CN 43-1258/TP
ISSN 1007-130X

计算机工程与科学
Computer Engineering & Science

第 44 卷第 6 期 2022 年 6 月
Vol. 44, No. 6, Jun. 2022

文章编号:1007-130X(2022)06-1097-08

基于正文和评论交互注意的微博案件方面识别*

段　玲[1,2],郭军军[1,2],余正涛[1,2],相　艳[1,2]

(1.昆明理工大学信息工程与自动化学院,云南 昆明 650500;
2.昆明理工大学云南省人工智能重点实验室,云南 昆明 650500)

摘　要:微博案件观点所涉方面的自动识别是了解互联网社交媒体新闻舆情的重要手段,但由于微博文本形式和内容均灵活多变,传统的方面识别方法通常只利用单一的正文或评论,使微博语义理解非常有限。针对涉案微博文本的方面识别问题开展研究,提出一种基于正文和评论交互注意的案件方面识别方法,通过融合社交媒体的上下文信息,实现对案件观点所涉方面的识别。首先基于 Transformer 框架对正文和评论分别进行编码;然后基于交互注意力机制,实现正文信息和评论信息的融合,并基于融合后的特征实现对评论文本案件方面的识别;最后基于 12 个案件构建的微博数据集进行实验,实验结果表明,采用交互注意力机制融合微博正文信息和评论信息可以显著提升案件方面识别的准确率,证明了所提方法的有效性。

关键词:社交媒体文本处理;微博案件;方面识别;交互注意;Transformer

中图分类号:TP391.1　　**文献标志码**:A

doi:10.3969/j.issn.1007-130X.2022.06.018

Aspect identification of microblog cases based on the interactive attention of contents and comments

DUAN Ling[1,2], GUO Jun-jun[1,2], YU Zheng-tao[1,2], XIANG Yan[1,2]

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500;
2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: The automatic identification of aspects involved in microblog cases is an important means to understand the public opinion of the Internet social media news. However, the text format and content of the microblog are flexible and changeable, and the traditional aspect identification methods usually use only a single text or comment, which brings great difficulties to the understanding of microblog semantics. This paper studies the identification of aspects in the microblog text involved in the cases, proposes an aspect identification method of microblog cases based on the interactive attention of contents and comments, and realizes the aspect identification of microblog cases by integrating the contextual information of social media. The paper firstly encodes the contents and comments individually based on the Transformer framework, realizes the fusion of the content information and the comment information based on the interactive attention mechanism, and realizes the aspect identification of the comment text based on the fused features. Finally, experiments were conducted based on the microblog dataset containing 12 cases. The experimental results show that using the interactive attention to fuse microblog content information can significantly improve the accuracy of aspect identification, which proves the effectiveness of the method proposed in the paper.

* **收稿日期**:2020-10-27;**修回日期**:2020-12-30
基金项目:**国家重点研发计划**(2018YFC0830105, 2018YFC0830100);**国家自然科学基金**(61972186, 61762056, 61472168, 61866020);**云南省科技厅省级人培项目**(KKSY201703015);**云南省科技厅面上项目**(202001AT070047)
通信作者:**余正涛**(ztyu@hotmail.com)
通信地址:650500 **云南省昆明市昆明理工大学信息工程与自动化学院**
Address: Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, Yunnan, P. R. China

Key words: text processing for social media; microblog case; aspect identification; interactive attention; Transformer

1 引言

近年来网民对于法律案件的关注度增加,特别是特大重大案件的评审工作,已经成为互联网关注的热点,微博媒体对此类案件的报道愈加频繁,并且有很多网民对此发表评论。目前这类新闻的文本较多,面对庞大的数据,通过人工阅读大量评论来把握舆论动向是不现实的,而且民众和司法机关往往关注案件的不同方面。因此,面向微博涉案数据的方面识别研究对迅速掌握互联网态势具有重要的研究意义。

但是,微博数据形式和表述方式通常灵活多变,基于传统自然语言处理技术的方法进行微博案件观点所涉方面的判别较为困难。事实上,微博正文是对案件事实的陈述,包括对案件各方面的描述,微博评论大都是围绕正文内容展开的讨论,因此结合正文的信息能够对涉案文本的理解提供帮助。例如,"丽江反杀案"的微博文本示例如表 1 所示。

Table 1 Microblog case text

表 1 微博案件文本

标题:丽江反杀案
微博正文:云南丽江。醉酒男李某湘持刀砸门,90 后退伍女兵唐雪将其反杀。8 月 28 日,劝架者告诉澎湃新闻,他们拖拽李某湘准备往回走时,此时唐雪开门,被挣扎中的李某湘踹了一脚,唐雪便扑上来,两人扭打到了一起,全程不超过两分钟,如果当时双方各退一步,悲剧就不会发生。唐雪父亲称,男子砸门后欲翻墙进家,女儿行为应属正当防卫。检察院通知律师说以故意伤害防卫过当起诉,我们对这个起诉有意见。希望法律还我们一个公正公平。检察院工作人员告诉澎湃新闻,已经向法院依法提起公诉。
评论 1:谁主动拿刀纠集几人半夜到谁家? 评论 2:唐某查看开门之后李某踢了唐某腹部一脚,然后唐某拿出红色小削刀警告,李某不理睬依然继续殴打唐某。 评论 3:法院在审判,但目前至少检方认为不是正当防卫。

