

# 基于注意力与门控机制相结合的细粒度情感分析



张瑾 段利国 李爱萍 郝晓燕

太原理工大学信息与计算机学院 太原 030024

(295511703@qq.com)

**摘要** 细粒度情感分析(fine-grained sentiment analysis)是自然语言处理领域的关键问题之一,其通过学习文本的上下文信息来进行特定方面的情感分析,可以帮助用户和商家更好地了解用户评论特定方面的情感。针对基于用户评论的方面级别细粒度情感分析任务,提出了 BiGRU-Attention 与门控机制(gated mechanisms)相结合的文本情感分类模型。首先,通过整合现有的情感资源,将 HOWNET 评价情感词典作为种子情感词典,利用 SO-PMI 算法扩充用户评论情感词典,结合否定词典以及词性信息扩充用户评论情感知识,将用户评价情感知识作为用户评论情感特征信息;其次,引入字词特征与情感特征信息,将它们联合作为模型输入,使用 BiGRU 对文本进行深层次的特征提取;然后,结合门控机制以及注意力机制,根据获取的方面词信息进一步提取与方面词相关的上下文情感特征信息;最后,在输出层进行文本情感分析,经过 softmax 获得最终的情感极性。在 AIchallenger2018 细粒度情感分析中文数据集上,所提模型的 Macro\_F1\_score 值达到了 0.7218,性能超过基线系统,获得了较好的实验结果。

**关键词:**深度学习;情感分析;注意力机制;门控机制;BiGRU

**中图法分类号** TP391

## Fine-grained Sentiment Analysis Based on Combination of Attention and Gated Mechanism

ZHANG Jin, DUAN Li-guo, LI Ai-ping and HAO Xiao-yan

College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China

**Abstract** The fine-grained sentiment analysis is one of the key problems in the area of natural language processing. By learning contextual information of the text to conduct sentiment analysis on specific aspects, it can help users and businesses to better understand the sentiment information of specific aspects of users' comments. Aiming at the task of fine-grained sentiment analysis on users' comments, a text sentiment classification model combining BiGRU-attention and Gated Mechanisms is proposed. By integrating existing sentiment resources, HOWNET evaluation sentiment dictionary is used as the seed sentiment dictionary to expand the user comment sentiment dictionary through SO-PMI algorithm, the negative dictionary and part of speech information are combined to expand the user comment sentiment knowledge as the users' comment sentiment characteristic information. Introducing word, character and sentiment characteristics as the model of input information, and using BiGRU to extract deep text features, then combined with gated mechanism as well as the attention mechanism, according to the acquired aspect word information to further extract the contextual sentiment characteristics related to aspect words, the final sentiment polarity is obtained by the softmax classifier. Experimental results show that the proposed model achieves better experimental results on the AIchallenger 2018 fine-grained sentiment analysis Chinese data sets, the Macro\_F1\_score value reaches 0.7218, and the performance exceeds the baseline system.

**Keywords** Deep learning, Sentiment analysis, Attention mechanism, Gated mechanism, BiGRU

## 1 引言

对用户生成的评论进行意见挖掘和情感分析可以为消费者和用户提供有价值的信息。情感分析又称为评论挖掘或意见挖掘,其主要任务是以文本中带有主观情感的词语、句子为基础,分析文本语义,识别情感极性。让计算机自动识别文本

的情感极性一直是自然语言处理领域的研究热点<sup>[1]</sup>。在 Web2.0 的时代<sup>[2]</sup>背景下,互联网用户不仅是网页的浏览者,更是网页内容的参与者。用户评论作为用户生成数据的一种主要形式,对于消费者以及商家来说都有着重要意义。根据分析的目的不同,情感分析可分为“粗粒度情感分析”和“细粒度情感分析”。粗粒度情感分析主要是针对句子级、篇章级整

收稿日期:2020-07-09 返修日期:2020-08-14 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:山西省科技厅基础研究计划项目(201801D121137)

This work was supported by the Basic Research Project of Shanxi Province(201801D121137).

通信作者:段利国(463035793@qq.com)

体来进行情感分析,并不考虑评价对象或属性的情感,而细粒度情感分析则是针对文本方面级别进行情感分析。现阶段的用户评论是非结构化数据,其表达自由,往往包含多种情感。传统的粗粒度情感分析方法并不能很好地完成细粒度情感分析任务。例如,在“这家店装修一般,菜品的味道还挺好的,就是上菜速度太慢了”这段评价文本中,如果进行细粒度情感分析,对于“装修情况”方面的情感极性是中性的,对于“口感”方面的情感极性是正向的,对于“服务态度”方面的情感是负向的,而对于其他方面则未提及;如果进行粗粒度分析,则不能真正将用户的情感倾向性准确地表达出来。通过对用户评论进行更细粒度级别的情感分析,可以进一步提取出用户的情感,从而更好地表达用户的观点,提取出的数据对于用户和商家都有重要意义。因此,本文主要是在细粒度级别对用户评论进行情感分析,以获得用户评论方面级别的情感倾向。传统的机器学习算法已无法满足大规模用户评论数据的处理需要。目前深度学习被广泛应用于情感分析方面,效果有所提升,但是仍存在一些不足。当前大部分深度学习模型都引入了注意力机制,也有门控单元的引入,但是很少有人将两者结合起来进行探索研究。门控单元结构比较简单,与注意力机制结合扩充了与方面词以及上下文相关的特征信息,结构相对来说不会很复杂,同时也弥补了门控机制参数获得的特征信息可能较少的不足。加入文本情感知识对于情感分析也至关重要,有效利用文本特征信息还需要我们进行进一步的探索。

本文实验选用 AIchallenger2018 中提供的一个面向餐饮领域的细粒度用户评论情感分析数据集。本文在该数据集上对用户评论细粒度情感分析进行了一些探索性研究,主要贡献如下:

(1)构建了 BiGRU-Attention 模型,引入卷积 Tanh-Relu 门控机制,并与基于方面词的 Attention 机制相拼接,进一步抓取与方面词相关的情感信息,以保留更多特征信息,弱化无关信息。

