计算机科学
COMPUTER SCIENCE

基于交互注意力和图卷积网络的方面级情感分析

王娅丽, 张凡, 余增, 李天瑞

引用本文

王娅丽, 张凡, 余增, 李天瑞.基于交互注意力和图卷积网络的方面级情感分析[J]. 计算机科学, 2023, 50(4): 196-203.

WANG Yali, ZHANG Fan, YU Zeng, LI Tianrui. Aspect-level Sentiment Classification Based on Interactive Attention and Graph Convolutional Network [J]. Computer Science, 2023, 50(4): 196-203.

相似文章推荐（请使用火狐或 IE 浏览器查看文章）

Similar articles recommended (Please use Firefox or IE to view the article)

基于抽象语法树裁剪的智能合约漏洞检测研究

Smart Contract Vulnerability Detection Based on Abstract Syntax Tree Pruning

计算机科学, 2023, 50(4): 317-322. https://doi.org/10.11896/jsjkx.220300063

基于多邻接图与多头注意力机制的短期交通流量预测

Short-time Traffic Flow Forecasting Based on Multi-adjacent Graph and Multi-head Attention Mechanism

计算机科学, 2023, 50(4): 40-46. https://doi.org/10.11896/jsjkx.220200079

传播树结构结点及路径双注意力谣言检测模型

Dual-attention Network Model on Propagation Tree Structures for Rumor Detection

计算机科学, 2023, 50(4): 22-31. https://doi.org/10.11896/jsjkx.220200037

深度学习在健康医疗中的应用研究综述

Review of Deep Learning Applications in Healthcare

计算机科学, 2023, 50(4): 1-15. https://doi.org/10.11896/jsjkx.220600166

深度学习模型的后门攻击研究综述

Backdoor Attack on Deep Learning Models:A Survey

计算机科学, 2023, 50(3): 333-350. https://doi.org/10.11896/jsjkx.220600031

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220100105

基于交互注意力和图卷积网络的方面级情感分析

王娅丽[1]　张　凡[1,2]　余　增[1,2]　李天瑞[1,2]
1 西南交通大学计算机与人工智能学院　成都 611756
2 综合交通大数据应用技术国家工程实验室　成都 611756
(18398618821@163.com)

摘　要　方面级情感分析是细粒度情感分析中的一项关键任务,旨在预测一个句子中不同方面术语的情感倾向。针对目前结合图卷积网络的研究忽略方面术语本身的含义以及方面术语与上下文之间的交互的问题,文中提出了基于交互注意力和图卷积网络的模型(Interactive Attention Graph Convolution Network,IAGCN)。该模型首先结合 BiLSTM 和修正动态权重层对上下文进行建模,其次在句法依存树上使用图卷积网络对句法信息进行编码,然后利用交互注意力机制学习上下文和方面术语中的注意力,重构上下文和方面术语的表示,最后通过 softmax 层获取给定方面术语的情感极性。与基线模型相比,所提模型在 5 个数据集中的准确率和 F1 值分别提高了 0.56%～1.75%和 1.34%～4.04%。同时,将预训练模型 BERT 应用到此任务中,相比基于 GloVe 的 IAGCN 模型,其准确率和 F1 值分别提高了 1.47%～3.95%和 2.59%～7.55%,模型效果有了进一步的提升。

关键词:方面级情感分析;深度学习;图卷积网络;交互注意力机制;BERT
中图法分类号　TP181

Aspect-level Sentiment Classification Based on Interactive Attention and Graph Convolutional Network

WANG Yali[1],ZHANG Fan[1,2],YU Zeng[1,2] and LI Tianrui[1,2]
1 School of Computing and Artificial Intelligence,Southwest Jiaotong University,Chengdu 611756,China
2 National Engineering Laboratory of Integrated Transportation Big Data Application Technology,Chengdu 611756,China

Abstract　Aspect-level sentiment analysis is a key task in fine-grained sentiment analysis,which aims to predict the sentiment tendency of different aspect terms in a sentence. In view of the fact that the current research combined with graph convolution network ignores the meaning of aspect terms themselves and the interaction between aspect terms and context,an interactive attention graph convolutional network model is proposed,named interactive attention graph convolution network(IAGCN). It firstly combines BiLSTM and modified dynamic weights to model context. Secondly,the syntactic information is encoded by exploiting graph convolutional network on syntactic dependency tree. Then,the attention among context and aspect terms is investigated through interactive attention mechanism and the representation of context and aspect term is reconstructed. Finally,the sentiment polarity of a given aspect term is obtained through a softmax layer. Compared with the baseline models,the accuracy rate and F1 score of the proposed model improves by 0.56%～1.75% and 1.34%～4.04% on 5 datasets,respectively. At the same time,the pre-training model BERT is applied to this task. Compared with the IAGCN based on GloVe model,its accuracy rate and F1 score increases by 1.47%～3.95% and 2.59%～7.55%,respectively. Thus,the model effect has been further improved.