对于刑事案件,民众关注的焦点一般聚焦在:嫌疑人、被害人、案由和其他四个方面,并且通过分析微博案件数据,证实了以上四个方面的可靠性。以"丽江反杀案"为例,评论 1 中提到两个"谁",仅凭此条评论识别所包含的方面,并且明确这两个"谁"分别指代哪个对象是有困难的。根据正文中"李某湘持刀砸门,90 后退伍女兵唐雪将其反杀",可以明确第一个"谁"指代李某湘,第二个"谁"指代唐雪,评论中包含嫌疑人和被害人两个方面。根据正文中"唐雪开门,被挣扎中的李某湘踹了一脚,唐雪便扑上来,两人扭打到了一起",可以知道评论 2 描述的是案发的过程和原因,除了嫌疑人和被害人,还包含案由共三个方面。根据正文中"检察院通知律师说以故意伤害防卫过当起诉",可以知道评论 3 提及的是其他的方面。

微博案件的方面识别就是将网民对司法案件的评论映射到上述四个不同的方面,但是由于评论中包含一方面或者多方面的信息,所以微博案件方面识别通常归类为文本多标签分类问题。然而,社交新闻媒体通常形式和内容都灵活多变,给文本语义理解带来极大的困难,本文针对涉案微博文本观点所含方面的识别问题开展研究,提出一种基于案件相关微博正文和评论交互注意的案件方面识别方法。首先基于 Transformer 框架对正文和评论分别进行编码;然后基于交互注意力机制,实现正文信息和评论信息的融合;最后基于融合后的特征实现对评论文本案件方面的识别。其具体内容性为:

(1) 提出一种文本多标签分类的模型,解决了微博案件的方面识别问题;

(2) 提出了融合正文信息和评论信息的交互注意力机制,提升和增强了评论的语义表征;

(3) 构建了一个刑事案件微博数据集,并基于该数据集验证了所提评论交互注意力机制的有效性。

2 相关工作

方面识别就是要自动识别出一条特定文本评论中粒度最细的评价对象。根据目前的研究工作,可以将方面识别的方法分为传统机器学习方法和基于深度学习的方法。

传统机器学习方法中,Hu 等[1]融合评论文本词性信息,基于 Apriori 算法实现评论文本高频名词和名词短语的自动挖掘。其他典型的机器学习方法有隐马尔可夫模型 HMM(Hidden Markov Model)、条件随机场模型 CRF(Conditional Random Field)等。如 Jin 等[2]采用一种编入词汇的 HMM 模型来提取显式方面;Zhang 等[3]在带监督学习的条件随机场模型中引入词性等特征进行训练,并构建相应领域词典,利用该词典识别产品方面。上述的研究工作多为有监督的学习方法,大多

数无监督方法基于 LDA(Latent Dirichlet Allocation)及其变体,例如,文献[4,5]将主题信息作为方面词项,实现方面信息的挖掘。然而,LDA 并不能很好地从短评论中找到连贯的主题,而且,即使主题和方面可能重叠,也不能保证它们是相同的。

近年来,深度学习方法在方面识别任务中取得了不错的成绩。He 等[6]提出了一种无监督的神经网络结构 ABAE(Attention-based Aspect Extraction),利用预先训练好的词嵌入来增强主题的连贯性,通过重构损失从词嵌入空间中学习各个方面的嵌入。Angelidis 等[7]提出了 MATE(Multi-Seed Aspect Extractor),它使用一些与方面相关的种子词的嵌入来确定 ABAE 中的方面嵌入。Zhao 等[8]借用方面嵌入和种子词的想法,提出了一个生成模型,能自动地从外部信息中收集种子词。除此之外,有部分学者将方面识别当作一个文本序列标注问题,如 Poria 等[9]提出一种基于七层深度卷积网络的模型来对句子进行标记训练,从而识别方面。也有学者将方面识别转化为文本分类问题,如 Liu 等[10]提出一种文本分类模型,用于识别产品评论中所包含的方面,并且取得了不错的效果。基于卷积神经网络 CNN(Convolutional Neural Network)和循环神经网络 RNN(Recurrent Neural Network)的深度学习网络结构在社交媒体文本分类任务中效果显著,但不足之处是不够直观,解释性不好。而注意力机制[11]能够很直观地给出每个词对结果的贡献,如 Yang 等[12]提出的分层注意网络(Hierarchical Attention Network)。Transformer[13]仅依靠注意力机制,能够比神经网络更快速地获取文本特征表达。近年来,有部分学者结合更为丰富的信息对文本进行分类,如 Chai 等[14]将标签类别与类别描述相关联,类别描述由手工制作的模板或使用从强化学习中所提取的抽象模型生成,使分类模型关注与标签描述相关的最显著文本。Wang 等[15]提出了一种标签嵌入注意模型,将词和标签联合嵌入在一潜在空间中,直接利用文本标签的兼容性构造文本表示。Sun 等[16]在基于方面的情感分析 ABSA(Aspect-Based Sentiment Analysis)任务中使用四种不同的句子模板构造辅助句,将 ABSA 转化为句子对分类任务。与单标签分类不同,多标签分类给一个数据样本标注多个标签,能够更加准确、有效地表达多重语义信息。通常一条特定文本评论中包含多个评价对象,因此方面识别任务更类似于文本多标签分类的问题。基于深度学习的多标签分类模型尚处于研究阶段。Kurata 等[17]使用 CNN 进行分类,Chen 等[18]使用 CNN 级联 RNN 获取文章的语义信息。Yang 等[19]提出将序列到序列模型应用到多标签分类中,序列生成模型 SGM(Sequence Generation Model)把标签相关关系考虑在内,并且提出一个新的解码结构的序列生成模型,能够捕获标签之间的相关关系,而且在预测的时候自动选择信息量最丰富的词。Yang 等[20]基于卷积神经网络对文本进行表征,并利用自注意力机制实现文本信息的交互及特征提取,以从源文本中提取细粒度的局部邻域信息和全局交互信息。Xiao 等[21]提出了一个标签特定注意网络 LSAN(Label-Specific Attention Network),LSAN 利用标签语义信息来确定标签和文档之间的语义联系,同时采用自注意力机制从文档内容信息中识别标签特定的文档表示形式。