(2)在 HOWNET 评价情感词典的基础上,通过 so-pmi 算法加入用户评论评价情感词,扩充用户评论情感词典。在给定的中文细粒度方面词的基础上,通过相关算法构建方面词词典。

(3)在嵌入层将字特征、词特征、情感极性特征、词性特征以及否定词特征相融合,并将结果作为编码层的输入。

## 2 相关研究

方面级别的情感分析是一种更细粒度的情感分类任务<sup>[3]</sup>,根据现有的细粒度情感分析模型,方面级别的情感分析任务可以分为两个子任务:方面-类别情感分析(Asspect-Category Sentiment Analysis, ACSA)和方面-术语情感分析(Asspect-Term Sentiment Analysis, ATSA)。ACSA 任务中的 aspect-category 指文本中实体的类别,是抽象出来的方面词,即 ACSA 任务是预测一段文本中给定方面词的情感极性,而 ATSA 中 Aspect-Term 是文本中直接存在的名词或实体,因此 ATSA 指分析与文本中出现的目标实体有关的情感极性。在过去的细粒度情感分析研究中,基于情感词典和基于机器学习的情感分析方法在情感分析任务中都取得了不错的效

果,但是这些方法的优劣都是基于对良好特征的选择。而近几年兴起的深度学习方法<sup>[4-5]</sup>很好地弥补了上述方法的不足,在图像和语音识别中取得了巨大的成就,并且逐渐被应用于自然语言处理领域。

相比一般的前馈神经网络而言,循环神经网络(Recurrent Neural Network, RNN)引入了记忆单元,使得网络具有一定的记忆性,可以很好地处理序列信息,但是在训练过程中会出现梯度爆炸或梯度弥散的问题。长短时记忆网络<sup>[6]</sup>(Long Short-Term Memory, LSTM)和门控循环单元<sup>[7]</sup>(Gated Recurrent Unit, GRU)在传统 RNN 的基础上引入门控机制,可以灵活捕获方面词及其上下文之间的语义关系,较好地弥补了 RNN 的不足<sup>[8]</sup>。很多研究者通过使用基于注意力<sup>[9]</sup>的神经网络来进行特定方面的情感分析,并取得了较好的效果。然而,这些工作只关注情感词的相关信息,忽略了方面词的相关信息。为解决上述问题, Tang 等<sup>[10]</sup>在 LSTM 的基础上,根据文本中方面目标词的上文信息和下文信息分别进行建模,提出了 TD-LSTM(Target-Dependent LSTM),其相比传统的 LSTM 分类效果更佳。TD-LSTM 虽然考虑了目标词信息的作用,但是,没有充分利用目标词与上下文词之间的关联信息,因此本文在 TD-LSTM 的基础上在输入端将目标词信息与上下文词汇整合起来,提出了 TC-LSTM(Target-Connection LSTM),以进行情感分类,并取得了比 TD-LSTM 更好的效果。但是,这种方法只考虑与目标词相关的信息,并没有真正将目标词信息应用到方面级别的情感分类上。为此, Wang 等<sup>[11]</sup>在 TD-LSTM 的基础上针对 ACSA 任务提出了 ATAELSTM(Attention-based LSTM)。该方法将 LSTM 与注意力机制相结合,通过注意力机制计算权重,对于不同的方面,能够通过关注句子的不同部分来进行方面级别的情感分析,取得了不错的效果。为了充分利用方面词及其上下文相关信息, Huang 等<sup>[12]</sup>提出了 Attention-over-Attention 神经网络,通过方面词注意力机制与上下文注意力机制联合的方法进行情感分析。Cheng 等<sup>[13]</sup>提出了 HEAT(Hierarchical Attention)网络,通过一个分层注意力模块同时学习方面词信息以及与方面词相关的情感信息,并取得了不错的效果。Wei 等<sup>[14]</sup>提出 GCAE(Gated Convolutional Networks)门控机制单元,通过卷积获得文本特征表示,然后通过门控机制获得与方面词相关的特征信息,较好地利用了方面词信息。Shuang 等<sup>[15]</sup>在 TD-LSTM 的基础上提出了 AELA-LSTMs(Attention-Enabled and Location-Aware DLSTMs),充分获得方面目标词之前和之后的上下文语义信息,通过为目标词上下文分配不同的位置权重来进行情感分析,并取得了不错的效果。上述方法中,基于方面的注意力机制以及门控机制在情感分析领域得到了很好的应用。

上述方法并没有将门控机制与注意力机制相结合来进行探索,也并未有效利用情感知识,因此,为了保留更多与方面词相关的情感信息,将门控机制与注意力机制相结合。基于以上研究,针对细粒度情感分析任务,本文将卷积门控机制与 BiGRU-Attention 相结合,以保留更多情感信息,弱化无关信息,并在嵌入层进一步融合了多种情感特征信息,如词性信息、情感极性信息、否定词信息等。

### 3 模型构建

本文主要进行方面-类别情感分析任务,可形式化描述为句子、方面词和情感极性三元组  $\langle S, A, C \rangle$ ,  $S = \{w_1, w_2, \dots, w_n\}$  表示由  $n$  个词  $w_i$  组成的句子,  $A = \{a_1, a_2, \dots, a_t\}$  表示事先定义好的 aspect-category 共有  $t$  个, 在本文中  $t = 20$ , 具体如表 2 所列.  $C = \{-2, -1, 0, 1\}$  表示情感极性,  $-2$  表示未提及 (not mention),  $-1$  表示负向 (negative),  $0$  表示中性 (neutral),  $1$  表示正向 (positive). 本文的主要任务是根据给定的评论文本以及细粒度方面词预测出评论文本特定细粒度方面的情感倾向. 本文模型的整体架构如图 1 所示. 本文模型共包含以下 4 层.

