其中:  $q, k, v$  为查询、键、值矩阵;  $C$  为标签描述信息的隐层表示,  $x'$  为情感文本的隐层表示;  $W_q, W_k, W_v$  为可训练的参数矩阵。  $d_w$  是  $k$  的维度。

### 1.2.5 特征融合

通过上文得到了文本的两种隐层表示  $h_t$  和  $h_q$ , 但两者的侧重点不同。  $h_t$  侧重于对文本上下文信息的利用, 而  $h_q$  则更关注对文本中情感信息和标签信息的利用。为了在提取到文本特征信息的同时突出文本中的标签信息和情感信息的作用, 本文在特征融合阶段将  $h_t$  和  $h_q$  进行拼接作为最终的词向量表示, 最终得到文本表示为

$$h = h_t \oplus h_q \quad (15)$$

将  $h$  输入到全连接层中并引入 Dropout 机制防止模型出现过拟合现象。最后使用分类器对文本表示进行分类, 得到文本的情感倾向, 全连接层的隐藏单元个数等于分类类别数。

### 1.2.6 输出层

将文本的最终向量表示使用 Softmax 函数生成情感标签上的概率分布, 为

$$y' = \text{Softmax}(w_{\text{out}}h' + b_{\text{out}}) \quad (16)$$

其中:  $w_{\text{out}}$  和  $b_{\text{out}}$  为权重矩阵和偏置。

训练阶段使用交叉熵损失函数对模型进行训练, 为

$$L = \sum_{(x,y) \in D} y \log y' \quad (17)$$

其中:  $x$  为目标句子,  $y$  为句子真实标签,  $y'$  为预测的句子标签,  $D$  为整体语料库。

## 2 实验与结果分析

### 2.1 实验环境

本文的实验是在 Linux 操作系统下使用 Python 语言编写完成的, 开发过程中所使用的深度学习框架为 TensorFlow, 实验所使用 GPU 为 RTX 1080。

### 2.2 数据集

采用的实验数据为 online\_shopping\_10\_cats 数据集, 该数据集共 62 774 条电商购物评论数据, 数据分为两类: 一类为积极评论数据, 共有 31 728 条; 一类为消极评论数据, 共有 31 046 条。数据集包含牛

奶、酒店、衣服、热水器等 10 个领域, 具体分布见表 1。本文对数据集进行进一步处理, 根据数据集包含的领域数量和标签数量为每个领域内的标签生成最合适的标签描述信息。本文中实验数据的训练集与测试集的比例设置为 8 : 2。

表 1 数据集统计信息

Tab. 1 Datasets statistics

类别	评论数		类别	评论数	
	积极	消极		积极	消极
书籍	2 100	1 751	热水器	475	100
平板	5 000	5 000	牛奶	992	1 041
手机	1 165	1 158	衣服	5 000	5 000
水果	5 000	5 000	计算机	1 996	1 996
洗发水	5 000	5 000	酒店	5 000	5 000

本文在大连理工大学的中文情感词本体库的基础上结合本文数据集的特点构建了 3 类词典, 词典中的部分关键词为:

(1) 正面情感倾向词: 推荐、方便、值、赞、不错、棒、好评、好用、满意、很好、可以、喜欢等。

(2) 负面情感倾向词: 太少、差评、不满意、失望、别买、差、不好、很差、垃圾、贵、骗人、坑等。

(3) 总结词: 总的来说、总之、总的感觉、总体、结果、综上所述、个人认为、反正、整体等。

(4) 甲类转折词: 但、可是、却、然而、所以等。

(5) 乙类转折词: 只是、而且、就是、虽然等。

### 2.3 参数设置

本文的网络模型中所涉及的参数设置: 训练轮数 36, 批大小 100, 学习率 0.001, 词向量维度 300, 隐藏层维度 300, 优化器 Adam, Dropout 0.5, 损失函数为交叉熵损失函数, 标签描述信息的长度  $M$  值为 30。

### 2.4 评价指标

本文采用准确率 ( $A$ )、召回率 ( $R$ )、精确率 ( $P$ ) 以及  $F_1$  值作为实验评价指标。本文的任务属于二分类任务, 用来计算评价指标的分类类别矩阵见表 3。