针对微博评论通常包含多方面观点,微博案件的评论都是围绕微博正文展开讨论,微博正文能给评论的语义理解提供帮助的特点,本文提出一种基于正文和评论交互注意的多标签分类模型,用于识别案件的方面。

3　基于正文和评论交互注意的案件方面识别方法

本文提出一种基于微博正文和评论交互注意思想的微博案件方面识别方法,通过自注意力机制,实现对正文和评论的编码,基于交互注意力机制,实现正文和评论的融合。具体网络结构如图 1 所示。

本文融合正文和评论的案件方面识别模型主要由三个部分组成:基于自注意力机制的正文和评论编码、正文和评论交互注意力编码网络和案件方面识别。(1)基于自注意力机制的正文和评论编码。微博案件的正文和评论作为模型两端的输入,对正文和评论采用同样的编码方式,将正文和评论分别进行表征,得到正文和评论的嵌入矩阵;然后采用多头注意力机制对其编码,计算句子中的每个词与所有词的关注。(2)正文和评论交互注意力编码网络包括 3 个模块:①正文到评论的编码模块表示哪些正文词与每个评论词最相关;②评论到正文的注意力编码模块表示哪些评论词与每个正文词最相似;③交互注意融合模块,通过正文和评论互相关注,融合正文和评论信息,得到包含有正文信息的评论的表征。(3)案件方面识别。对融合后的特征通过线性变换和 sigmoid 函数,将评论映射到

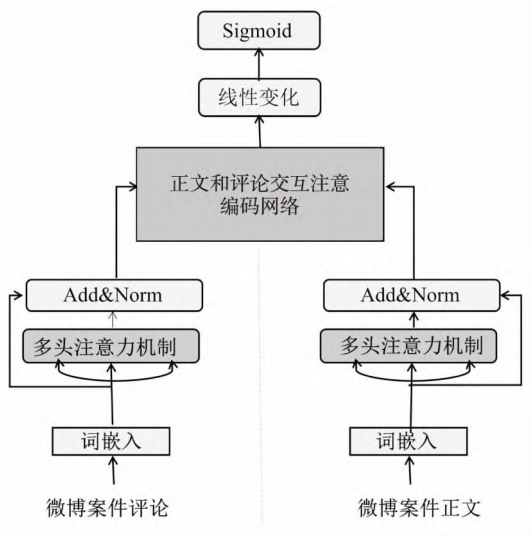

Figure 1 Network structure of aspect identification based on the contents and comments interactive attention network

图 1 基于正文和评论交互注意力网络的案件方面识别网络结构

相应的标签,得到案件方面识别的结果。

3.1 正文和评论交互注意力编码

正文和评论交互注意力编码包括两个模块:基于自注意力机制的正文和评论编码、正文和评论交互注意力编码网络。对正文和评论采用自注意力机制的方式进行编码,得到关于正文和评论的两个表征,再通过交互注意力机制对这两个表征进行融合,得到包含有正文信息的评论语义表征。

3.1.1 基于自注意力机制的正文、评论编码

正文和评论自注意力编码将微博案件的正文和评论作为编码端的两个输入。假设一个句子,其中有 n 个词,可以将其用序列来表示,则正文和评论分别用式(1)和式(2)表示:

$$\boldsymbol{L}=(l_1,l_2,\cdots,l_n) \tag{1}$$

$$\boldsymbol{K}=(k_1,k_2,\cdots,k_n) \tag{2}$$

其中,$\boldsymbol{L}$ 和 $\boldsymbol{K}$ 分别表示正文和文本的序列,n 为序列长度。

对输入的句子序列进行词嵌入,得到词嵌入序列表示,如式(3)和式(4)所示:

$$\boldsymbol{E}_L=(\boldsymbol{w}_{l_1},\boldsymbol{w}_{l_2},\cdots,\boldsymbol{w}_{l_n}) \tag{3}$$

$$\boldsymbol{E}_K=(\boldsymbol{w}_{k_1},\boldsymbol{w}_{k_2},\cdots,\boldsymbol{w}_{k_n}) \tag{4}$$