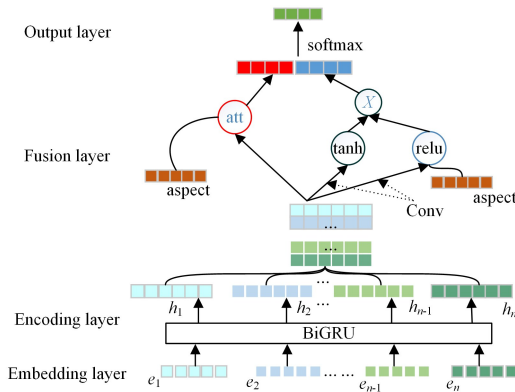


图 1 网络模型结构

Fig. 1 Structure of network model

(1)词嵌入层(word embedding layer):使用 word2vec 将评论文本中的每个词以及方面词映射到向量空间,形成初始化词向量,并在此处引入字向量、情感极性特征、否定词特征以及词性特征等情感特征信息.

(2)编码层(encoding layer):将上一层的输出输入 Bi-

GRU,利用词语的上下文重新定义词嵌入。

(3)融合层(fusion layer):通过注意力机制(基于方面词的注意力机制)与卷积门控机制进行拼接融合来进一步捕获方面词相关信息与情感相关信息。

(4)输出层(output layer):将上一层的信息通过全连接层得到情感评分,最后经过 softmax 得到情感分类。

#### 3.1 词嵌入层

首先对原始数据进行分词、过滤,本文使用的是哈尔滨工业大学发布的预处理工具 LTP<sup>1)</sup>进行分词,停用词表使用的是哈尔滨工业大学提供的公开的哈尔滨工业大学停用词表数据.通过停用词表筛选并过滤掉语料中与情感信息或者方面词信息相关度不强的词,以便于进行情感分析.

本层的输入主要由两部分构成:文本输入联合情感特征信息以及方面词输入.

##### 3.1.1 字词联合

一般的自然语言处理任务中常使用的是单纯的词输入,中文文本中,每个字都有自己的含义,因此本文尝试采用字向量与词向量联合的方法,如图 2 所示.某个词语由  $c_1, c_2, c_3$  3 个字组成,首先采用 word2vec 将词向量和字向量全部映射为  $d_w$  维向量(本文实验设置为 100),即  $W_i \in \mathbb{R}^{d_w \times v}$ ,  $v$  表示语料中互不相同的词数目,将文本序列  $W = [w_1, w_2, \dots, w_m]$  中的各个单词和各个字都映射到对应的  $d_w$  维向量,得到  $E^w \in \mathbb{R}^{d_w}$ ,将字向量与词向量进行匹配,通过语言模型 BiGRU 获得各个词语的字向量,然后进行最大池化操作,最后通过拼接的方式,将字向量与词向量联合起来,如式(1)~式(3)所示:

$$c_i = BiGRU(s_i) \tag{1}$$

$$e_i^c = [\max(c_1), \dots, \max(c_{2d_h})] \tag{2}$$

$$e_i = [e_i^w; e_i^c] \tag{3}$$

其中,  $C \in \mathbb{R}^{n \times 2d_h}$ ,  $e_i^c \in \mathbb{R}^{2d_h}$ ,  $e_i \in \mathbb{R}^d$ ,  $n$  为一个词中包含  $n$  个字,  $d_h$  为 BiGRU 的隐藏层大小(此处设置为 100).

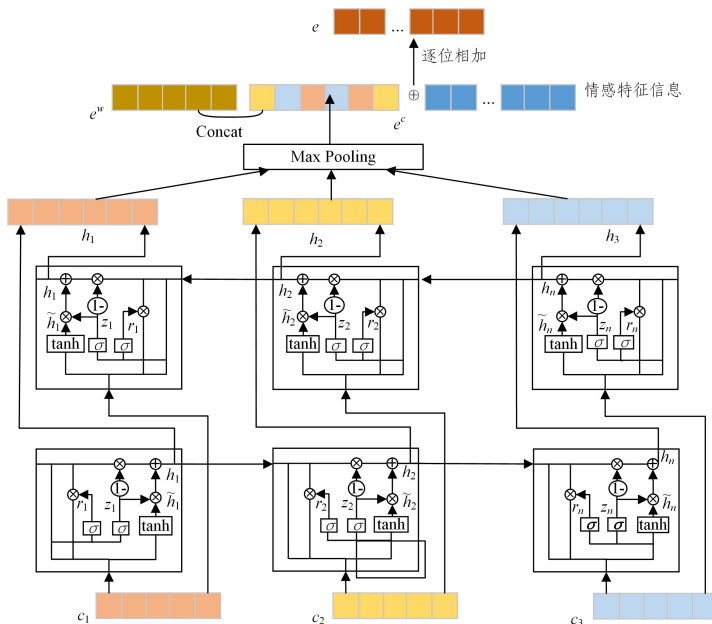


图 2 字向量与词向量联合并添加情感特征信息

Fig. 2 Embedding layer concatenated word vectors with character vectors and combined sentiment feature information

<sup>1)</sup> <http://ltp.ai/index.html>

### 3.1.2 情感特征嵌入

本文共使用 3 种情感特征,分别是词性特征、否定词特征以及情感极性特征。词性特征使用哈尔滨工业大学发布的预处理工具 LTP;否定词词表主要通过收集现有的互联网情感资源得到,共包含 243 个否定词;情感极性特征主要由知网(HOWNET)发布的“情感分析用词语集”中的中文正面评价词典和中文负面评价词典得到,其中正面评价词语有 3 730 个,负面评价词语有 3 116 个。

本文通过对数据集进行分析,来进一步扩充情感评价词典,引入与数据集相关的情感词,如“比得上”“无敌”“给力”等,并使用知网评价情感词作为种子情感词,采用 so-pmi<sup>[16]</sup>算法作为情感词扩展的算法,如式(4)、式(5)所示:

$$PMI(word1, word2) = \log_2 \left( \frac{P(word1 \& word2)}{P(word1)P(word2)} \right) \quad (4)$$

$$SO-PMI(word1) = \frac{\sum_{pword \in pwords} PMI(word1, pword) - \sum_{mword \in mwords} PMI(word1, mword)}{\sum_{pword \in pwords} PMI(word1, pword) + \sum_{mword \in mwords} PMI(word1, mword)} \quad (5)$$

