

面向方面抽取与情感分类的多任务联合建模

孟甜甜, 韩 虎⁺, 吴渊航

兰州交通大学 电子与信息工程学院, 兰州 730070

+ 通信作者 E-mail: hanhu_lzjtu@mail.lzjtu.cn

摘要:基于方面的细粒度情感分析包括方面术语抽取和方面情感分类两项任务,以独立方式解决以上两项任务的研究方法无法利用彼此之间的关联信息,同时也会造成训练冗余和资源浪费。针对上述问题,在多任务学习框架下提出一种基于位置嵌入和图卷积网络的联合模型(PE-GCN),以端到端方式整体解决方面术语抽取和方面情感分类。该模型首先通过双向门控循环单元网络学习句子的语义特征表示;随后利用位置嵌入增强句子中方面术语的识别,同时使用图卷积网络生成包含句法信息的上下文表示;最后通过交互注意力网络建模上下文和方面术语之间的语义关系,并通过softmax输出方面术语的情感极性。在SemEval-2014公开数据集上的实验结果表明,提出的模型与其他现有模型相比性能有显著的提升。

关键词:方面术语抽取;方面情感分类;位置嵌入;图卷积网络;交互注意力

文献标志码:A **中图分类号:**TP391

Joint Modeling Based on Multi-task Learning for Aspect Term Extraction and Sentiment Classification

MENG Tiantian, HAN Hu⁺, WU Yuanhang

School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: Fine-grained aspect-based sentiment analysis involves aspect term extraction and aspect sentiment classification. Most existing research methods address them in an independent fashion, which lack a mechanism to account for the relevant information between each other, resulting in training redundancy and waste of resources. To solve the above problems, a joint model based on position embedding and graph convolutional network (PE-GCN) under the framework of multi-task learning is proposed, which is an end-to-end approach to the overall solution of aspect term extraction and aspect sentiment classification. Firstly, the model learns the semantic feature representation of sentence through a bidirectional gated recurrent unit network. Then, it exploits positional embedding to enhance the recognition of aspect terms in sentence, and uses the graph convolutional network to generate a contextual representation containing syntactic information. Finally, interactive attention network is used to model the semantic relationship between context and aspect terms, and the sentiment polarity of aspect terms is output through softmax. Experimental results on the SemEval-2014 public datasets show that the performance of the proposed model has a significant improvement compared with other existing models.

Key words: aspect term extraction; aspect sentiment classification; position embedding; graph convolutional network; interactive attention

基金项目:国家自然科学基金(62166024)。

This work was supported by the National Natural Science Foundation of China (62166024).

收稿日期:2022-03-07 **修回日期:**2022-05-06

基于方面的细粒度情感分析(fine-grained aspect-based sentiment analysis, ABSA)是对评论语句中特定评价方面给出情感极性的研究^[1],其包含两个基本任务,即方面术语抽取和方面情感分类。其中,方面术语抽取(aspect term extraction, ATE)是检测评论语句中提到的评价方面^[2];方面情感分类(aspect sentiment classification, ASC)则是预测给定方面的情感极性^[3]。例如,在句子“Although the food in this restaurant is delicious, the price is very expensive.”中,ATE的任务就是将“food”“price”标记为两个方面术语,ASC的任务则是对方面术语“food”和“price”分别给出它们对应的情感极性“positive”和“negative”。

传统研究通常以独立的方式解决以上两个任务^[4-6]。在方面术语抽取问题上,目前大部分研究将其看作一个序列标注问题,条件随机场(conditional random fields, CRF)和深度学习理论成为解决方面术语抽取的主流方法。Jakob等人^[7]考虑到评论文本包含的词特征、词性特征、依赖句法特征、词距特征和观点句特征等外部知识可以提供丰富的指示信息,提出基于CRF的意见目标提取方法,解决了领域可移植性问题。Ma等人^[8]指出基于序列标注的方法不能充分利用句子的整体含义,并且在处理标签依赖问题时存在局限性,提出采用门控单元网络将相应的单词表示加入到解码器中,并利用位置感知注意力对目标单词的相邻单词加以关注。在方面情感分类问题上,主要考虑到评价方面与上下文之间的互相作用,基于注意力机制的深度学习方法取得了较好的成绩。例如, Ma等人^[9]针对方面术语由多个单词构成时,采用双注意力交互方式分别对上下文和目标词进行建模,并通过注意力机制对目标短语中的不同词进行区别处理。Zhang等人^[10]提出一种基于图卷积神经网络的情感分类模型,该模型利用图卷积网络(graph convolutional network, GCN)对方面术语与上下文的长距离依赖关系进行了建模,并利用注意力机制实现权重分配。

虽然方面术语抽取和方面情感分类研究都取得了一定进展,但单一任务的独立学习往往忽略了任务间的信息交互和参数共享,致使训练冗余重复和学习资源的浪费,也限制了其性能的提升。因此,基于多任务学习的联合建模成为方面级情感分析的研究热点^[11-14]。例如, Li等人^[15]采用联合标签机制对ABSA任务进行统一预测,模型包含两个堆叠的递归神经网络,上层网络预测联合标签,下层网络辅助方面边界预测,并引导上层网络提高情感预测任务的

性能。Luo等人^[16]指出联合标签容易使学习到的表示产生混淆,提出利用双递归神经网络(recurrent neural network, RNN)生成方面术语和上下文的各自表示,并设计交叉共享单元考虑两者之间的关系,然后集成方面长度增强和情感增强两个辅助任务提高方面术语和上下文的表示,最后利用门控机制输出方面极性对。He等人^[17]指出上述方法仅从方面级别的实例中学习,且忽略了方面术语抽取和方面情感分类任务的相关性建模问题,提出使用交互式多任务学习网络,引入消息传递机制同时解决方面术语抽取和方面情感分类。Akhtar等人^[18]提出的方法首先使用双向长短时记忆网络(bidirectional long short-term memory network, Bi-LSTM)和自注意力机制识别句子中的方面术语,随后利用卷积神经网络(convolutional neural network, CNN)预测方面术语的情感极性。以上研究针对方面抽取与情感分类任务之间的内在联系,在多任务学习框架下建立联合模型,通过联合训练和优化促进两个任务之间的信息交互,从而达到相互学习的目的。

然而,上述联合模型没有考虑评论语句的句法结构信息,尤其是对评论语句中上下文间的依赖关系缺少考虑。此外,方面术语的周围通常会伴有一些修饰词,但是目前广泛使用的注意力机制通常忽略了位置对于每个单词权重的影响。为此,本文提出一个基于位置嵌入和图卷积网络的方面情感分析联合模型。本文主要贡献如下:

(1)提出一种基于位置嵌入和图卷积网络的联合模型(joint model based on position embedding and graph convolutional network, PE-GCN),以端到端的方式整体解决方面术语抽取和方面情感分类。

(2)利用图卷积网络生成了包含句法信息的上下文表示,利用位置嵌入改善了方面术语的识别,通过交互注意力网络促进了上下文和方面术语之间的信息交互。

(3)在SemEval-2014公开数据集上的实验结果表明,该模型对不同数据集的准确率(accuracy, Acc)与F1值均有所提升,有力证明了本文提出模型的有效性。

1 相关工作

1.1 方面级情感分析

近几年,基于方面的细粒度情感分析研究获得了很大的关注,同时也取得了一定进展。循环神经网络、注意力机制和图神经网络被广泛使用。例如:

Tang 等人^[13]提出将一条评论语句划分为包含有特定目标的左右两部分,然后通过两个 LSTM(long short-term memory)模型分别对以上两部分进行建模,在建模过程中分别针对是否将目标词嵌入到上下文的情况,给出了两种模型结构 TC_LSTM(target-connection long short-term memory)和 TD_LSTM(target-dependent long short-term memory)。Ma 等人^[19]提出一种基于 LSTM 和注意力机制的交互式网络模型,该方法使用两个注意力网络对目标词和上下文进行交互建模,可以密切关注目标词和上下文中的重要部分。Zhang 等人^[19]提出利用双重图卷积网络完成方面级情感分析任务,除了普通的句法图外,还利用词汇图获取训练语料库中的全局词共现信息。Bai 等人^[20]在关系图注意力网络的基础上集成了类型化的句法依赖信息。与此同时,研究人员对方面术语抽取也开展了一系列研究,主要包括下面三类:基于规则的方法、基于传统机器学习的方法和基于深度学习的方法。基于规则的方法通过人工构建规则模板来抽取方面术语,基于传统机器学习的方法主要使用条件随机场抽取方面术语,基于深度学习的方法大多将方面抽取问题看作序列标注问题。例如:Li 等人^[5]考虑到评论文本包含的意见词以及方面历史信息等可以提供丰富的指示信息,提出基于意见摘要和方面检测历史的方面抽取模型。Wang 等人^[11]提出使用递归神经网络和条件随机场相结合的方法进行方面词和意见词的抽取。Ma 等人^[8]提出一种基于序列到序列(Seq2Seq)的学习方法来解决方面术语抽取任务,解决了单词间的标签依赖问题。Phan 等人^[21]利用自注意力机制学习句法信息,并结合词性嵌入、依赖嵌入和上下文嵌入提高了方面提取器的性能。

然而,上述方法都是针对 ATE 或 ASC 提出的,但单一任务的独立学习往往忽略了任务间的信息交互。因此,Li 等人^[15]采用统一标签方案,以端到端的方式解决了基于特定目标的情感分析任务。Luo 等人^[16]重点研究了两种独立路径之间的交互作用,提出一种交叉共享 RNN 框架,同时生成输入句子中的各

方面极性对。He 等人^[17]提出一个交互式多任务学习网络,引入消息传递架构,通过一组共享的潜在变量将消息迭代地传递给不同的任务。Akhtar 等人^[18]提出以端到端机制处理方面术语抽取和情感分类任务,旨在一次性识别方面术语并对其进行情感分类。虽然这些结合注意力机制的深度学习模型取得了比仅使用神经网络模型更好的效果,但此类方法多数使用单一层面注意力机制,没有对上下文信息进行更深层次的挖掘,且使用的神经网络存在训练速度慢,无法获得全局句法结构信息等缺点。

1.2 注意力机制

注意力机制(attention mechanism)作为一种资源分配方案,将有限的计算资源用来处理更重要的信息^[22]。在句子级的文本语义特征表示中,假设 S_i 表示通过 Attention 机制得到的句子语义特征表示,则计算如下:

$$S_i = \sum_{j=1}^l \alpha_j^i h_j^i \quad (1)$$

其中, α_j^i 表示第 j 个词的重要度, h_j^i 表示第 j 个 GRU 隐藏单元的输出。Attention 权重 α_j^i 的计算如下:

$$\alpha_j^i = \frac{\exp(g(h_j^i))}{\sum_{k=1}^l \exp(g(h_k^i))} \quad (2)$$

假设, $g(\cdot)$ 是计算单词重要性程度的函数,定义如下:

$$g(h_j^i) = V^T \tanh(W_h h_j^i + b_h) \quad (3)$$

其中, V 是参数向量, V^T 是转置向量, W_h 是参数矩阵, b_h 是偏置项。

1.3 图卷积网络

1.3.1 句法依赖树

本文使用 spaCy 库生成句子的句法依赖树。以“Although the food is delicious, the price is very expensive.”为例,其句法依赖树结构如图 1 所示。其中,绿色字体表示方面术语,黄色字体表示意见词。根据单词之间的句法依赖信息,方面术语“food”通过



图1 句法依赖树

Fig.1 Syntactic dependency tree

依赖关系“nsubj”和“acomp”关注到对应的意见词“delicious”，方面术语“price”通过相应的依赖关系关注到意见词“expensive”。为了充分利用句法结构，本文使用图卷积网络编码句法依赖树，以捕获单词之间的句法依赖关系。

1.3.2 图卷积网络

图卷积网络是卷积神经网络的改进，用于处理非结构化数据的局部信息^[23]。对于给定 n 个节点的图，通过枚举法获得邻接矩阵 $A \in \mathbb{R}^{n \times n}$ 。为方便起见，将节点 i 的初始状态记为 h_i^0 ，在第 l 层的输出记为 h_i^l 。对于 L 层图卷积网络， $l \in (1, 2, \dots, L)$ ， h_i^l 是节点 i 的最终状态。则第 l 层的节点更新定义如下：

$$h_i^l = \sigma \left(\sum_{j=1}^n A_{ij} W^l h_j^{l-1} + b^l \right) \quad (4)$$

其中， W^l 和 b^l 分别是权重和偏置项， σ 是非线性函数，例如 ReLU。

2 PE-GCN 模型

受文献[18]的启发，本文在多任务学习框架下建立联合模型，在抽取方面术语的同时对方面术语的情感极性进行判断，模型结构如图2所示。主要包括词嵌入与 Bi-GRU 层、位置嵌入层、图卷积网络层、交互注意力层和标签分配五部分。