其中,$\boldsymbol{E}_L$ 和 $\boldsymbol{E}_K$ 是将正文和评论句子表示成一个二维嵌入矩阵的序列,它将句子的所有嵌入连接在一起,其维度大小为 $n\times d$,n 是词的个数,d 为句子嵌入的维度。序列中的每个元素都是相互独立的。

使用多头注意力机制读取每个文本序列并计算每个词与所有词的关注。将二维嵌入矩阵 $\boldsymbol{E}_L$ 和 $\boldsymbol{E}_K$ 转化为查询($\boldsymbol{Q}$)、键($\boldsymbol{K}$)和值($\boldsymbol{V}$),并对其进行线性变化后输入到头数为 8 的多头注意力机制中;然后将所有头的输出值进行拼接;最后通过线性变换层转换成和单头一样的输出值,如式(5)和式(6)所示:

$$\boldsymbol{A}_L=Linear(MH(\boldsymbol{Q}_L,\boldsymbol{K}_L,\boldsymbol{V}_L)) \tag{5}$$

$$\boldsymbol{A}_K=Linear(MH(\boldsymbol{Q}_K,\boldsymbol{K}_K,\boldsymbol{V}_K)) \tag{6}$$

其中,矩阵 $\boldsymbol{A}_L$ 和 $\boldsymbol{A}_K$ 分别表示基于多头注意力网络表征后的正文和评论表征向量。

基于多头注意交互的文本语义表征对每个词进行词嵌入得到关于句子的嵌入矩阵,对嵌入矩阵分别进行线性变换得到相应的查询($\boldsymbol{Q}$)、键($\boldsymbol{K}$)和值($\boldsymbol{V}$)。$\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 为三个完全相同的语义特征向量,分别经过一次线性变换,然后同时并行进行 h 次点积注意交互运算,也就是所谓的多头,头之间参数不共享,每次 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 进行线性变换的参数都是不一样的。然后将 h 次缩放点积注意的结果进行拼接,再进行一次线性变换后得到的值是多头注意力机制的结果,如式(7)所示:

$$MH(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=Concat(\boldsymbol{head}_1,\boldsymbol{head}_2,\cdots,\boldsymbol{head}_h)\boldsymbol{W}^o \tag{7}$$

$$\boldsymbol{head}_i=Attention(\boldsymbol{Q}\boldsymbol{W}_i^{\boldsymbol{Q}},\boldsymbol{K}\boldsymbol{W}_i^{\boldsymbol{K}},\boldsymbol{V}\boldsymbol{W}_i^{\boldsymbol{V}}) \tag{8}$$

其中,d_Q 表示 $\boldsymbol{Q}$ 的特征维度,d_K 表示 $\boldsymbol{K}$ 的特征维度,d_V 表示 $\boldsymbol{V}$ 的特征维度,d_{model} 表示 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 进行线性变换后的维度(即 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 变换为相同维度的向量),参数矩阵 $\boldsymbol{W}_i^{\boldsymbol{Q}}\in\mathbf{R}^{d_{\text{model}}\times d_{\boldsymbol{Q}}}$,$\boldsymbol{W}_i^{\boldsymbol{K}}\in\mathbf{R}^{d_{\text{model}}\times d_{\boldsymbol{K}}}$,$\boldsymbol{W}_i^{\boldsymbol{V}}\in\mathbf{R}^{d_{\text{model}}\times d_{\boldsymbol{V}}}$,$\boldsymbol{W}^o\in\mathbf{R}^{hd_{\text{model}}\times d_{\boldsymbol{V}}}$。

输入由维度为 d_K 的 $\boldsymbol{Q}$ 和 $\boldsymbol{K}$ 以及维度为 d_V 的 $\boldsymbol{V}$ 组成,计算 $\boldsymbol{Q}$ 与 $\boldsymbol{K}$ 的点积,每个键除以 $\sqrt{d_K}$,并应用 softmax 函数来获得这些值的权重,如式(9)所示:

$$Attention(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=softmax(\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}/\sqrt{d_K})\boldsymbol{V} \tag{9}$$

其中,$1/\sqrt{d_K}$ 是缩放因子,起到调节作用,使得分子上的内积不至于太大。

3.1.2 正文和评论交互注意编码网络

正文和评论交互注意编码的思想受 Seo 等[22]在机器阅读理解任务上提出的双向注意力流 BiDAF(Bi-Directional Attention Flow)的启发。正文和评论交互注意力编码网络负责链接和融合正文和评论的信息,编码结构如图 2 所示,其中,$\boldsymbol{K}=\{K_1,\cdots,K_i,\cdots,K_T\}$ 表示基于正文加权后的评论文本表征,$i=1,\cdots,T$ 表示文本序列序号。