Keywords　Aspect-level sentiment analysis,Deep learning,Graph convolutional network,Interactive attention mechanism,BERT

1　引言

随着互联网时代的到来,各种网络平台层出不穷,越来越多的人在平台上发表自己的评价和建议,就此产生了海量的评价数据。对这类数据进行情感分析处理,可以为用户提供建议,同时为各平台提供决策支持。

文本情感分析[1]以往主要针对篇章和句子级别,将文章或句子看作一个信息单元进行整体情感分析,但针对一篇文章或单个句子中存在多种情感极性的情况,其情感分析效果并不好。因此,目前许多学者将研究内容从文本情感分析转向方面级情感分析。方面级情感分析[2]旨在识别出给定句子中给定方面的情感极性。例如,在餐厅评论“The environ-

到稿日期:2022-01-11　返修日期:2022-09-01
基金项目:国家自然科学基金(61773324);四川省重点研发项目(2020YFG0035)
This work was supported by the National Natural Science Foundation of China(61773324) and Sichuan Key R&D Project(2020YFG0035).
通信作者:李天瑞(trli@swjtu.edu.cn)

ment is great but the service attitude is poor”中，“environment”和“service attitude”这两个方面术语对应的情感极性分别是积极和消极。注意力机制[3]能够在模型训练的过程中高度关注指定的目标，已经被广泛应用到基于深度学习的自然语言处理的各项子任务中。在方面级情感分析领域，结合注意力机制的模型[4-5]可以关注给定方面术语的重要特征信息，增强句子中各单词与方面术语之间的联系。然而，注意力机制容易忽略句子的句法结构，当给定的方面术语与情感描述词相隔较远时，注意力机制往往识别不到正确的情感描述词。而句法依存树可以表达句子中词与词之间的相互依存关系，也可被视作图结构[6]，可以利用图卷积网络（Graph Convolution Network，GCN）[7]提取依存信息。因此，本文提出了一种结合交互注意力机制和图卷积网络的模型 IAGCN。该模型通过图卷积网络来捕获句子中单词之间的依存关系，较好地提取了结合上下文句法信息的方面术语，利用修正动态权重对句法位置的距离进行修正，以提取更多的上下文语义信息，结合交互注意力机制学习上下文和方面术语中的注意力，捕捉上下文和方面术语之间的相关性，正确识别上下文中给定方面术语的情感极性。在 5 个标准的数据集上验证了此模型的有效性，本文的主要工作如下：

（1）为了解决模型错误地将与句法无关的词识别为情感描述词的问题，本文在句法依存树上利用 GCN 进行建模，结合句法依存信息和位置信息，有效地学习单词之间的关系，提取方面术语特征。句法依存树包含了上下文单词之间的词汇依赖，位置信息隐含了重要的信息，方面术语和单词之间的相对距离对目标词有重要的影响，将 GCN 应用到句法依存树上能够融合邻居节点的特征信息，结合位置信息能够提取上下文各单词的不同的语义特征。

（2）经过 BiLSTM 建模的上下文提取了丰富的语义信息，但上下文各单词与方面术语的句法距离不同会影响情感的识别。为了减轻上下文中不相关单词对情感识别的不利影响，本文增加修正动态权重层使模型更加关注在句法上与方面术语相关的上下文。利用句法依存树中不同单词与方面术语之间的句法距离，设置一定的距离阈值，添加修正动态权重层对上下文中各单词的初始权重进行修正，针对语义贫乏的单词进行特征信息弱化，有效地提取了上下文特征表示。

（3）由于简单的注意力机制不能提取到方面术语和上下文之间的交互关系，以及忽略了方面术语内部的重要性，本文利用交互注意力机制提取上下文和方面术语中的语义和情感特征。将交互注意力机制应用于由修正动态权重层建模提取的上下文和方面特定掩码层建模提取的方面术语特征中，交互学习两者的注意力，捕获重要特征信息。