其中,  $pwords$  是知网评价情感词库筛选后的情感极性为正向的情感词库,  $mwords$  是筛选后情感极性为负向的情感词库。

用户评论情感词库扩展的主要流程如图 3 所示。

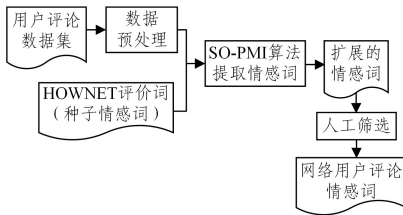


图 3 获得用户评论情感词的流程图

Fig. 3 Flow chart of getting user comment sentiment words

经 SO-PMI 算法扩充情感词并进行人工筛选后,其中正面评价词语扩充为 5 891 个,负面评价词语扩充为 2 963 个,具体如表 1 所列。

表 1 用户评论情感词典

Table 1 User comment sentiment dictionary

	数量/个	举例
否定词表	243	不用、不是、没有
HOWNET 评价情感词	6 846	好心、逼真、正直
扩充后的情感词	8 854	无奈、根本、爆棚

设置 7 种情感特征信息,分别为名词、动词、副词、形容词、否定词、正向情感标签、负向情感标签。根据文本序列  $W = [w_1, w_2, \dots, w_m]$ ,对于每个词  $w_i$ ,若包含情感特征信息,则为其初始化一个  $d$  维向量,若不包含则初始化为  $d$  维全零向量,7 种情感特征信息依次类推。特征融合一般有两种方式,按点逐位相加和向量拼接,本文在此处选择按点逐位相加的方式,将 7 种情感特征向量与字词联合向量  $e_i$  按点逐位相加,最终得到  $E = [e_1, e_2, \dots, e_m]$  作为编码层的输入,如图 2 所示。其中  $e_i$  表示评论文本中第  $i$  ( $i \in [1, m]$ ) 个词语,  $d = 2d_h + d_w$ 。

### 3.1.3 方面词嵌入

对于方面词输入,前期工作主要是将语料中提供的 20 个

细粒度方面进行分词、清洗,人工筛选出与各个方面相关的词。紧接着本文在前期工作的基础上通过 SO-PMI 以及 Tf-idf 方法分别提取出方面目标词并筛选方面词。在此,先进行二分类,分别将各个方面词类别下的情感标签为 -2 的词分出来,再通过 SO-PMI 和 Tf-idf 方法寻找各个细粒度类别下特征重要性较高的词,根据特征重要性排序为每个细粒度要素分别抽取前 50 个相对重要的方面目标词, Tf-idf 方法的主要计算公式如式(6)~式(8)所示:

$$TF_w = \frac{\text{在语料中词条 } w \text{ 出现的次数}}{\text{该语料中所有词条的数目}} \quad (6)$$

$$IDF = \log \left( \frac{\text{语料中的文本总数}}{\text{包含词条 } w \text{ 的文本数} + 1} \right) \quad (7)$$

$$TF-IDF = TF * IDF \quad (8)$$

最后计算前期工作中抽取出的与各个方面相关的词和提取出的方面目标词的相似度值,通过基于 word2vec 词向量相似度的计算,来计算方面词与方面目标词的相似度,保留高质量的方面目标词,选出最终方面词信息。本文采用 python 中第三方模块 gensim<sup>1)</sup> 中的 word2vec 词向量的余弦相似度的计算函数,计算每个细粒度类别提取出的方面词与方面目标词的相似度并对其进行排序,然后进行筛选。主要计算公式如下:

$$\text{sim}_{\cos}(u, v) = \frac{u \cdot v}{\|u\|_2 \|v\|_2} \quad (9)$$

其中,  $\|u\|_2$  和  $\|v\|_2$  表示向量  $v$  和  $u$  的各个元素的平方和再开平方。最终,为细粒度类别筛选了相关方面词共计 140 个。获得部分方面目标词的具体例子如表 2 所列。筛选出各个细粒度要素的方面目标词之后,与上述文本输入类似,每个词通过字特征与词特征联合,得到每个方面目标词的嵌入矩阵,通过各个维度求和平均获得方面词的输入矩阵。最终得到  $a_i \in \mathbb{R}^d$ ,  $a_i$  表示第  $i$  个细粒度方面词 ( $i \in [1, 20]$ )。  $A = [a_1, a_2, \dots, a_{20}]$  作为融合层的输入。

表 2 获得方面目标词

Table 2 Obtain aspect target words

细粒度要素	预处理过程	方面目标词举例
交通是否便利	交通	交通、地铁站、地铁、公交车
距离商圈远近	商圈	位置、百货、商圈、附近
是否容易寻找	寻找	位置、寻找、找到、地理位置
排队等候时间	排队/等候	排队、等、小时、分钟
服务人员态度	态度	态度、服务员、服务、老板
是否容易停车	停车	停车、车位、停车场、开车
点菜/上菜速度	上菜/速度	上菜、速度、分钟、等
价格水平	价格	价格、价位、价钱
性价比	性价比	性价比、物美价廉、价廉物美
折扣力度	折扣	折扣、打折、团购、券、优惠
装修情况	装修	装修、装饰、布置、店面、装潢
就餐空间	空间	空间、面积、装修、环境
嘈杂情况	嘈杂	嘈杂、吵闹、安静、氛围、音乐
卫生情况	卫生	卫生、环境、干净、脏
分量	分量	分量、一份、每份、足
口感	口感	口感、味道、入味、香味
外观	外观	外观、造型、卖相、颜色
推荐程度	推荐	推荐、强烈推荐、值得一试
本次消费感受	感受	消费、感受、满意
再次消费意思	再次	再次、还会、下次