2.1 任务定义

本文的目的是针对方面术语抽取和方面情感分类之间密切相关性，在多任务学习框架下建立联合模型，通过共同训练和优化同时解决方面术语抽取和情感分类的问题。即给定包含 n 个单词的评论语句 $S = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ ，对于单词 x_i ，ATE 的目标是给 x_i 分配一个标签 $y_i^a \in Y^a$ ，ASC 的目标是给 x_i 分配一个标签 $y_i^p \in Y^p$ ，其中， $Y^a \in \{B, I, E, S, O\}$ ， $Y^p \in \{POS, NEG, NEU\}$ 。标签 B 、 I 、 E 表示方面术语组块的开始、内部和结束， S 表示方面术语只有一个词， O 表示组块外的词语。 POS 、 NEG 、 NEU 代表极性类别，分别是积极的、消极的和中立的。

2.2 词嵌入与 Bi-GRU 层

给定包含 n 个单词的评论语句 $S = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ ， $x_i \in \mathbb{R}^d$ 是单词嵌入。本文首先使用嵌入矩阵 $E \in \mathbb{R}^{|\mathcal{V}| \times d}$ 将每个单词映射到一个低维的实值向量空间中，其中， $|\mathcal{V}|$ 代表词汇表的大小， d_e 代表词嵌入的维度；其次，为了学习整个句子的词嵌入，将初始化的单词向量输入到 Bi-GRU 层获得句子 S 的隐藏状态 $H = \{h_1, h_2, \dots, h_i, \dots, h_n\}$ ，其中， $h_i \in \mathbb{R}^s$ 代表在 Bi-GRU 处于时间 t 的隐藏状态向量， s_e 是隐藏状态向量输出的维度大小；同时通过 copy 操作生成上下文的隐藏状态 $H^c = \{h_1^c, h_2^c, \dots, h_i^c, \dots, h_n^c\}$ 。

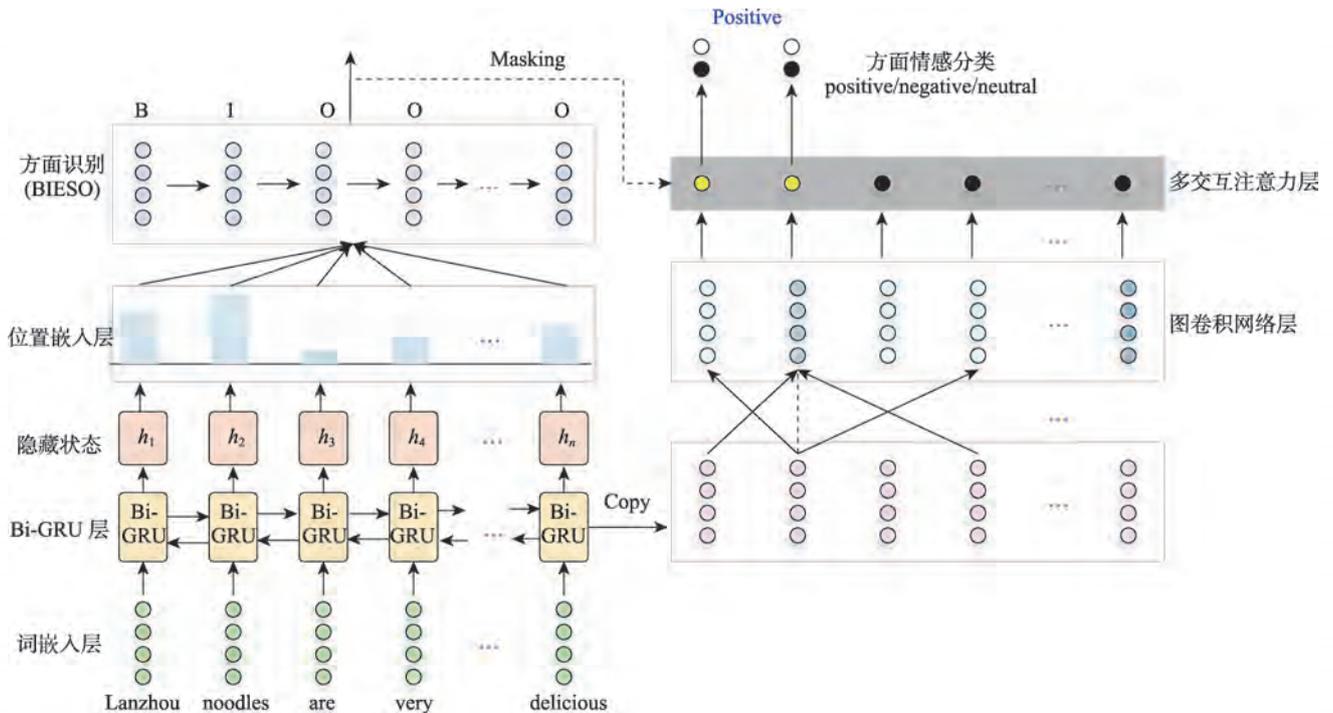


图2 PE-GCN 模型

Fig.2 PE-GCN model

2.3 位置嵌入层

在方面术语抽取任务中,每个单词的相邻单词对预测其标签具有重要的影响,而距离较远的词语则对其标签预测的影响较小。原因是,方面术语的周围通常会伴有一些修饰词,但是目前广泛使用的注意力机制通常忽略了位置对于每个单词权重的影响。因此,本文提出用位置嵌入(position embedding, PE)根据单词 i 和单词 t 之间的相对距离有规律地降低单词 i 的权重。假定计算位置 t 处的上下文向量 c_t , PE 首先计算每个单词的权重:

$$\alpha_i = \frac{\exp(f(s_{t-1}, h_i))}{\sum_{j=1}^n \exp(f(s_{t-1}, h_j))} \quad (5)$$

其中, $f(s_{t-1}, h_i)$ 是得分函数,用来计算先前结合句子 S 的标签隐藏状态 s_{t-1} 和相应距离得出的 h_i 的权重。得分函数定义如下:

$$f(s_{t-1}, h_i) = \frac{1}{d(w_i, w_t)} (\mathbf{W}_s [s_{t-1}, h_i] + \mathbf{b}_s) \mathbf{V}_s^T \quad (6)$$