（4）在 5 个标准数据集上进行实验，结果表明，IAGCN 模型相比基线模型具有较好的效果。同时，将预训练语言模型 BERT（Bidirectional Encoder Representation from Transformer）应用到 IAGCN 中，模型的效果取得了显著的提升。

2 相关工作

早期的方面级情感分析主要基于词典和规则。Hu 等[8]利用关联规则挖掘评论的产品特性和意见。Ding 等[9]利用基于词典的方法，能够同时针对显示和隐式的方面术语进行情感极性分类。然而，这些方法人工设计的依赖度高，比较耗费时间和人力[10]。随着深度学习在自然语言中的广泛应用，Dong 等[11]首次将递归神经网络应用于方面级情感分析任务中，并取得了很好的效果。Tang 等[12]提出了基于目标依赖的长短期记忆网络（TD-LSTM）和基于目标关联的长短期记忆网络（TC-LSTM）两种模型，对上述工作进行了改进。基于注意力机制的方法可以使模型高度关注给定方面术语的特征信息，挖掘上下文与方面术语之间的关系，如 Wang 等[13]首次将注意力机制应用到方面级情感分析任务中，提出了结合 LSTM 和注意力机制的模型，将特定的方面术语嵌入到注意力权重的计算中。Zhang 等[14]认为，传统的长短期记忆网络和注意力相结合未考虑方面术语和上下文之间的联系，提出了有序神经元长短时记忆和自注意力机制相结合的方面情感分析模型，用于获取单词之间的依赖关系。随着记忆网络在方面级情感分析应用中的有效性得到验证，Tang 等[15]将记忆网络和注意力机制相结合，在保留上下文信息的同时也能捕捉到方面术语与上下文的关联性。Ma 等[16]认为，单注意力模型不足以捕捉到上下文和目标之间的联系，设计了交互注意力网络，使上下文和目标词之间获得对应的注意力权重。Song 等[17]避免了复杂的递归神经网络，采用多个多头注意力机制提取上下文和方面术语之间的隐含状态和语义交互。

近年来，卷积神经网络（Convolution Neural Network，CNN）在计算机视觉领域迅速发展。Fan 等[18]提出了一种卷积记忆网络，将 CNN 和注意力机制相结合，同时捕捉到单词和多个单词的表达，但是基于 CNN 的模型存在只能将多单词特征感知为连续单词的局限，不足以确定多个不连续单词表达的情感极性。GCN 是一种以图结构为输入的深度学习框架，在自然语言处理领域中，将文本表示成图结构数据是一种重要思想，利用 GCN 处理句法依存树，可以融合邻居节点信息以及捕捉非连续、长距离语义单词的特征。Zhang 等[19]首次将 GCN 运用到方面级情感分析任务中，提出了特定方面术语图卷积网络（Aspect-specific Graph Convolutional Networks，ASGCN），其通过引用句法依存树，利用 GCN 挖掘多词关系和句法信息，取得了很好的效果。Sun 等[20]添加多层 GCN 来提取语义信息，仅聚合方面术语有效信息进行情感极性分类。Zhao 等[21]结合 GCN 提出了一个新的方面级情感分析模型，该模型在注意力机制上使用 GCN 来捕获一个句子中不同方面之间的情感依赖关系，以此正确判断情感极性。Qi 等[22]利用带权的矩阵来存储更完整的句法关系，以此区分方面术语和非方面术语在特定方面情感极性判断中的重要性。Zhou 等[23]提出了基于语法和知识的图卷积网络，将常识知识也作为图结构输入到 GCN 中，联合建模提取方面术语的特征表示。

但是，以往结合交互注意力机制的模型针对方面术语没有过多的处理，忽略了方面术语本身在上下文中的语义信息，结合图神经网络的模型更加注重方面术语的句法依存信息，忽略了位置信息和方面术语而单独建模，没有考虑上下文和方面术语之间的交互以及方面术语本身包含的语义信息。在 Zhang 等[19]的启发下，本文提出的 IAGCN 模型利用图卷积