### 3.2 编码层

在获得文本词嵌入的基础上,需要进一步获取基于上下

<sup>1)</sup> <https://radimrehurek.com/gensim/>

文本文语义的编码表示, LSTM 和 GRU 是两个非常优秀的 RNN 变体, 通过使用一些门来解决梯度消失问题。很多之前的工作也显示出与 LSTM 和 GRU 在许多任务中的执行情况相似, 也有很多实验<sup>[17-18]</sup>表明, 在参数相同的情况下, GRU 的实验效果会比 LSTM 稍好一些, 本文也进行了相应的实验对比。

本文使用 GRU 来获取文本的上下文语义编码。GRU 的模型公式如式(10)~式(13)所示:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (10)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (11)$$

$$\tilde{h}_t = \tanh(W_h x_t + r_t \circ U_h h_{t-1} + b_h) \quad (12)$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \tilde{h}_t \quad (13)$$

其中,  $\circ$  表示点积,  $\sigma$  是 sigmoid 函数,  $z_t$  是更新门,  $r_t$  是重置门,  $h_t$  是输出的隐含状态。本文使用的是 BiGRU, 即双向 GRU, 可以同时两个方向的序列进行操作, 每个节点会生成正向和反向两个不同方向门控单元输出向量, 将这两个方向的向量进行拼接, 获得当前节点的向量。  $h_t = [\vec{h}_t, \overleftarrow{h}_t]$ , 其中  $\vec{h}_t$  为前向输出结果,  $\overleftarrow{h}_t$  为后向输出结果。  $\vec{H} = [\vec{h}_1, \vec{h}_2, \dots, \vec{h}_T] \in \mathbb{R}^{2d_h \times T}$ ,  $d_h$  为隐含层维度, 对于经过 BiGRU 获得上下文语义的编码的操作, 可以表示为  $\vec{H} = \text{BiGRU}(E)$ 。

### 3.3 融合层

#### 3.3.1 注意力层

在编码层获得文本的包含上下文语义表示的词编码以及词嵌入层的方面词嵌入后, 注意力机制计算基于方面词嵌入以及上下文特征的每个词的权重值, 如式(14)、式(15)所示:

$$g_t = u^T \tanh(W^{(a)} [a; \vec{h}_t] + b^{(a)}) \quad (14)$$

$$\alpha_t = \frac{\exp(g_t)}{\sum_{j=1}^m \exp(g_j)} \quad (15)$$

其中,  $W^{(a)} \in \mathbb{R}^{d_h \times (2d_h + d)}$  是权值矩阵,  $b^{(a)} \in \mathbb{R}^{d_h}$  是偏置向量,  $u \in \mathbb{R}^{d_h}$  是权值向量,  $\alpha_t$  是词语  $w_t$  包含目标方面词信息的权值。依据权值, 对序列中所有节点的隐含向量进行加权, 具体如式(16)所示, 输出维度与隐含层大小相关,  $v \in \mathbb{R}^{2d_h}$ 。

$$v = \sum_{t=1}^m \alpha_t \vec{h}_t \quad (16)$$

#### 3.3.2 门控机制

门控机制与注意力层一样, 输入是编码层提供的文本序列  $\vec{H}$  以及词嵌入层提供的方面目标词向量  $A$ 。Dauphin 等<sup>[19]</sup>提出的门控单元(Gated Tanh Unit, GTU)和门控线性单元 GLU(Gated Linear Unit, GLU)显示了门控机制的有效性。门控机制主要是在文本词嵌入经过编码层后通过卷积, 卷积核大小设为  $k$ , 分别取 1, 2, 3, 步长设为 1, 卷积核数量设为  $l=128$ , 然后将从词嵌入层获得的方面词与上下文的隐含向量通过  $\text{relu}$  门获得方面词门控特征  $a_t$ , 同时上下文的隐含向量通过  $\tanh$  函数获得情感门控特征  $S_t$ , 最后将  $S_t, a_t$  元素一一相乘获得最终门控特征  $c_t$ 。

与注意力机制的作用类似, 本文致力于聚焦与方面词相

关的情感词, 如式(17)~式(20)所示:

$$a_t = \text{relu}(\overrightarrow{h_{t,t+k-1}} * W_a + aU_a + b_a) \quad (17)$$

$$S_t = \tanh(\overrightarrow{h_{t,t+k-1}} * W_s + b_s) \quad (18)$$

$$c_t = S_t \times a_t \quad (19)$$

$$g_t = [\max(c_1), \dots, \max(c_l)] \quad (20)$$

其中,  $W_a, W_s \in \mathbb{R}^{2d_h \times k \times l}$ ,  $U_a \in \mathbb{R}^{d^* \times l}$ ,  $b_a, b_s \in \mathbb{R}^l$ ,  $*$  表示卷积操作, 最后通过池化函数获得最终门控机制表示, 分别将不同大小的卷积核得到的  $g$  进行拼接,  $g \in \mathbb{R}^{3l}$ 。

最后将注意力层与门控机制进行融合, 即拼接得到  $f = [v; g]$ ,  $f \in \mathbb{R}^{2d_h + 3l}$ 。

### 3.4 输出层

本层主要是在获得上一层的注意力特征与门控机制特征后, 通过全连接层 softmax 函数来预测基于方面的文本情感极性, 如式(21)、式(22)所示:

$$y = \text{softmax}(W_y f + b_y) \quad (21)$$

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_{j=1}^{d_s} \exp(x_j)} \quad (22)$$

其中,  $W_y \in \mathbb{R}^{d_s \times (2d_h + 3l)}$ ,  $b_y \in \mathbb{R}^{d_s}$ ,  $y \in \mathbb{R}^{d_s}$ ,  $d_s$  是语料中情感极性的分类, 本实验用到的语料  $d_s = 4$ 。本文模型是通过减小预测值与真实值的交叉熵损失函数值来进行训练的, 具体如式(23)所示:

$$\text{loss} = -\frac{1}{P} \sum_i^P \sum_j^{d_s} y_i^j \log \hat{y}_j^i \quad (23)$$