其中, $\frac{1}{d(w_i, w_t)}$ 用来计算单词 i 的权重衰减率, \mathbf{W}_s 和 \mathbf{b}_s 分别是权重矩阵和偏差。 \mathbf{V}_s^T 是权重向量 \mathbf{V}_s 的转置。最后,上下文向量 c_t 由句子 S 的隐藏状态的加权总和计算,定义如下:

$$c_t = \sum_{i=1}^n \alpha_i h_i \quad (7)$$

PE 根据距离调整每个单词的权重,因此,和自注意力相比,本文模型更加关注其相邻单词对预测词的影响。

2.4 图卷积网络层

2.4.1 距离编码

在句法依赖树中,语法距离反映方面术语与上下文的语法关联程度,即与方面术语越相关的上下文语法距离越小,与方面术语越不相关的上下文语法距离越大。语法距离可通过句法依赖树得到,根据文献[24],语法距离向量可以表示为 $D = \{d_1, d_2, \dots, d_i, \dots, d_n\}$, 其中, d_i 为上下文词 x_i 与方面术语 x_a 间的语法距离。则句子中每个词的语法距离权重计算公式如下:

$$l_i = 1 - \frac{d_i}{2d_{\max}}, 1 \leq i < a + 1, a + m < i \leq n \quad (8)$$

根据每一个方面术语可计算得到一个与句子 S 长度相同的语法距离特征向量 $L = (l_1, l_2, \dots, l_i, \dots, l_n)$ 。 d_{\max} 表示语法距离向量 D 中的最大值。由式(8)可知,语法距离权重 $l_i \in [0.5, 1.0]$, 这样可以使距离方面

术语远的上下文的最小权重在一定范围内,并且保证了与方面术语距离更近的上下文权重更高,避免了在预测方面术语的情感极性时被与该方面词无关的情感词误导。

假定方面术语由 m 个词组成,将距离句法依赖树根节点最近的单词作为方面词。相对距离可以由上下文词 x_i 和方面术语 x_a 的位置长度得到。则相对距离权重计算如下:

$$l'_i = \begin{cases} 1 - \frac{a+1-i}{n}, & 1 \leq i < a+1 \\ 0, & a+1 \leq i \leq a+m \\ 1 - \frac{i-a-m}{n}, & a+m < i \leq n \end{cases} \quad (9)$$

由此可得到方面术语和上下文之间的相对距离特征向量 $L' = (l'_1, l'_2, \dots, l'_i, \dots, l'_n)$ 。然后,结合语法距离权重和相对距离权重生成新的距离权重表示:

$$q_i = \frac{l_i + l'_i}{2} \quad (10)$$

其中, q_i 表示句子 S 中方面术语与上下文之间的语法距离权重。最后,引入 q_i 更新句子的隐藏状态,更新后的句子隐藏状态中包含了单词的位置信息和语法信息。更新定义如下:

$$h_i = q_i h_i \quad (11)$$

2.4.2 句法依赖

在图卷积网络中,其卷积核对每个节点的计算中权重参数都是共享的,每个节点都能通过相邻节点更新当前单词的表示。当图卷积网络是多层堆叠时,每个节点都可以接收直接相连的邻居节点以外的节点信息。

假设GCN有 L 层,在PE-GCN模型中,将句子的隐藏状态信息、依赖树的邻接矩阵和语法距离权重向量输入图卷积网络,第 l 层的节点更新定义如下:

$$\tilde{h}_i^l = \sum_{j=1}^n A_{ij} \mathbf{W}^l h_j^{l-1} \quad (12)$$

$$h_i^l = \text{ReLU} \left(\frac{\tilde{h}_i^l}{\text{degree}_{(i)} + 1} + \mathbf{b}^l \right) \quad (13)$$

其中, A_{ij} 表示单词 i 和单词 j 在句法上的依赖关系, h_j^{l-1} 是 $l-1$ 层GCN中第 j 个词的特征表示, $\text{degree}_{(i)}$ 表示依赖树中第 i 个词的度, h_i^l 是当前 l 层GCN中第 i 个词的输出。权重 \mathbf{W}^l 和偏置项 \mathbf{b}^l 是可训练参数。则 L 层GCN的最终输出结果为 $H^L = (h_1^L, h_2^L, \dots, h_i^L, \dots, h_n^L)$, $h_i^L \in \mathbb{R}^s$ 。

2.5 多交互注意力层

多交互注意力(multi-interactive attention, MIA)

既考虑了方面术语的语义特征,又考虑了上下文的语义特征,并对两部分特征进行融合,挖掘对情感分析最重要的语义信息,提高了细粒度情感分析的准确性。首先,对方面术语和上下文的语义表示进行平均池化,分别得到两部分的初始整体表示状态 \bar{a} 和 \bar{c} ,计算公式如下:

$$\bar{a} = \frac{s_i^a}{k} \quad (14)$$

$$\bar{c} = \frac{s_i^c}{n} \quad (15)$$

其中, s_i^a 和 s_i^c 分别表示方面术语和上下文的最终语义表示特征, n 表示句子中的单词个数。其次,以方面术语的初始表示作为输入,利用 MIA 选择上下文中对判断情感极性比较重要的信息,给定上下文的语义表示 $S^c = \{s_1^c, s_2^c, \dots, s_i^c, \dots, s_n^c\}$ 和方面术语的初始整体表示 \bar{a} 生成注意力向量 α , 定义如下:

$$\alpha_i = \frac{\exp(\Phi(s_i^c, \bar{a}))}{\sum_{j=1}^n \exp(\Phi(s_j^c, \bar{a}))} \quad (16)$$

其中, Φ 是计算上下文重要程度的得分函数,计算方式如下:

$$\Phi(s_i^c, \bar{a}) = \tanh(s_i^c \cdot W_c \cdot \bar{a}^T + b_c) \quad (17)$$

其中, \tanh 是非线性激活函数, W_c 和 b_c 表示可学习参数, \bar{a}^T 是向量 \bar{a} 的转置。然后学习上下文的最终语义表示,其包含了判断情感极性的信息,定义如下:

$$f_c = \sum_{i=1}^n \alpha_i s_i^c \quad (18)$$

同样地,以上下文的初始表示作为输入,利用 MIA 选择方面术语中对判断情感极性比较重要的信息,给定方面术语的语义表示 $S^a = \{s_1^a, s_2^a, \dots, s_i^a, \dots, s_n^a\}$ 和上下文的初始整体表示 \bar{c} 生成注意力向量 β , 定义如下:

$$\beta_i = \frac{\exp(\Phi(s_i^a, \bar{c}))}{\sum_{j=1}^k \exp(\Phi(s_j^a, \bar{c}))} \quad (19)$$