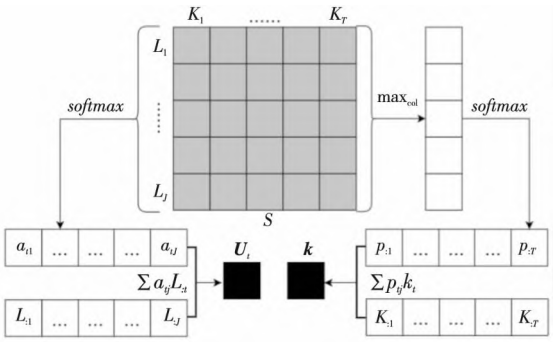

Figure 2　Structure of the contents and comments interaction attention coding network

图 2　正文评论交互注意力编码网络结构

正文和评论交互注意力编码网络包括三个模块:正文到评论的注意力编码模块、评论到正文的注意力编码模块和交互注意融合模块。

交互注意力编码模块的输入是经过多头注意力机制编码得到的正文和评论的表征。编码网络从两个方向计算关注:从评论到正文的关注和从正文到评论的关注,这两个关注来自于正文和评论的上下文嵌入信息之间的共享相似度矩阵 $\boldsymbol{S}$,相似度矩阵计算如式(10)所示:

$$S_{tj}=\alpha(\boldsymbol{A}_{\boldsymbol{K}_{:t}},\boldsymbol{A}_{\boldsymbol{L}_{:j}}) \tag{10}$$

其中,S_{tj} 表示第 t 个评论词和第 j 个正文词之间的相似度;$\alpha(\cdot)$ 表示可训练标量函数,编码两个输入向量之间的相似度;$\boldsymbol{A}_{\boldsymbol{K}_{:t}}$ 表示 $\boldsymbol{A}_{\boldsymbol{K}}$ 的第 t 列向量,$\boldsymbol{A}_{\boldsymbol{L}_{:j}}$ 表示 $\boldsymbol{A}_{\boldsymbol{L}}$ 的第 j 列向量。

(1)正文到评论的注意力编码模块。正文到评论的关注表示哪些正文词与每个评论词最相关。令 $\boldsymbol{a}_t$ 表示所有的正文词与第 t 个评论词之间的注意力权重向量,对于所有的 t,$\sum_t a_{tj}=1$。注意力权重通过 $a_t=softmax(\boldsymbol{S}_{t:})$ 计算求得,则每个参与关注的正文向量表示如式(11)所示:

$$\boldsymbol{U}=\sum\nolimits_j a_{tj}\boldsymbol{A}_{\boldsymbol{L}_{:j}} \tag{11}$$

其中,$\boldsymbol{U}$ 中包含了同所有评论词计算关注的正文向量。

(2)评论到正文的注意力编码模块。评论到正文的注意力表示哪些评论词与每个正文词最相似。通过 $p=softmax(\max_{\text{col}}(\boldsymbol{S}))$ 来获得评论词的注意力权值,其中求最大值的函数 $\max_{\text{col}}(\boldsymbol{S})$ 是跨列执行的。则每个参与关注的评论向量基于加权平均后的表示如式(12)所示:

$$\boldsymbol{U}'=\sum\nolimits_t p_t\boldsymbol{A}_{\boldsymbol{K}_{:t}} \tag{12}$$

其中,t 为正文文本序列序号,p_t 表示第 t 个词的注意力权值,$\boldsymbol{U}'$ 表示某一时刻与正文最相关的评论词的权重和。

(3)交互注意融合模块。交互注意融合模块最后将评论词嵌入和正文评论交互式注意向量进行拼接,得到的矩阵用 $\boldsymbol{G}$ 来表示,如式(13)所示:

$$\boldsymbol{G}=\beta(\boldsymbol{K},\boldsymbol{U},\boldsymbol{U}') \tag{13}$$

其中,$\boldsymbol{G}$ 中的每个列向量可以被看作是每个评论词有正文信息的表征,$\beta(\cdot)$ 是任意可训练的神经网络,它融合了三个输入向量。

3.2　融合正文和评论交互注意信息的案件方面识别

3.2.1　方面识别

通过正文和评论相互关注的过程,融合正文和评论交互注意信息得到的矩阵 $\boldsymbol{G}$ 是包含有正文信息的评论表征。将这个表征经过一个四层的线性变换和一个非线性激活函数,可得到四个类的向量 $\boldsymbol{F}$。再通过 sigmoid 函数对每个类进行二分类,将四个类的向量映射到(0,1),得到类别预测矩阵 $\boldsymbol{P}$。分类过程如式(14)和式(15)所示:

$$\boldsymbol{F}=\tanh(Linear(\boldsymbol{G})) \tag{14}$$

$$\boldsymbol{P}=sigmoid(\boldsymbol{F}) \tag{15}$$

当评论的类别预测结果 $P_{i,j}$ 大于阈值 0.47 时,则认为该评论属于该类别;反之则不属于。将每条评论映射到一类或多类,可实现微博案件方面的识别。

3.2.2　损失函数

本文使用汉明损失 HL(Hamming Loss)来评估模型的精度。汉明损失统计了被误分类的样本数,即不属于这个类别的标签被预测,或者属于这个类别的标签没有被预测。汉明损失的值越小,性能越好。对于一个测试集 $X=\{(\boldsymbol{x}_1,\boldsymbol{Y}_1),(\boldsymbol{x}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{x}_n,\boldsymbol{Y}_n)\}$,其损失计算如式(16)所示:

$$HL=\frac{1}{n}\sum_{i=1}^{n}\frac{XOR(Y_i,P_i)}{L} \tag{16}$$

其中,n 是样本的数量,L 是标签的数量,Y_i 表示第 i 个样本的真实类别标签,P_i 表示第 i 个样本的类别预测标签。$XOR(\cdot)$ 表示异或运算,且满足 $XOR(0,1)=XOR(1,0)=1$,$XOR(0,0)=XOR(1,1)=0$。

4　实验

本节首先介绍本文提出的数据集、实验参数、评价指标、实验细节和所有的基线方法;然后,将本文方法与基线方法进行比较,并通过消融实验验证文本和评论的双向交互注意力模块的有效性;最

后,对实验结果进行分析和讨论。

4.1 数据集

本文从微博爬取了 12 个热点案件的 7 000 条数据,每条数据均包含一个正文及相应的一条评论。在 7 000 条数据中,2 973 条数据包含对嫌疑人的评价,1 872 条数据包含对被害人的评价,879 条数据包含对案由的评价,3 772 条数据包含对其他方面的评价。大部分的评论数据有 2 个标签,部分数据有 3 或 4 个标签,少部分数据有 1 个标签,平均标签数约为 2 个。按照 8∶1∶1 的比例划分数据集,训练集、验证集和测试集中均包含 12 个案件的数据,其中训练集 5 600 条,验证集和测试集各 700 条。

4.2 实验参数设置

本文使用 PyTorch 深度学习框架编码模型。微博案件文本的词向量维度为 512 维,词表大小为 3 076。多头注意力机制中的 d_{model} 大小为 512 维,头数 h 为 8。交互式注意力机制融合正文和评论信息,对得到的交互注意信息矩阵进行四层线性变换,预测到四个类别,然后使用 sigmoid 函数输出标签空间上的概率分布。其中每层线性变换的隐藏单元数为 128。模型的学习率设置为 0.000 3,随机失活率设置为 0.5。

4.3 基线方法

针对本文所提出的数据集,本文设置了以下几个对比实验。基线方法的对比实验均不融合案件正文信息。

(1)CNN[17]:使用多个卷积核提取微博案件评论文本特征,输入到线性变换层,然后输出标签空间上的概率分布。

(2)CNN-RNN[18]:利用 CNN 和 RNN 获取全局和局部的案件评论文本语义,并对标签相关性进行建模。

(3)Transformer[13]:在文本分类中只用到 Transformer 的编码端,获取微博案件评论的文本特征,再输出标签的概率。

(4)ABAE[6]:将词共现统计数据显式地编码为词嵌入,使用降维方法提取评论语料中最重要的方面,并通过方面嵌入的线性组合来重构每个句子,使用注意力机制去除不相关的词,以进一步提高各方面的连贯性。

4.4 实验分析

4.4.1 评价指标

本文采用准确率 P(Precision)、召回率 R(Recall)和 $Weigthed\text{-}F1$ 值作为所提方法有效性的评价指标。

$Weight\text{-}F1$ 的计算方法为:先计算出每一个类别的 P、R 和 $F1$,然后通过求均值得到整个样本集上的 $F1$ 值。在此基础上,再根据每一类别的数据大小进行加权平均。

4.4.2 对比实验

在本文所提出的数据集上,本文所提方法与不同基线方法进行了对比实验,实验结果如表 2 所示。

Table 2 Experimental results of different approaches aspects identification

表 2 不同方面识别方法的实验结果

方面识别方法	P	R	$F1$
CNN	0.725 8	0.821 3	0.781 9
CNN-RNN	0.713 6	0.839 4	0.778 2
Transformer	0.732 4	0.816 9	0.765 8
ABAE	0.734 3	0.872 9	0.791 9
本文方法	0.764 1	0.854 5	0.803 5

表 2 的实验结果表明,在本文所提出的数据集上,相比基线方法,本文所提方法通过交互注意力机制融合正文信息,在主要评价指标上性能最好。本文提出的基于正文和评论交互注意的方面识别方法,与最常用的 CNN 方法相比,准确率提高了 3.83 个百分点,召回率提高了 3.32 个百分点,$F1$ 值提高了 2.16 个百分点。相比于 CNN-RNN 方法,本文所提方法精确率提高了 5.05 个百分点,召回率提高了 1.51 个百分点,$F1$ 值提高了 2.53 个百分点。本文只使用了 Transformer 模型中的多头注意力部分,采用 Transformer 模型的编码端对评论进行编码,然后对其进行多标签分类。相比之下,本文所提方面识别方法在准确率上提高了 3.17 个百分点,召回率提高了 3.76 个百分点,$F1$ 值提高了 3.77 个百分点。基于 LDA 的无监督方法难以从评论中获取连贯的主题。ABAE 是方面提取的经典方法,对于本文所使用的数据集,ABAE 模型首先将微博案件的评论数据通过 word2vec 训练成词向量,使用注意力层关注句子中最重要的信息,实验结果准确率为 73.43%,召回率为 87.29%,$F1$ 值为 79.19%。由表 3 可知,ABAE 没有融入正文信息,本文所提方法在方面识别性能上优于 ABAE 方法。