网络和方面特征掩码提取结合上下文语义的方面术语，同时针对上下文单词之间不同的句法距离，利用修正动态权重自动分配权重，提取结合句法位置信息的上下文特征，最后利用交互注意力机制同时考虑上下文和方面术语中的重要特征信息，为正确识别给定方面术语的情感极性提供足够且有效的信息。

3 方面情感分析模型

假设一个包含 n 个单词的语句 $s^c=\{\boldsymbol{e}_1^c,\boldsymbol{e}_2^c,\cdots,\boldsymbol{e}_n^c\}$，其中含有由 m 个单词组成的方面术语。方面级情感分析的目的是识别出特定的方面术语在整个句子中的情感极性。

本文提出的 IAGCN 模型如图 1 所示，模型主要分为 3 个模块，具体包括：

(1)利用 BiLSTM 对上下文建模，获取丰富的语义信息，结合修正动态权重层对上下文中各单词权重进行修正，弱化在句法依存树上与方面术语距离较远且无关的单词。

(2)考虑到单词之间的依赖关系，在句法依存树上利用 GCN 建模，经过方面特定掩码层学习方面术语的特征表示。

(3)通过交互注意力模块学习上下文和方面术语之间的注意力权重，更新两者的表示，最后连接上下文和方面术语的最终表示送入到全连接层，采用 softmax 函数进行情感极性分类。

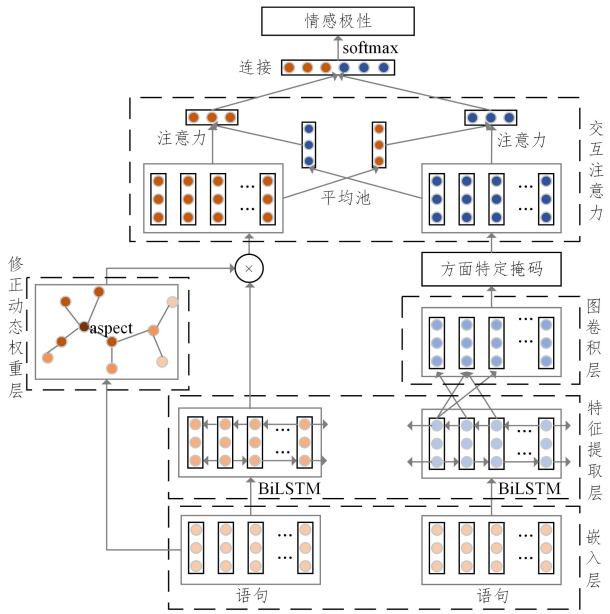

图 1 IAGCN 模型的结构

Fig. 1 Structure of the model for interactive attention graph convolution network

3.1 嵌入层和 BiLSTM 层

3.1.1 嵌入层

给定包含 n 个单词的语句 $s^c=\{\boldsymbol{e}_1^c,\boldsymbol{e}_2^c,\cdots,\boldsymbol{e}_n^c\}$ 和包含 m 个单词的方面术语 $\boldsymbol{s}^a=\{\boldsymbol{e}_{(t+1)}^a,\boldsymbol{e}_{(t+2)}^a,\cdots,\boldsymbol{e}_{(t+m)}^a\}$，其中 $\boldsymbol{s}^c$ 包含 $\boldsymbol{s}^a$。嵌入层将输入语句中的每个单词映射为低维实值向量，以此得到上下文词向量矩阵，本文在嵌入层分别应用了 GloVe 嵌入和 BERT 嵌入。以此，模型分别命名为 IAGCN 和 IAGCN-BERT。

(1) GloVe 嵌入。本文利用 GloVe[24] 嵌入矩阵 $\boldsymbol{E}\in R^{|V|\times d_e}$，将每一个单词嵌入到一个低维实值向量空间中，其中 $|V|$ 表示词汇量的大小，d_e 表示词嵌入的维度。

(2)BERT 嵌入。本文利用预先训练好的 BERT[25] 生成序列词向量，并将其作为 BERT 嵌入，为了 BERT 模型的训练和微调，模型将输入语句转换为“[CLS]+语句+[SEP]”。

3.1.2 BiLSTM 层

为了捕获上下文语义特征，针对每一个输入语句的词嵌入，模型构造了 BiLSTM 层以得到上下文相关隐藏状态向量 $\boldsymbol{H}^c=\{\boldsymbol{h}_1^c,\boldsymbol{h}_2^c,\cdots,\boldsymbol{h}_n^c\}$。其中，$\boldsymbol{h}_t^c\in R^{2d_h}$ 表示时间步 t 的隐藏状态向量，$\boldsymbol{d}_h$ 表示单向 LSTM 输出的隐藏状态向量的维数。