其中,  $P$  是样本数量,  $\hat{y}$  是预测值,  $y$  是真实的值。

## 4 实验结果及分析

### 4.1 数据集与评价指标

本文选用 AIchallenger2018<sup>1)</sup> 提供的面向餐饮领域的细粒度用户评论情感分析数据集进行分析研究, 数据集共包含 20 个细粒度要素, 分别为: 交通是否便利、距离商圈远近、是否容易寻找、排队等候时间、服务人员态度、是否容易停车、点菜/上菜速度、价格水平、性价比、折扣力度、装修情况、就餐空间、嘈杂情况、卫生情况、分量、口感、外观、推荐程度、本次消费感受、下次消费意向。每个细粒度要素的情感倾向有 4 种状态: 未提及、负向、中性、正向, 使用  $[-2, -1, 0, 1]$  这 4 个值对情感倾向进行描述。数据集全部来自大众点评的真实公开用户评论, 其中训练集有 10.5 万条, 验证集有 1.5 万条。由于测试集没有答案, 因此本文将 AIchallenger2018 提供的数据集整合起来, 按照近似 8:1:1 的比例将数据集分为训练集、验证集和测试集, 数据集中每条数据都是由评论文本、细粒度方面词以及情感极性组成。在整个实验中, 除使用扩充的情感知识库外, 没有使用任何外部数据。

本文后续和其他系统进行了实验对比, 将 Macro\_Precision, Macro\_Recall 以及 Macro\_F1\_score 值作为本次实验的评价指标。对于 Precision 和 Recall 值, 先计算各个细粒度要素中每个情感类别的 precision 与 recall 值, 然后求均值, 最后得到各个细粒度要素的 Macro\_Precision, Macro\_Recall 和 Macro\_F1\_score, 接着求 20 个细粒度要素的均值, 获得本次

<sup>1)</sup> <https://challenger.ai/dataset/fsaouord2018>

实验最终的评价指标。具体公式如下:

$$recall_i^j = \frac{TP_i^j}{TP_i^j + FN_i^j} \quad (24)$$

$$precision_i^j = \frac{TP_i^j}{TP_i^j + FP_i^j} \quad (25)$$

$$macro\_recall_i = \frac{\sum_{j=1}^4 recall_i^j}{4} \quad (26)$$

$$macro\_precision_i = \frac{\sum_{j=1}^4 precision_i^j}{4} \quad (27)$$

$$macro\_F1_i = \frac{2macro\_precision_i \times macro\_recall_i}{macro\_precision_i + macro\_recall_i} \quad (28)$$

$$Macro\_recall = \frac{\sum_{i=1}^{20} macro\_recall_i}{20} \quad (29)$$

$$Macro\_precision = \frac{\sum_{i=1}^{20} macro\_precision_i}{20} \quad (30)$$

$$Macro\_F1\_score = \frac{\sum_{i=1}^{20} macro\_F1_i}{20} \quad (31)$$

其中,  $TP_i^j$  是在第  $i$  个细粒度要素的第  $j$  个情感类别下预测为正且实际也为正的样本数,  $FP_i^j$  是在第  $i$  个细粒度要素的第  $j$  个情感类别下预测为正且实际为负的样本数,  $FN_i^j$  是在第  $i$  个细粒度要素的第  $j$  个情感类别下预测为负且实际为正的样本数,  $TN_i^j$  是在第  $i$  个细粒度要素的第  $j$  个情感类别下预测为负且实际为负的样本数。

## 4.2 模型参数设置

本文词向量采用 Word2Vec 开源工具中的 Skip-gram 模型进行预训练,维度设置为 100,字特征中 Bi-GRU 的隐含层维度设置为 100,在神经网络模型训练的过程中,词向量以及融合的特征向量都会不断学习更新。编码层 Bi-GRU 隐含层的维度设置为  $d_h = 64$ 。在训练过程中,模型采用 Adamax<sup>[18]</sup> 作为优化器,学习率设置为 0.001。误差采用批处理, batch-size 设置为 32。为避免过拟合, dropout 设置为 0.5, epoch 设置为 100, earlystop = 2, 如果 Macro\_F1\_score 连续 3 次没有上升,则停止训练,记录最优的 epoch,保存最优模型。

## 4.3 实验结果及分析

首先将本文模型与 SVM, CNN, ATAE, HEAT 以及 GCAE 等常用的且较前沿的模型方法进行对比,本文还做了几个基准实验,作为模型的对比实验的参考,最后进行自对比消融实验。

**SVM:** 机器学习常用的分类器,本文中正则化常数  $C$  设置为 1.0,核函数选择 RBF 核函数,采用词向量进行输入。

**CNN:** 参数设置与本文的卷积参数设置一致,采用词向量进行输入。

**ATAE:** 融合方面词,使用 LSTM 进行编码,引入注意力机制聚焦与方面词相关的情感信息,采用 LSTM 与 max-pooling 进行字向量与词向量联合,作为编码输入,其中,字词向量联合中 LSTM 的维度设置为 64, LSTM 隐含层维度设置为 128。

**HEAT:** 融合注意力模块学习方面词信息以及与方面词相关的情感信息,使用 BiGRU 进行编码,融合注意力机制进

一步聚焦情感信息,采用词向量进行输入,方面词输入主要是数据集给定的方面词,其余维度参数设置与本文一致。

**GCAE:** 使用门控机制进行方面级别的情感分析,先进行卷积,通过门控单元,经过池化层,线性层获得最终情感极性,此处仅采用词向量嵌入,方面词输入使用的是 3.2 节中提到的方面词输入方式,其余参数设置与本文模型一致。

**BiGRU-ATT+w:** 作为基准对比实验模型,使用 BiGRU 进行编码,融合注意力模块。这里的注意力机制是普通的注意力机制,未融合方面词信息,仅仅聚焦于情感信息,最后通过全连接层计算获得情感极性。其中,模型的输入采用词向量输入,其余参数设置与本文模型的参数设置一致。