表 2 分类类别矩阵

Tab. 2 Classification matrix

预测结果	真实结果	
	积极	消极
积极	$r$	$t$
消极	$s$	$z$

准确率 ( $A$ ) 表示所有情感极性都被正确分类的样本占总样本的比值, 为

$$A = (r + z) / (r + s + t + z) \quad (18)$$

召回率 ( $R$ ) 表示正样本中正确预测的样本数

量,为

$$R = r / (r + s) \tag{19}$$

精确率( $P$ )表示在所有被预测为正的样本中实际为正样本的概率,为

$$P = r / (r + t) \tag{20}$$

$F_1$  值是对精确率与召回率加权调和得到的结果, $F_1$  值越大,实验结果越好,其计算公式为

$$F_1 = 2 \times P \times R / (P + R) \tag{21}$$

### 2.5 实验结果分析

#### 2.5.1 随机失活率对模型性能的影响

为防止模型在训练过程中出现过拟合现象,本文在模型中加入 Dropout,并通过实验记录了 Dropout 值对模型性能的影响,实验结果如图 2 所示。从图 2 可以看出,随着 Dropout 值的增加,准确率先升高后降低,当 Dropout 值为 0.5 时,模型性能最佳。这是由于过低的随机失活率不能缓解模型的过拟合现象,而过高的随机失活率又会导致模型欠拟合。

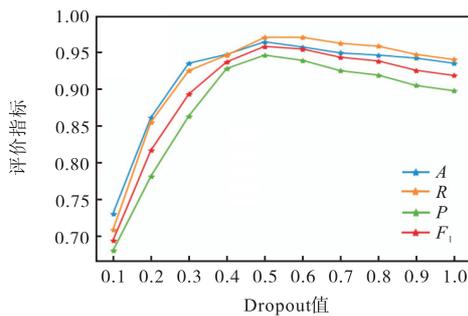


图 2 随机失活率对模型性能的影响

Fig. 2 Effect of random deactivation rate on model performance

#### 2.5.2 训练轮数对模型性能的影响

实验过程中模型的训练轮数对模型性能的影响较大,图 3 展示了随着训练轮数的增加,模型性能的变化。

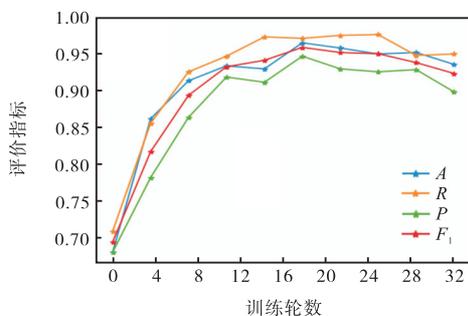


图 3 训练轮数对模型性能的影响

Fig. 3 Effect of number of epoch on model performance

从图 3 可以看出,随着训练轮数的增加,模型的

准确率不断上升,在训练轮数为 20 时,模型的准确率达到最高,随后模型准确率逐步下降。这是因为当模型的训练轮数较少时,模型不能完全拟合数据,而当训练轮数过多时,模型表达能力过剩,就会出现过拟合现象,导致模型的性能下降。

#### 2.5.3 M 值对模型性能的影响

为了探究标签描述信息的长度对模型性能的影响,实验结果如图 4 所示。从 4 图可以看出,随着  $M$  值增大,模型的性能先升高后降低,在  $M$  值为 30 时,模型表现出最佳的性能。这主要是因为当  $M$  取值比较小时,用来描述标签的词语过少,标签描述信息无法充分描述标签的含义;当  $M$  取值过大时,过长的描述信息导致一些不具有代表性的词语进入描述信息中进而产生噪声,影响了模型性能。

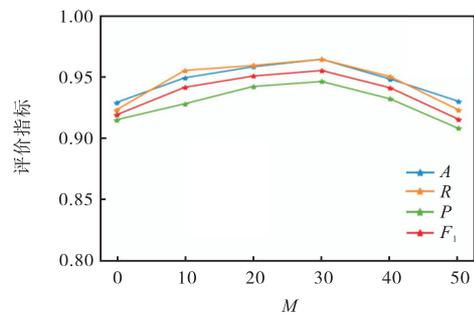


图 4 M 值对模型性能的影响

Fig. 4 Effect of M value on model performance

#### 2.5.4 消融实验

为了验证本文模型中的句法规则和标签描述可以提高中文文本情感分析的效果,本节通过设计消融实验验证了模型中句法规则部分和标签描述生成部分对最后分类结果的影响。首先在保证本文模型中参数不变的前提下对文本模型进行处理:

(1) 去除本文模型中的句法规则模块得到模型 Model\_L。

(2) 去除本文模型中的标签描述生成模块得到模型 Model\_R。

(3) 同时去除本文模型中的句法规则模块和标签描述生成模块得到模型 Model\_D。其中模型 Model\_R 用来验证句法规则模块对最后分类结果的影响,Model\_L 用来验证标签描述生成模块对最后分类结果的影响。