3.2 句法依存树上的图卷积网络

句法指句子中各个组成部分的相互关系，句法分析分为句法结构分析和依存关系分析。现有研究主要利用依存关系分析来获取词汇之间的依存关系。本文利用 spaCy 工具得到依存句法分析图，它是带标签的有向图。图 2 为一个句子的依存句法分析图，箭头从中心词指向从属词，标签表示从属的语法功能。

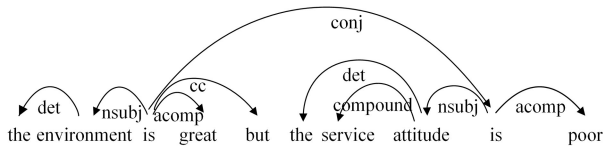

图 2 句法依存树

Fig. 2 Syntactic dependency tree

本文在给定句子的依存树上建立图卷积网络[7]，各单词作为图中的节点，通过句法依存树得到句子中与各单词相关的邻接矩阵 $\boldsymbol{A}_{ij}\in R^{n\times n}$，其中 n 是图的节点数。本文的依存树采用无向图，认为每个单词之间是相邻的，即设置 $\boldsymbol{A}_{ij}$ 的对角线值均为 1。GCN 根据每个节点的邻居节点相关信息进行建模，通过多层 GCN 的堆叠，每个节点最终的隐藏状态表示可以接收来自更远的邻居节点信息，如式(1)所示：

$$\boldsymbol{h}_i^l=\sigma\left(\sum_{j=1}^{k}\boldsymbol{A}_{ij}\boldsymbol{W}^l\boldsymbol{h}_j^{l-1}+\boldsymbol{b}^l\right) \tag{1}$$

其中，$\boldsymbol{A}_{ij}$ 表示邻接矩阵，l 表示 GCN 的层数，$\boldsymbol{h}_i^l$ 表示某节点在 l 层的输出，$\boldsymbol{W}^l$ 表示权重矩阵，$\boldsymbol{b}^l$ 表示偏置项，σ 表示非线性函数。

IAGCN 模型中 GCN 层在 BiLSTM 层的输出上继续建模，因此将 BiLSTM 层得到的隐藏状态向量 $\boldsymbol{H}^c$ 作为节点的初始状态 $\boldsymbol{H}^0$ 输入 GCN 中，并使用归一化因子的图卷积运算更新每个节点的表示，计算方法如式(2)、式(3)所示：

$$\tilde{\boldsymbol{h}}_i^l=\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{W}^l\boldsymbol{p}_j^{l-1} \tag{2}$$

$$\boldsymbol{h}_i^l=\mathrm{ReLU}(\tilde{\boldsymbol{h}}_i^l/(\boldsymbol{d}_i+1)+\boldsymbol{b}^l) \tag{3}$$

其中，$\boldsymbol{p}_j^{l-1}\in\boldsymbol{R}^{2d_h}$ 表示节点在 $l-1$ 层的输出，$\boldsymbol{W}^l$ 表示权重矩阵，$\boldsymbol{d}_i=\sum_{j=1}^{n}\boldsymbol{A}_{ij}$ 表示第 i 个节点的度数和，$\boldsymbol{b}^l$ 表示偏置项。

为了增强上下文中与方面术语相邻单词的重要程度，$\boldsymbol{h}_i^l$ 没有被直接当作下一层的输入，而是在每层输入前对其进行位置权重编码，如式(4)所示：

$$\boldsymbol{p}_i^l=F(\boldsymbol{h}_i^l) \tag{4}$$

其中，F 是位置权重(Position Weight，PW)编码函数，用于减少

依存句法分析过程中产生的噪声和偏差，如式(5)、式(6)所示：

$$F(\boldsymbol{h}_i^l)=\boldsymbol{q}_i\boldsymbol{h}_i^l \tag{5}$$

$$q_i=\begin{cases}1-\dfrac{t+1-i}{n}, & 1\leqslant i<t+1\\ 0, & t+1\leqslant i\leqslant t+m\\ 1-\dfrac{i-t-m}{n}, & t+m<i\leqslant n\end{cases} \tag{6}$$