**BiGRU-ATT+c:** 与 BiGRU-ATT+w 模型架构一致,但此处采取字向量作为模型输入。

**BiGRU-ATT(asp)+w&c:** 在此处注意力机制聚焦于与方面词相关的情感词,融合了方面词信息,采取字向量与词向量联合作为模型输入,参数设置与本文模型的参数设置一致。

**BiLSTM-ATT(asp)+w&c:** 为增加对比性,加入了 BiLSTM 来做对比实验,参数设置与本文模型的参数设置一致。

**OUR MODEL:** 本文模型的词嵌入层采用上下文字向量与词向量联合的方式,融合情感特征信息特征作为编码层的输入,编码层使用 BiGRU 对输入的信息进行上下文编码,将预先获得的方面词特征以及编码层获得的上下文编码作为融合层的输入,将基于方面的注意力机制与基于方面门控机制融合,进一步聚焦方面信息与情感信息,最后通过全连接层和 softmax 函数来预测情感极性。

表 3 列出了实验结果。由实验结果可知,本文模型明显优于其他模型。

表 3 实验结果

Table 3 Experimental results

Method	Macro-Precision	Macro-Recall	Macro-F1 score
SVM	0.1709	0.25	0.2010
CNN	0.6564	0.5898	0.6213
ATAE	0.6647	0.6617	0.6632
HEAT	0.6673	0.6749	0.6728
GCAE	0.6717	0.6864	0.6790
BiGRU-ATT+w	0.6175	0.6762	0.6438
BiGRU-ATT+c	0.6381	0.6806	0.6561
BiLSTM-ATT(asp)+w&c	0.6603	0.7142	0.6822
BiGRU-ATT(asp)+w&c	0.6685	0.7136	0.6872
OUR MODEL	0.7217	0.7239	0.7218

BiGRU 在本次任务中略优于 BiLSTM,在未融入门控机制和情感特征时使用 BiGRU 比使用 BiLSTM 的效果略好一些。对比 HEAT 与 BiGRU-ATT+w 可以看出,基于方面词的注意力机制明显优于常规注意力机制,表明加入方面词信息对细粒度情感分析有很大的作用;对比 ATAE 和 BiLSTM-ATT(asp)+w&c 可以看出, BiLSTM 明显优于单向 LSTM;对比 BiGRU-ATT+w 和 BiGRU-ATT+c 可以看出,字特征相对来说效果略优于词特征。一直以来, NLP 任务中都是使用词进行建模的,但词并不是最小的语言单位,它是由字组成的,每个字也有各自的意义,因此采用字输入可以弥补词输入 OOV(Out of Vocabulary)的不足。

将本文模型的 Macro-F1 score 值与对比算法的 Macro-F1 score 值进行比较,证明了本文实验的有效性,通过引入情感特征信息以及门控机制与注意力机制拼接,可以更进一步地抓取方面信息与上下文情感信息。

为了进一步对比本文方法,做消融实验以与对比测试集的 Macro\_Precision, Macro\_Recall 以及 Macro\_F1\_score 值进行对比,结果如表 4 所列。可以看到,分别去掉本文模型中的情感特征信息、门控机制以及注意力机制, Macro\_Precision, Macro\_Recall 以及 Macro\_F1\_score 值都会下降,这充分验证了本文模型的有效性。

表 4 消融实验 Macro\_F1\_score 值的对比结果

Table 4 Comparison of Macro\_F1\_score values of ablation experiments

Method	Macro-Precision	Macro-Recall	Macro-F1_score
Our model	0.7217	0.7239	0.7218
Our model without feature in formation	0.6971	0.7177	0.7051
Our model without gated	0.6895	0.7180	0.7035
Our model without attention	0.6853	0.7096	0.6968

为了进一步分析实验结果,图 4 给出了本实验模型计算的各个细粒度要素下的 Macro\_Precision, Macro\_Recall 以及 Macro\_F1\_score 值。由图 4 可以看出,在距离商圈远近、菜品外观以及本次消费感受这 3 个细粒度要素上的分类效果较低,进一步对比了在不同的实验模型下 20 个细粒度要素的分类效果,这里主要是对 Macro\_F1\_score 值进行对比,如图 5 所示。

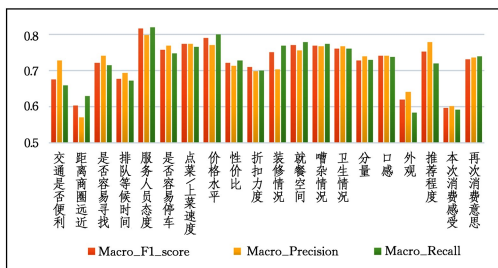


图 4 本实验模型在 20 个细粒度要素下的 Macro\_Precision, Macro\_Recall, Macro\_F1\_score 值

Fig. 4 Macro\_Precision, Macro\_Recall, Macro\_F1\_score values of 20 fine-grained elements in this model

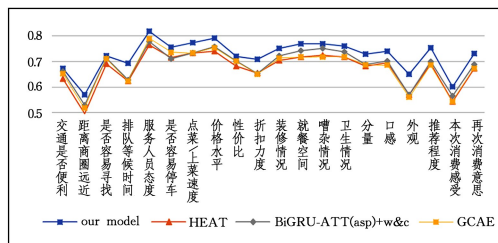


图 5 不同实验模型 Macro\_F1\_score 值的对比

Fig. 5 Comparison of Macro\_F1\_score values of different models

分析图 4、图 5 可以得出,在用户评论中,距离商圈远近、菜品外观以及本次消费感受这 3 个细粒度要素方面的分类效果不及其他方面,而服务人员态度以及价格水平方面的分类效果较好。

统计数据集中各个细粒度要素情感值为 -2 的文本数,即情感倾向为未提及的文本数量在总文本数量中的占比并进行排序,如图 6 所示。用户评论中未提及本次消费感受、菜品味道、服务人员态度、价格水平的数量较少,而未提及是否容易停车、排队等候时间以及点菜/上菜速度的数量较多,说明用户在评论时更关注菜品味道、本次消费感受、服务人员态度和价格水平,而对于是否容易停车、排队等候时间以及点菜/上菜速度的关注较少。