在上述模型中使用表 1 中的数据集进行实验,将取得的实验结果与本文模型的实验结果进行比较,消融实验结果见表 4。

通过比较模型 Model\_R、Model\_L 和 Model\_D

的实验结果,可以观察到句法规则信息和标签描述信息对提高中文文本的情感分类效果有积极的作用。由表4可知,在比较的几种模型中,本文模型获得的分类效果最好,说明其能够将句法规则和标签信息很好地融入模型中。与Model\_L相比,本文模型的准确率提高0.018,召回率提高0.016, $F_1$ 值提高0.010;与Model\_R相比,本文模型的准确率提高0.030,召回率提高0.027, $F_1$ 值提高0.019。将Model\_L和Model\_R的实验结果比较可以发现,相对于句法规则部分,标签描述生成模块对中文文本情感分类的影响更大。

表4 消融实验结果

Tab. 4 Results of ablation experiments

模型	准确率	精确率	召回率	$F_1$
Model_D	0.920	0.917	0.921	0.919
Model_R	0.933	0.935	0.942	0.938
Model_L	0.945	0.941	0.953	0.947
本文模型	0.963	0.945	0.969	0.957

### 2.5.5 本文模型与其他模型的性能对比

为了进一步验证本文模型的有效性,将本文模型与以下几篇文献模型在表1数据集上进行对比实验,进而验证本文模型在情感分类任务上的优越性。分类结果对比见表5。

表5 分类结果对比

Tab. 5 Comparison of classification results

模型	准确率	精确率	召回率	$F_1$
CB_Rule	0.951	0.946	0.960	0.953
TLE	0.948	0.947	0.953	0.950
SVW-BERT	0.958	0.952	0.958	0.955
AW Ave	0.946	0.941	0.947	0.944
本文模型	0.963	0.945	0.969	0.957

CB\_Rule<sup>[21]</sup>: 邱宁佳等<sup>[21]</sup>提出的中文情感分析模型,将根据情感词典构建的语法规则和神经网络模型进行融合,在中文情感分析任务中取得了优秀的表现。

TLE<sup>[24]</sup>: 王嫒等<sup>[24]</sup>提出的一种融合文本内容和标签引导文本编码的文本分类方法,通过使用自注意力机制得到文本表示,再使用交互注意力机制得到融合了标签信息的文本表示,最后将两种文本表示融合后输入到分类器中得到文本的情感倾向。

SVW-BERT<sup>[26]</sup>: 张小艳等<sup>[26]</sup>提出的情感分析方法。通过融合文本的字、词级别文本表示得到文本的最终表示,并使用文本的情感值对其进行加权融合,构建情感值加权融合字词向量的文本表示,然后将该表示输入到分类器中得到文本的情感倾向。

AW Ave<sup>[23]</sup>: Sheng 等<sup>[23]</sup>提出的同时学习词汇和

情感标签信息的嵌入方法。该方法使用1组和标签相关的词汇表示情感标签信息。

由表5可以看出,对于中文文本的情感分类任务,本文模型整体上优于其他代表性的深度学习模型。这是因为模型CB\_Rule和模型SVW-BERT只使用了句法规则信息,没有利用标签描述信息,而模型TLE和模型AW Ave只使用了标签描述信息,没有使用句法规则信息。本文模型不仅使用了句法规则信息降低了文本复杂度,而且使用了标签描述信息为模型带来全局分类信息,提高了模型的性能。

以上实验结果表明,本文提出的模型能够很好地利用中文文本中的句法规则信息和标签描述信息,同时结合深度学习提取文本的深层特征,在总体性能上优于其他模型。

## 3 结语

针对目前中文文本情感分析注重使用深度学习模型进行情感分析而忽略了使用句法规则和标签信息等知识的问题,本文提出了一种融合了句法规则和标签描述信息的中文文本情感分析模型。首先使用句法规则对文本进行处理,提取情感倾向更加明确的信息,并将其与标签描述信息进行融合,进一步突出情感信息,然后将得到的隐层特征与深度学习部分提取的文本特征进行融合,最后将得到的文本表示输入分类器中,得到文本的情感极性。通过设置消融实验,探究了本文模型中各个模块对情感分析结果的影响,将本文模型与近几年文献中的模型进行比较,验证了本文模型在中文文本情感分析任务上的有效性。