由此可见,本文通过融合正文信息的方式,增强了评论的语义表征,使分类的性能得到了提升。

此外,本文构建的数据集中包含嫌疑人和其他的评论较多,包含案由的评论较少,存在数据不均衡的问题,因此还计算了每一类的准确率、召回率和 $F1$ 值。各类别的有效性验证结果如表 3 所示。

Table 3　Effectiveness verification results of each category

表 3　各类别的有效性验证结果

类别	P	R	$F1$
嫌疑人	0.781 0	0.998 0	0.876 2
被害人	0.822 2	0.758 8	0.762 2
案由	0.413 4	0.431 9	0.448 6
其他	0.657 3	0.724 1	0.805 9

分析表 3 可得,嫌疑人、被害人和其他的评论数量较大,其 $F1$ 值较高,案由评论数量较少,其 $F1$ 值较低,但在加权平均后对总体的 $F1$ 值影响较小。

4.4.3　消融实验

为验证本文所提方法的有效性,本节将交互注意力层消去进行比较。

在表 4 中,"消去交互注意力网络"表示去掉正文和评论交互注意力网络,"消去评论到正文注意力编码"表示交互注意力网络中只保留正文到评论的注意力,"消去正文到评论注意力编码"表示交互注意力网络中只保留评论到正文的注意力。分析表 4 实验结果可知,通过交互注意力的方式融合正文的信息,增强了评论的语义表征,同时,相比于单向的关注,双向关注对评价对象方面识别的性能有明显的提升。

Table 4　Results of ablation experiment

表 4　消融实验结果

方面识别方法	P	R	$F1$
本文方法	0.764 1	0.854 5	0.803 5
消去交互注意力网络	0.701 8	0.823 2	0.748 1
消去评论到正文注意力编码	0.734 3	0.857 4	0.766 7
消去正文到评论注意力编码	0.774 3	0.831 0	0.754 2

4.4.4　案例分析

在对不同标签进行预测时,不同词语的贡献也存在差异。表 5 给出几个例子的注意力层可视化,表 6 给出这几个例子使用不同方法得到的相应标签。当预测标签"嫌疑人"时,它可以自动为信息更丰富的词分配更大的权重,比如十四岁、男孩和施暴者等。对于标签"被害人",所选的信息词为十一岁、女孩和受害者等。对于标签"案由",所选的信息词为性侵和杀人等。对于标签"其他",所选的信息词为正文中除上述三方面外的关键词、正文中没有出现的信息词等。这表明本文所提方法在预测不同的标签时,能够自动在评论中选择信息最丰富的词语。

Table 5　Visualization of interactive attention mechanism

表 5　交互注意力机制可视化

微博正文:近日,有网友在微博反映,10 月 20 日,大连市沙河口区有一名*11 岁女孩被一名14 岁男孩杀害,女孩身中 7 刀*。当时男孩将美术班补课结束路过的女孩带至家中,想要与其发生性关系遭拒绝,便将女孩杀害并抛尸。依据《刑法》第十七条第二款之规定,加害人蔡某某未满 14 周岁,未达到法定刑事责任年龄,依法不予追究刑事责任。同时,公安机关依据《刑法》第十七条第四款之规定,按照法定程序报经上级公安机关批准,于 10 月 24 日依法对蔡某某收容教养。
评论 1:十四岁的施暴者需要被法律保护,那十一岁的被害者却没有得到真正意义上的保护,十分期待法律的完善,不要给罪恶一点庇护。
评论 2:建议性侵和杀人等恶劣行为的处罚不应该受刑法最低年龄的影响。杀人就是杀人,跟加害人多少岁没有关系。若是低于 14 岁完全不受刑法管辖,那他为所欲为,杀人放火,等他接受教育完出来,又是一条"好汉",搬个家之后他完全没有受到任何影响,那受害者呢。
评论 3:这应该属于不可回收垃圾了吧!