图 4 和图 5 所示的结果显示在本次消费感受方面的分类效果较差,通过对数据集本次消费感受方面具有情感倾向的文本进行分析可以发现,该方面的情感标注上的 4 种情感界限不是很清晰,更偏向于整体的情感倾向。当出现多种情感共现时,情感倾向判断不是很明确,可能导致分类效果不佳。而对于距离商圈远近以及菜品外观,通过对数据集进行分析可以发现,这两个方面的各个用户评论与方面相关的情感特征相对来说不太明显,因此可能导致分类效果欠佳。

对于商家来说,在餐厅的运营过程中需要更多关注菜品味道、价格水平以及服务人员态度这 3 个方面的提升。在店面位置的选取上,需要多关注店铺能否让消费者容易找到,在店内环境方面,店铺的装修相对来说也比较重要。因此,假如商家在这几个方面都能有所提升,则对于提升餐厅的效益会有很大的帮助。

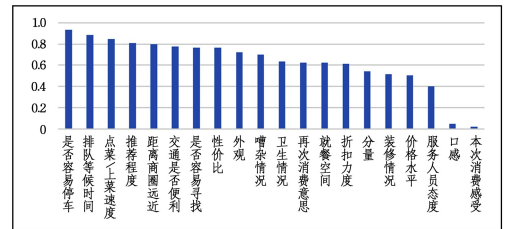


图 6 数据集中各细粒度情感倾向为未提及的数据占比排序

Fig. 6 Data proportion ranking of fine-grained sentiment polarity of hot mention in dataset

**结束语** 针对方面级别细粒度情感分析任务,本文提出了一个基于注意力机制与门控机制相融合的神经网络模型来进行方面级别的细粒度情感分析。首先,扩充用户评论情感词、词性信息以及否定词信息作为模型的情感特征信息;然后,字向量与词向量融合作为情感特征信息,经过 BiGRU 进一步获得包含上下文语义信息的编码表示,融合方面词注意力机制以及门控机制,获取文本中与方面词紧密相关的特征信息;最后通过输出层获得文本情感极性。

虽然本文在面向餐饮领域的细粒度用户评论情感分析数据集上从整体上取得了较好的结果,但是对于个别细粒度方面的 Macro\_F1\_score 值有待进一步的提升。考虑到融合情感特征信息的方式不同也会对结果产生影响,接下来的工作将选择性地加入多种特征信息,以寻找更优的融合方法进行进一步的实验,验证融合多种特征信息对基于方面的细粒度情感分析的效果的影响。

## 参考文献

- Sentiment Words Detection Based on Attention Mechanism [J]. *Computer Science*, 2020, 47(1): 186-192.
- [2] ZHAO Y Y, QIN B, LIU T. Sentiment analysis [J]. *Journal of Software*, 2010, 21(8): 1834-1848.
- [3] TANG X B, LIU G C. A Review of Studies on Fine-grained Sentiment Analysis [J]. *Library and Information Service*, 2017(5): 132-140.
- [4] DOHAIHA 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(1): 272-299.
- [5] LI Y H, XIE M, YI Y. Fine-grained Sentiment Analysis for Social Network Platform based on Deep-learning model [J]. *Application Research of Computers*, 2017, 34(3): 743-747.
- [6] MALHOTRA P, VIG L, SHROFF G, et al. Long Short Term Memory Networks for Anomaly Detection in Time Series [C] // *European Symposium on Artificial Neural Networks*. 2015.
- [7] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling [J]. arXiv: 1412. 3555, 2014.
- [8] ZHANG Z, ROBINSON D, TEPPER J. Detecting hate speech on Twitter using a convolution-GRU based deep neural network [C] // *ESWC 2018*. Springer, Cham, 2018.
- [9] RAFFEL C, ELLIS D P W. Feed-Forward Networks with Attention Can Solve Some Long-Term Memory Problems [J]. arXiv: 1512. 08756, 2015.
- [10] TANG D, QIN B, FENG X, et al. Effective LSTMs for Target-Dependent Sentiment Classification [J]. arXiv: 1512. 01100, 2015.
- [11] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for Aspect-level Sentiment Classification [C] // *Conference on Empirical Methods in Natural Language Processing*. 2016.
- [12] HUANG B, OU Y, CARLEY K M. Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks [J]. arXiv: 1804. 06536, 2018.
- [13] CHENG J, ZHAO S, ZHANG J, et al. Aspect-level Sentiment Classification with HEAT (HiErarchical ATtention) Network [C] // *ACM*. 2017: 97-106.
- [14] WEI X, TAO L. Aspect Based Sentiment Analysis with Gated Convolutional Networks [J]. arXiv: 1805. 07043, 2018.
- [15] SHUANG K, REN X, YANG Q, et al. AELA-DLSTMs: Attention-Enabled and Location-Aware Double LSTMs for aspect-level sentiment classification [J]. *Neurocomputing*, 2019, 334: 25-34.
- [16] TURNEY P D, LITTMAN M L. Measuring praise and criticism: inference of semantic orientation from association [J]. *Acm Transactions on Information Systems*, 2003, 21(4): 315-346.
- [17] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling [J]. arXiv: 1412. 3555, 2014.
- [18] JOZEFOWIC Z, RAFA L, ZAREMB A, et al. An Empirical Exploration of Recurrent Network Architectures [C] // *Proceedings of the 32nd International Conference on Machine Learning*. 2015: 2342-2350.
- [19] DAUPHIN Y N, FAN A, AULI M, et al. Language Modeling with Gated Convolutional Networks [J]. arXiv: 1612. 08083, 2016.



**ZHANG Jin**, born in 1996, postgraduate. Her main research interests include sentiment analysis and natural language processing.



**DUAN Li-guo**, born in 1970, Ph.D, associate professor, is a senior member of China Computer Federation. His main research interests include automatic question answering system, text sentiment analysis, entity relationship extraction and knowledge mapping.