在未来的工作中,可以尝试对标签描述信息生成部分的模型进行改进,从而获得更加合理的标签描述信息,并尝试对模型的深度学习部分进行优化从而进一步提高模型分类效果。

## 参考文献:

- [1] 李然,林政,林海伦,等. 文本情绪分析综述[J]. 计算机研究与发展,2018,55(1):30-52.
- [2] CHAIMA M, ZAHIA G, LOTFI B R. Opinion mining in online social media: a survey[J]. Social network analysis and mining, 2022, 12(1):1-18.
- [3] 钟佳娃,刘巍,王思丽,等. 文本情感分析方法及应用综述[J]. 数据分析与知识发现,2021,5(6):1-13.
- [4] CAI Y, YANG K, HUANG D P, et al. A hybrid model for opinion mining based on domain sentiment dictionary

- ary[J]. International journal of machine learning and cybernetics, 2019, 10(8): 2131–2142.
- [5] XU G X, YU Z H, YAO H S, et al. Chinese text sentiment analysis based on extended sentiment dictionary[J]. IEEE Access, 2019, 7: 43749–43762.
- [6] 梅莉莉, 黄河燕, 周新宇, 等. 情感词典构建综述[J]. 中文信息学报, 2016, 30(5): 19–27.
- [7] RUZ G A, HENRIQUEZ P A, MASCARENO A. Sentiment analysis of twitter data during critical events through Bayesian networks classifiers[J]. Future generation computer systems, 2020, 106: 92–104.
- [8] 刘丽, 岳亚伟. 面向高校学生微博的跨粒度情感分析[J]. 计算机应用研究, 2019, 36(6): 1618–1622.
- [9] 唐莉, 刘臣. 基于 CRF 和 HITS 算法的特征情感对提取[J]. 计算机技术与发展, 2019, 29(7): 71–75.
- [10] KAUR S, SIKKA G, AWASTHI L K. Sentiment analysis approach based on N-gram and KNN classifier[C]//IEEE. 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC). New York: IEEE, 2018: 8703350.
- [11] DO H H, PRASAD P, MAAG A, et al. Deep learning for aspect-based sentiment analysis: a comparative review[J]. Expert systems with applications, 2019, 118: 272–299.
- [12] YADAV A, VISHWAKARMA D K. Sentiment analysis using deep learning architectures: a review[J]. Artificial intelligence review, 2020, 53(6): 4335–4385.
- [13] NIU C M, ZHAN G H, LI Z H. Chinese weibo sentiment analysis based on deep neural network[J]. Computer systems & applications, 2018, 27: 205–210.
- [14] 刘龙飞, 杨亮, 张绍武, 等. 基于卷积神经网络的微博情感倾向性分析[J]. 中文信息学报, 2015, 29(6): 159–165.
- [15] 滕飞, 郑超美, 李文. 基于长短期记忆多维主题情感倾向性分析模型[J]. 计算机应用, 2016, 36(8): 2252–2256.
- [16] ZHAI S P, YANG Y Y, QIU C, et al. Bilingual text sentiment analysis based on attention mechanism Bi-LSTM[J]. Computer applications and software, 2019, 36(12): 251–255.
- [17] 陈珂, 谢博, 朱兴统. 基于情感词典和 Transformer 模型的情感分析算法研究[J]. 南京邮电大学学报(自然科学版), 2020, 40(1): 55–62.
- [18] 吴杰胜, 陆奎, 王诗兵. 基于多部情感词典与 SVM 的电影评论情感分析[J]. 阜阳师范学院学报(自然科学版), 2019, 36(2): 68–72.
- [19] 徐康庭, 宋威. 结合语言知识和深度学习的中文文本情感分析方法[J]. 大数据, 2022, 8(3): 115–127.
- [20] CHEN L C, LEE C M, CHEN M Y. Exploration of social media for sentiment analysis using deep learning[J]. Soft computing, 2020, 24: 8187–8197.
- [21] 邱宁佳, 王晓霞, 王鹏, 等. 融合法语规则的双通道中文情感模型分析[J]. 计算机应用, 2021, 41(2): 318–323.
- [22] 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[J]. 清华大学学报(自然科学版), 2018, 58(2): 122–130.
- [23] SHENG Y, TAKASHI I. Joint embedding of words and labels for sentiment classification[J]. International journal of Asian language processing, 2021, 31(1): 2150005.
- [24] 王嫻, 周宇博, 徐涛, 等. 融合文本内容和标签引导文本编码的分类方法[J]. 华中科技大学学报(自然科学版), 2021, 49(8): 92–97.
- [25] ZHOU G, MOU N, FAN Y, et al. Deep interest evolution network for click-through rate prediction[C]//AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI, 2019: 5941–5948.
- [26] 张小艳, 白瑜. 基于加权融合字词向量的中文在线评论情感分析[J]. 计算机应用研究, 2022, 39(1): 31–36.

责任编辑: 郎婧