在计算注意力权重的基础上,学习方面术语的最终语义表示,其包含了判断情感极性的信息,定义如下:

$$f_a = \sum_{i=1}^k \beta_i s_i^a \quad (20)$$

最后将上下文和方面术语的最终语义进行拼接,得到 $r = [f_c, f_a]$ 。

2.6 标签分配

在方面术语抽取阶段,生成当前单词的标签时

基于之前单词的标签。经过位置嵌入层后,得到最终的句子特征表示 $S = \{s_1, s_2, \dots, s_t, \dots, s_n\}$ 。将其输入一个全连接网络层,并利用 softmax 进行标签分类。由此,方面术语标签 y_t^a 的条件概率定义如下:

$$P(y_t^a | \mathbf{y}_{[1:t-1]}^a, \mathbf{c}_t) = \text{softmax}(W_o s_t + b_o) \quad (21)$$

其中, $\mathbf{y}_{[1:t-1]}^a = [y_1^a, y_2^a, \dots, y_{t-1}^a]$, W_o 和 b_o 分别表示权重和偏置项, s_t 表示第 t 个单词标签的隐藏状态向量。假定 l_t 代表第 t 个单词的真实标签,使用 Adam 算法优化方面术语抽取任务,通过最小化负对数似然函数对该任务进行训练:

$$\text{Loss}(a) = -\frac{1}{n} \sum_{t=1}^n l_t \ln(P(y_t^a | \mathbf{y}_{[1:t-1]}^a, \mathbf{c}_t)) \quad (22)$$

在方面情感分类阶段,经过多交互注意力层之后,得到最终表示 r , 将其输入全连接层,再通过 softmax 预测方面术语的类别,定义如下:

$$P(y_i^p) = \text{softmax}(W_p r + b_p) \quad (23)$$

其中, W_p 和 b_p 分别表示可学习的参数, $P \in \mathbb{R}^{d_c}$ 是极性决策空间的概率分布, d_c 表示情感标签的维数。采用 Adam 算法和 l_2 正则化优化情感分类,使用交叉熵损失函数对该任务进行训练:

$$\text{Loss}(p) = -\sum_{i=1}^{d_c} \ln P(y_i^p) + \lambda \|\theta\|_2 \quad (24)$$

其中, θ 代表所有可训练参数, λ 是 l_2 正则化的系数。

2.7 联合训练

本文模型的性能依赖于每个任务损失之间的相对权重,手工调优是一个费时费力的过程。本文结合方面术语得分和情感分类得分的损失作为 PE-GCN 框架的训练目标:

$$\text{Loss} = \lambda_a \text{Loss}(a) + \lambda_p \text{Loss}(p) \quad (25)$$

其中, λ_a 和 λ_p 是平衡两个任务的训练参数。

3 实验

3.1 实验数据

为了验证 PE-GCN 模型的有效性,本文采用 SemEval-2014 公开数据集^[1]进行实验验证,该数据集包括 Restaurant 和 Laptop 两个领域的用户评论,数据样本都包含 Positive、Neutral 和 Negative 3 个标签(即 $d_c = 3$)。数据集统计结果如表 1 所示。

3.2 实验设置与评价指标

本文实验采用 300 维预先训练的 Glove 向量^[25]初始化单词嵌入,模型的所有权重参数均用 $U \sim (-0.1, 0.1)$ 的均匀分布进行初始化,GCN 的层数设置为 1,此时模型的性能表现最好。模型的超参数设置如表 2 所示。

表1 数据集统计

Table 1 Statistics of datasets

Dataset	Positive	Neutral	Negative
Laptop-train	994	464	870
Laptop-test	341	169	128
Restaurant-train	2 164	637	807
Restaurant-test	728	196	196

表2 超参数设置

Table 2 Setting of hyperparameters

超参数	参数值
词嵌入维度	300
隐藏状态向量维度	300
优化器 Optimizer	Adam
学习率 Learning rate	1E-3
L2正则化系数	1E-5
Batch size	16
训练迭代次数	50
Dropout rate	0.5

模型对方面术语抽取任务采用 $a-F1$ 值作为评价指标, 对方面情感分类任务采用 $s-F1$ 值和准确率 (accuracy, Acc) 作为评价指标。其定义如下:

$$Precision = \frac{TP}{TP + FP} \quad (26)$$

$$Recall = \frac{TP}{TP + FN} \quad (27)$$

$$a-F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (28)$$

$$Acc = \frac{\sum_{i=1}^d TP_i}{\sum_{i=1}^d (TP_i + FP_i + FN_i)} \quad (29)$$

$$s-F1 = \frac{1}{d} \sum_{c=1}^d (a-F1_c) \quad (30)$$

其中, TP 是正确预测的样本数量, FN 是将当前样本类别预测成其他类别的数量, FP 是将其他类别的样本预测为当前样本类别的数量。

3.3 对比实验

为了评估 PE-GCN 模型的有效性, 本文将其与单任务和任务的多组对比模型进行比较, 其中, 文献[1-2,4-5]是单任务方面术语抽取, 文献[3,6,9-10,19]是方面情感分类, 文献[11-14,18,21]是多任务。简介如下:

Baseline^[1], 对方面术语抽取任务, 首先采集训练句子中出现的一组方面, 然后找出测试句子是否有相同的方面集。对于情感分类任务, 以训练数据集中出现相同方面最频繁的为目标方面。

CNN+LP^[2], 一个将深度学习方法引入方面术语抽取任务中的应用。使用深度卷积神经网络标记句子中的方面词和非方面词, 并使用语言模式 (linguistic patterns, LP) 进一步提高该方法的性能。

PSO-EN (particle swarm optimization based ensemble learning method)^[4], 建立在不同特征集上的各种分类器的集合。应用粒子群算法选择最优的分类器集, 使用了 CRF、最大熵马尔可夫模型和支持向量机三个分类器。

HAST (history attention and selective transformation)^[5], 该模型包含意见总结和方面检测历史两个有用的线索。意见总结辅助标记进行方面术语预测, 方面检测历史信息利用坐标结构和标记模式增强方面预测。

LSTM (long short-term memory)^[3], 面向方面的长短时记忆模型。在生成句子表示时可以捕获上下文和方面词之间的关系, 以用于方面情感分类。

IAN (interactive attention networks)^[9], 利用两个交互注意力网络交互建模上下文和方面词之间的语义关系, 为判断方面术语的情感极性提供了足够的信息。

AOA (attention-over-attention)^[6], 对方面术语和上下文进行联合建模, 利用注意力过渡注意力捕获彼此之间的相互作用, 共同学习方面术语和上下文的表示, 自动关注句子中比较重要的部分。

ASGCN (aspect-specific graph convolutional network)^[10], 一种面向特定方面的情感分类框架。在句子依赖树上构建图卷积网络学习句子的语法信息和单词依赖关系。

BiGCN (bi-level interactive graph convolution network)^[19], 设计了双层交互式图卷积网络模型。充分利用层次句法图和词法图集成词共现信息和依赖类型信息。