Table 6　Examples of label generation

表 6　标签生成案例

评论	真实标签	CNN	本文方法	消去交互式注意力网络
评论 1	嫌疑人、被害人、其他	嫌疑人、被害人、其他	嫌疑人、被害人、其他	嫌疑人、被害人
评论 2	嫌疑人、被害人、案由、其他	嫌疑人、被害人、其他	嫌疑人、被害人、案由、其他	嫌疑人、被害人、其他
评论 3	嫌疑人	其他	其他	其他

以表 5 中的 3 条评论生成的标签序列为例,将本文方法与最常用的基线方法 CNN 进行了比较,结果如表 6 所示。与真实标签相比,CNN 模型能够预测评论 1 这样信息词比较明显的相关标签,然而因为未加入正文信息,CNN 在预测评论 2 时没有预测出完整的标签,在预测评论 3 时甚至出现了错误。消去交互注意力机制导致方法没有正文信息,会出现与 CNN 模型一样的相关标签预测不完整或者错误的问题。本文所提方法融合了正文的信息,可以准确地预测出评论 1 和评论 2 这样信息词较为明显的标签,在预测评论 3 时,由于主语不明确,模型预测出现了错误。但是,总体而言,本文所提方法由于融入了正文信息,增强了评论的语义表征,相比于其他方面识别方法,预测标签的准确率有明显的提升。

5　结束语

针对微博案件观点所涉方面的智能识别问题,

本文提出一种基于微博案件正文和评论交互注意力的多标签分类方法,实现对微博案件评论文本中方面的自动识别。通过与基线方法进行比较、消去交互注意力机制各项实验,结果表明交互注意力机制融合了正文信息,对评论中方面的自动识别有良好的效果。

参考文献:

[1] Hu M,Liu B. Mining and summarizing customer reviews[C]//Proc of the 10th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,2004:168-177.

[2] Jin W,Ho H H,Srihari R K. Opinion miner:A novel machine learning system for Web opinion mining and extraction[C]//Proc of the 15th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,2009:1195-1204.

[3] Zhang S,Jia W J,Xia Y J,et al. Opinion analysis of product reviews[C]//Proc of the 6th International Conference on Fuzzy Systems & Knowledge Discovery,2009:591-595.

[4] Mei Q,Ling X,Wondra M,et al. Topic sentiment mixture:Modeling facets and opinions in weblogs[C]//Proc of the 16th International Conference on World Wide Web,2007:171-180.

[5] Wang S,Chen Z,Liu B. Mining aspect-specific opinion using a holistic lifelong topic model[C]//Proc of the 25th International Conference on World Wide Web,2016:167-176.

[6] He R,Lee W S,Ng H T,et al. An unsupervised neural attention model for aspect extraction[C]//Proc of the 55th Annual Meeting of the Association for Computational Linguistics,2017:388-397.

[7] Angelidis S,Lapata M. Summarizing opinions:Aspect extraction meets sentiment prediction and they are both weakly supervised[C]//Proc of the 23rd Conference on Empirical Methods in Natural Language Processing,2018:3675-3686.

[8] Zhao C,Chaturvedi S. Weakly-supervised opinion summarization by leveraging external information[C]//Proc of the 34th AAAI Conference on Artificial Intelligence,2020:9644-9651.

[9] Poria S,Cambria E,Gelbukh A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis[C]//Proc of the 2015 Conference on Empirical Methods in Natural Language Processing,2015:2539-2544.

[10] Liu P,Joty S,Meng H,et al. Fine-grained opinion mining with recurrent neural networks and word embeddings[C]//Proc of the 2015 Conference on Empirical Methods in Natural Language Processing,2015:1433-1443.

[11] Bahdanau D,Cho K,Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv:1409.0473,2014.

[12] Yang Z,Yang D,Dyer C,et al. Hierarchical attention networks for document classification[C]//Proc of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics:Human Language Technologies,2016:1480-1489.

[13] Vaswani A,Shazeer N,Parmar N,et al. Attention is all you need[C]//Proc of the 2017 Neural Information Processing Systems,2017:5998-6008.

[14] Chai D,Wu W,Han Q H,et al. Description based text classification with reinforcement learning[J]. arXiv:2002.03067,2020.

[15] Wang G,Li C,Wang W,et al. Joint embedding of words and labels for text classification[C]//Proc of the 56th Annual Meeting of the Association for Computational Linguistics,2018:2321-2331.

[16] Sun C,Huang L,Qiu X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence[C]//Proc of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics:Human Language Technologies,2019:380-385.

[17] Kurata G,Xiang B,Zhou B,et al. Improved neural network-based multi-label classification with better initialization leveraging label cooccurrence[C]//Proc of the 2016 Conference of the North American Chapter of the Association for Computational Linguistic:Human Language Technologies,2016:521-526.

[18] Chen G,Ye D,Xing Z,et al. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization[C]//Proc of the 2017 International Joint Conference on Neural Networks,2017:2377-2383.

[19] Yang P,Sun X,Li W,et al. SGM:Sequence generation model for multi-label classification[C]//Proc of the 9th International Conference on Computational Linguistics,2018:3915-3926.

[20] Yang Z Y,Liu G J. Hierarchical sequence-to-sequence model for multi-label text classification[J]. IEEE Access,2019,7:153012-153020.

[21] Xiao L,Huang X,Chen B L,et al. Label-specific document representation for multi-label text classification[C]//Proc of the 24th Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP),2019:466-475.

[22] Seo M,Kembhavi A,Farhadi A,et al. Bidirectional attention flow for machine comprehension[J]. arXiv:1611.01603,2016.

作者简介:

段玲(1996-),女,云南大理人,硕士生,研究方向为自然语言处理和深度学习。**E-mail**:1401307663@qq.com

DUAN Ling,born in 1996,MS candidate,her research interests include natural language processing,and deep learning.