CSAE (contextualized syntax-based aspect extraction)^[21], 提出了基于多通道上下文语义的方面抽取模型。将语法方面提炼为上下文特征, 以改进序列标记。

RNCRF (recursive neural conditional random fields)^[11], 提出了递归神经条件随机场模型。使用 RNN 对方面意见对的关系进行编码, 并提出一种基于最大似然和反向传播的联合优化方法同时学习 RNN 和 CRF 组件。

CMLA (coupled multi-layer attentions)^[12], 提出一种耦合多层注意力的深度学习模型。每一层由两个带有张量算子的注意力模块组成, 分别用于提取方面术语和意见术语, 且通过交互学习在彼此之间传播信息。

SpanMlt (span-based multi-task learning framework)^[13], 提出一种基于跨度的多任务神经网络框架。通过共享跨度表示, 在跨度边界和类标签的监督下提取方面/意见术语。同时, 通过计算跨度之间的对应关系识别成对关系。

MNN (multi-task neural network)^[14], 一种新颖的基于方面的情感分析多任务神经网络学习框架。用于联合处理方面术语抽取和情感预测子任务, 并利用注意力机制学习方面-情感关系的联合表示。

E2E (end-to-end architecture)^[18], 该模型以端到端机制处理多任务学习。采用 Bi-LSTM 模块学习句子的序列信息并预测句子中每个方面术语的边界。随后, CNN 模块利用已识别方面术语的局部特征进行情感分类。

3.4 实验结果与分析

表 3 给出了 PE-GCN 模型与单任务方面术语抽取相关模型的比较结果。可以看出, PE-GCN 模型在两个数据集上的实验性能均高于其他对比模型。对比考虑实体及其方面的 Baseline 模型, 在两个数据集上 $a-F1$ 值平均提升了 44.3%, 这表明本文提出的 PE-GCN 模型不需要费力的特征工程和外部语言资源就能有效运行。对比利用意见信息和方面检测历史的 HAST 模型, $a-F1$ 值平均提升了 3.23%, 这表明位置信息对模型性能影响较大。分析其原因, 认为这是由于 LSTM 学习出的隐藏状态中包含的多是短语级别的语义信息, 很难帮助模型准确识别出给定方面术语相关的意见词, 而本文引入的位置嵌入可以对神经网络学习出的隐藏状态进行细化微调, 使之能够更精确地捕获句子中方面词以及意见词特征。对比利用自注意力机制学习语法信息的 CSAE 模型, $a-F1$ 值平均提升了 3.54%, 这是由于在方面术语抽取任务中, 每个单词的相邻单词对预测其标签具有重要的影响, 距离远的词语则对其标签预测的影响较小, 而广泛使用的自注意力机制通常忽略了

表 3 PE-GCN 与方面术语抽取的对比

Task	Model	$a-F1$	
		Laptop	Restaurant
Aspect term extraction	Baseline	35.64	47.15
	CNN+LP	82.32	87.17
	PSO-EN	74.93	84.52
	HAST	79.52	85.41
	CSAE	77.65	86.65
Ours	PE-GCN	82.72	88.67

位置对于每个单词权重的影响, 因此, PE-GCN 模型根据单词之间的相对距离有规律地降低了上下文之间的权重, 有效提高了模型的性能。

表 4 给出了 PE-GCN 模型与单任务方面情感分类相关模型的比较结果。首先, PE-GCN 模型在两个数据集上的性能与 ASGCN 模型的性能相当, 但与 AOA 和 IAN 等其他模型相比, 在两个数据集上一致取得最优效果, 这表明 PE-GCN 模型的泛化能力更好。其次, 基于记忆网络的 LSTM 模型的性能表现较差, 不仅低于 PE-GCN 模型, 而且弱于 AOA 模型和 IAN 模型, 这说明引入注意力机制可以使模型在训练中高度关注特定方面术语的上下文信息, 从而有效改善 LSTM 模型在方面情感分类任务中的性能。此外, 在 Restaurant 数据集上, 由于评论语句大多语法结构较弱, 而 BiGCN 模型使用双重图卷积关注单词之间的共现信息, 对用户惯用表达比较敏感, 因此在句子语法结构较弱的情况下, BiGCN 模型对比 PE-GCN 模型取得了更好的效果。

表 4 PE-GCN 与方面情感分类的对比

Table 4 Comparison of PE-GCN and aspect sentiment classification 单位 %

Task	Model	Laptop		Restaurant	
		$s-F1$	Acc	$s-F1$	Acc
Aspect sentiment classification	LSTM	63.09	69.28	67.47	78.13
	AOA	67.52	72.62	70.42	79.97
	IAN	67.38	72.05	70.09	79.26
	ASGCN	71.05	75.55	72.02	80.77
Ours	BiGCN	71.84	74.59	73.48	81.97
	PE-GCN	72.64	75.92	71.61	80.66

表 5 给出了 PE-GCN 模型与多任务相关模型的比较结果。虽然 RNCRF、CMLA 和 SpanMlt 模型通过联合训练对方面术语和意见术语进行提取, 但仍然需要对提取的方面术语进行进一步的情感预测,

表 5 PE-GCN 与多任务的对比

Table 5 Comparison of PE-GCN and multi-task 单位 %

Task	Model	Laptop14		Restaurant14	
		$a-F1$	Acc	$a-F1$	Acc
Aspect term & Opinion term	RNCRF	78.42	—	84.93	—
	CMLA	77.80	—	85.29	—
	SpanMlt	77.87	—	85.24	—
Aspect term & Sentiment classification	MNN	79.91	72.85	85.84	79.14
	E2E	78.57	60.56	83.36	71.40
Ours	PE-GCN	82.72	75.92	88.67	80.66

容易导致错误传播和外部资源依赖。此外,对比利用自注意力机制和卷积神经网络提取特征的MNN模型和E2E模型,在两个数据集上 $a-F1$ 值平均提升了3.78%,准确率平均提升了7.30%。这表明利用卷积神经网络做特征提取构造分类模型,虽然能够在性能上接近或优于SpanMlt模型,但与PE-GCN模型相比性能优势并不显著,说明在多任务学习中采用位置嵌入和句法结构特征对分类准确性有着不容忽视的重要影响。

3.5 消融实验

为了进一步检查PE-GCN的每个组件对模型整体性能的影响,本文对该模型进行了消融研究,具体如下:

Bi-GRU+MIA: 模型由GRU层和交互注意力层组成,GRU层用来学习整个句子的语义特征表示,交互注意力层用来学习方面术语和上下文之间的语义关系。

Bi-GRU+PE+MIA: 模型在Bi-GRU层后新增了一个位置嵌入层,用来捕获抽取方面术语时位置信息对每个单词权重的影响。

Bi-GRU+GCN+MIA: 模型在Bi-GRU层后新增了一个图卷积网络层,用来学习上下文之间的句法依赖关系。

PE-GCN: 本文完整的模型,考虑了单词之间的位置信息和句法依赖关系对模型整体性能的影响。

如表6所示,Bi-GRU+MIA模型在Laptop和Restaurant数据集上的 $a-F1$ 值和 Acc 均不及其他模型,这表明句子中单词之间的位置信息和上下文之间的句法依赖关系都是不可忽略的。观察Bi-GRU+GCN+MIA模型,其性能整体上优于Bi-GRU+MIA模型,但仍不及PE-GCN模型,这表明虽然句法依赖关系有助于模型的整体建模,但仍不能忽略单词之间

表6 消融实验结果

Table 6 Ablation study results 单位: %

Model	Laptop		Restaurant	
	$a-F1$	Acc	$a-F1$	Acc
Bi-GRU+MIA	79.99	69.33	82.16	76.67
Bi-GRU+PE+MIA	81.57	74.15	85.31	78.33
Bi-GRU+GCN+MIA	80.32	73.66	85.41	78.33
PE-GCN	82.72	75.92	88.67	80.66

的位置信息。Bi-GRU+PE+MIA模型的性能在Laptop数据集上优于Bi-GRU+GCN+MIA,在Restaurant数据集上的性能和Bi-GRU+GCN+MIA模型相近,这表明相较于句法依赖关系,Laptop对位置信息更敏感。由此可见,位置信息和句法结构对于本文模型的构建同等重要。

3.6 GCN层数分析

实验将GCN层数分别设置为 $L=\{1,2,3,4,5,6,7,8,9,10\}$,在Laptop和Restaurant数据集上的 $a-F1$ 值和 Acc 分别如图3(a)和图3(b)所示。PE-GCN模型在GCN深度为1层时性能表现最佳,但随着GCN深度的增加,模型在两个公共数据集上的性能存在一定的波动,整体呈下降趋势。在GCN深度为10层时,相比最佳性能, $a-F1$ 值平均下降2.95%,准确率平均下降1.95%,由于引入更多参数,模型产生过拟合现象。

3.7 样例分析

从Laptop-test数据集中选取评论语句“The screen is nice, and the images comes very clear.”用作位置注意力可视化分析,结果如图4所示。其中,区域颜色越深,代表单词的注意力权重越大。

根据经验可知,方面术语大多是名词或名词短语,意见术语大多是形容词或动词。例如句子中的“screen”和“images”,“nice”和“very clear”。从图4中

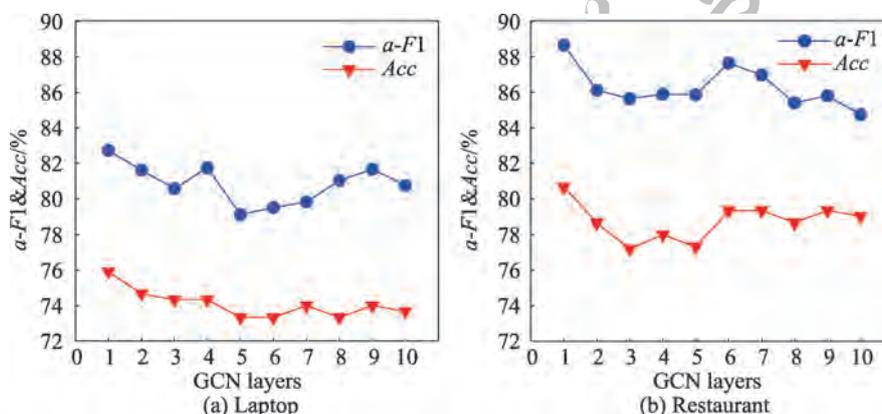


图3 GCN层数对模型性能的影响

Fig.3 Effect of the number of GCN layers on model performance

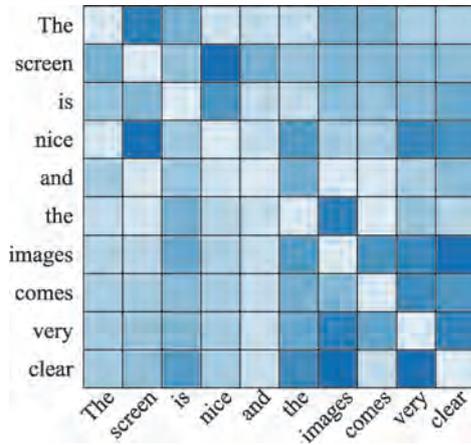


图4 上下文位置注意力权重

Fig.4 Context position attention weight

可以清晰地看出,这些词都受到了更多的关注。本文模型通过Bi-GRU捕捉句子的语义信息,再借助位置注意力为“screen”和“images”这两个方面术语分配了更高的权重。此外,由于意见术语与方面术语存在高度的语义相关性,例如,“screen”和“nice”相关度极高,“images”和“clear”相关度极高,因而它们也会获得更多的关注。

选取评论语句“I thought the price was great for specs.”用作特定方面语法距离可视化说明,结果如图5所示。区域颜色越深,表示该词的语法距离权重越大,对方面术语越重要。

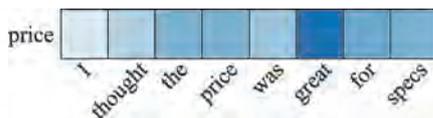


图5 语法距离权重

Fig.5 Grammatical distance weight

由图1可知,方面术语与意见词之间存在特定的句法依存关系,本文模型采用GCN生成包含句法信息的上下文表示,并通过距离编码为特定方面的上下文词分配合理的权重。从图5中可以看出,意见词“great”的区域颜色极深,其他区域颜色较浅,说明“great”与“price”在句法上相关度较大。这表明句法信息有利于模型的整体建模,图卷积网络能够有效学习与特定方面相关的上下文信息。

4 结束语

本文在多任务学习框架下提出一种基于位置嵌入和图卷积网络的联合模型,以端到端的方式整体解决 ABSA 任务,有效提高了复杂语句的分类准确

性。首先通过考虑单词之间的位置信息增强了方面术语的识别,同时通过构建句法依赖树和图卷积网络实现了句法依赖信息的融合;最后利用交互注意力建模方面术语和上下文之间的关系,实现了 ATE 和 ASC 两个任务之间的协调优化。实验结果表明,单词之间的位置信息和句法依赖关系均不可忽略,且交互建模方面术语和上下文是有利的,所提模型在 SemEval-2014 公开数据集上的准确率和 F1 值均得到了显著的提升。

参考文献:

- [1] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. SemEval-2014 task 4: aspect based sentiment analysis[C]//Proceedings of the 8th International Workshop on Semantic Evaluation, Dublin, Aug 23-24, 2014. Stroudsburg: ACL, 2014: 27-35.
- [2] PORIA S, CAMBRIA E, GELBUKH A. Aspect extraction for opinion mining with a deep convolutional neural network [J]. Knowledge-Based Systems, 2016, 108: 42-49.
- [3] TANG D Y, QIN B, FENG X C, et al. Effective LSTMs for target-dependent sentiment classification[C]//Proceedings of the 26th International Conference on Computational Linguistics, Osaka, Dec 11-17, 2016. Stroudsburg: ACL, 2016: 3298-3307.
- [4] AKHTAR M S, GUPTA D, EKBAL A, et al. Feature selection and ensemble construction: a two-step method for aspect-based sentiment analysis[J]. Knowledge-Based Systems, 2017, 125: 116-135.
- [5] LI X, BING L D, LI P J, et al. Aspect term extraction with history attention and selective transformation[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, Jul 13-19, 2018: 4194-4200.
- [6] HUANG B, OU Y L, CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks [C]//LNCS 10899: Proceedings of the 11th International Conference on Social, Cultural, and Behavioral Modeling, Washington, Jul 10-13, 2018. Cham: Springer, 2018: 179-206.
- [7] JAKOB N, GUREVYCH I. Extracting opinion targets in a single and cross-domain setting with conditional random fields[C]//Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, Oct 9-11, 2010. Stroudsburg: ACL, 2010: 1035-1045.
- [8] MA D H, LI S J, WU F Z, et al. Exploring sequence-to-sequence learning in aspect term extraction[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Jul 28-Aug 2, 2019. Stroudsburg: ACL, 2019: 3538-3547.
- [9] MA D H, LI S J, ZHANG X D, et al. Interactive attention networks for aspect-level sentiment classification[C]//Procee-

- dings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Aug 19-25, 2017: 4068-4074.
- [10] ZHANG C, LI Q C, SONG D W. Aspect-based sentiment classification with aspect-specific graph convolutional networks[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, Hong Kong, China, Nov 3-7, 2019. Stroudsburg: ACL, 2019: 4568-4578.
- [11] WANG W Y, PAN S J, DAHLMEIER D, et al. Recursive neural conditional random fields for aspect-based sentiment analysis[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Nov 1-4, 2016. Stroudsburg: ACL, 2016: 616-626.
- [12] WANG W Y, PAN S J, DAHLMEIER D, et al. Coupled multi-layer attentions for co-extraction of aspect and opinion terms [C]//Proceedings of the 31st AAAI Conference on Artificial Intelligence, San Francisco, Feb 4-9, 2017. Menlo Park: AAAI, 2017: 3316-3322.
- [13] ZHAO H, HUANG L T, ZHANG R, et al. SpanMIt: a span-based multi-task learning framework for pair-wise aspect and opinion terms extraction[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Jul 5-10, 2020. Stroudsburg: ACL, 2020: 3239-3248.
- [14] WANG F X, LAN M, Wang W T. Towards a one-stop solution to both aspect extraction and sentiment analysis tasks with neural multi-task learning[C]//Proceedings of the 2018 International Joint Conference on Neural Networks, Rio de Janeiro, Jul 8-13, 2018. Piscataway: IEEE, 2018: 1-8.
- [15] LI X, BING L D, LI P J, et al. A unified model for opinion target extraction and target sentiment prediction[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence, the 31st Innovative Applications of Artificial Intelligence Conference, the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, Honolulu, Jan 27-Feb 1, 2019. Menlo Park: AAAI, 2019: 6714-6721.
- [16] LUO H S, LI T R, LIU B, et al. DOER: dual cross-shared RNN for aspect term-polarity co-extraction[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Jul 28-Aug 2, 2019. Stroudsburg: ACL, 2019: 591-601.
- [17] HE R D, LEE W S, NG H T, et al. An interactive multi-task learning network for end-to-end aspect-based sentiment analysis[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Jul 28-Aug 2, 2019. Stroudsburg: ACL, 2019: 504-515.
- [18] AKHTAR M S, GARG T, EKBAL A. Multi-task learning for aspect term extraction and aspect sentiment classification[J]. Neurocomputing, 2020, 398: 247-256.
- [19] ZHANG M, QIAN T Y. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, Nov 16-20, 2020. Stroudsburg: ACL, 2020: 3540-3549.
- [20] BAI X F, LIU P B, ZHANG Y. Investigating typed syntactic dependencies for targeted sentiment classification using graph attention neural network[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021, 29: 503-514.
- [21] PHAN M H, OGUNBONA P O. Modelling context and syntactical features for aspect-based sentiment analysis[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Jul 5-10, 2020. Stroudsburg: ACL, 2020: 3211-3220.
- [22] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the Annual Conference on Neural Information Processing Systems 2017, Long Beach, Dec 4-9, 2017. Red Hook: Curran Associates, 2017: 5998-6008.
- [23] YOU J X, YING R, REN X, et al. GraphRNN: a deep generative model for graphs[J]. arXiv:1802.08773, 2018.
- [24] 苏锦钿, 欧阳志凡, 余珊珊. 基于依存树及距离注意力的句子属性情感分类[J]. 计算机研究与发展, 2019, 56(8): 1731-1745.
- SU J D, OUYANG Z F, YU S S. Aspect-level sentiment classification for sentences based on dependency tree and distance attention[J]. Journal of Computer Research and Development, 2019, 56(8): 1731-1745.
- [25] PENNINGTON J, SOCHER R, MANNING C. GloVe: global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Oct 25-29, 2014. Stroudsburg: ACL, 2014: 1532-1543.



孟甜甜 (1996—), 女, 甘肃庆阳人, 硕士研究生, CCF 学生会员, 主要研究方向为自然语言处理。

MENG Tiantian, born in 1996, M.S. candidate, student member of CCF. Her research interest is natural language processing.



韩虎 (1977—), 男, 甘肃兰州人, 博士, 教授, 主要研究方向为自然语言处理。

HAN Hu, born in 1977, Ph.D., professor. His research interest is natural language processing.



吴渊航 (1997—), 男, 山西运城人, 硕士研究生, CCF 学生会员, 主要研究方向为自然语言处理。

WU Yuanhang, born in 1997, M.S. candidate, student member of CCF. His research interest is natural language processing.