其中，q_i 是第 i 个单词的位置权重，因此在 l 层的最终输出为：$\boldsymbol{H}^L=\{\boldsymbol{h}_1^l,\boldsymbol{h}_2^l,\boldsymbol{h}_{(t+1)}^l,\cdots\boldsymbol{h}_{(t+m)}^l,\cdots\boldsymbol{h}_n^l\}$，$\boldsymbol{h}_t^l\in\boldsymbol{R}^{2d_h}$。

3.3 修正动态权重和方面特定掩码

3.3.1 修正动态权重

BiLSTM 层的建模提供了上下文丰富的序列信息和语义信息。为了减轻上下文中与方面术语之间距离较远以及在句法上无关的单词的影响，模型采用修正动态权重(Modified Dynamic Weight，MDW)弱化上下文中语义相对较少的单词的信息。修正动态权重的初值由上下文中各单词与方面术语之间的句法相对距离(Syntactic Relative Distance，SRD)[26]决定。特别地，如果方面术语由多个单词组成，则输入词与多词方面术语之间的 SRD 以每个方面术语组成词与输入词之间的平均距离计算。例如，图 2 中方面术语“service attitude”与情感描述词“poor”之间的句法相对距离 SRD 的计算式如下：

SRD(service，poor)＝3.0，SRD(attitude，poor)＝2.0＝＞SRD(service attitude，poor)＝2.5

具体指“service”与情感描述词“poor”的句法相对距离等于 3.0，“attitude”与情感描述词的句法相对距离为 2.0。则整个方面术语“service attitude”与情感描述词“poor”之间的距离取平均值，等于 2.5。

修正动态权重考虑到固定的句法相对距离在识别情感极性时可能会出现的误差，采用了激活函数在训练过程中对距离进行修正。具体计算式如式(7)、式(9)所示：

$$\boldsymbol{v}_i^m=\begin{cases}1-\dfrac{\boldsymbol{SRD}_i-\boldsymbol{\alpha}}{N}\cdot\boldsymbol{I}, & \boldsymbol{SRD}_i>\alpha\\ \boldsymbol{I}, & \boldsymbol{SRD}_i\leqslant\alpha\end{cases} \tag{7}$$

$$\boldsymbol{V}=[\boldsymbol{v}_1^m,\boldsymbol{v}_2^m,\cdots,\boldsymbol{v}_n^m] \tag{8}$$

$$\boldsymbol{h}_i^m=\mathrm{ReLU}(\boldsymbol{W}^m\ \boldsymbol{v}_i^m+\boldsymbol{b}^m) \tag{9}$$

其中，$\mathbf{I}$ 表示全为 1 的向量；N 表示句子的长度；SRD_i 代表在句法依存树中上下文第 i 个单词与方面术语之间的最短距离，α 表示 SRD 阈值。

经过修正动态权重层得到修正句法距离为：

$\boldsymbol{H}^M=\{\boldsymbol{h}_1^m,\boldsymbol{h}_2^m,\cdots,\boldsymbol{h}_{t+1}^m,\cdots,\boldsymbol{h}_{t+m}^m,\cdots,\boldsymbol{h}_n^m\}$，$\boldsymbol{h}_t^m\in\boldsymbol{R}^{2d_h}$。为了弱化不相关单词的影响，将修正句法距离运用在 BiLSTM 的输出上，如式(10)所示：

$$\boldsymbol{H}_m^C=\boldsymbol{H}^C\odot\boldsymbol{H}^M \tag{10}$$

其中，$\boldsymbol{H}^C$ 表示 3.1 节中 BiLSTM 层的输出，$\boldsymbol{H}^M$ 表示修正句法距离。通过修正动态权重层可以获得结合句法位置信息的上下文特征表示。

3.3.2 方面特征掩码

在整个任务中，方面术语的特征表示对最终的情感极性识别有很大的影响，为了获得结合上下文语义的方面术语的特征，对式(3)中 GCN 的输出 h_i^l 添加了方面特定掩码层，以获取结合位置信息和句法信息的方面术语特征。方面特定掩码层旨在对上下文中非方面术语单词的隐藏状态向量进行掩码操作，保留方面术语的隐藏状态，以便于后续的方面术语与上下文之间的交互，具体如式(11)所示：

$$\boldsymbol{h}_a^L=0;1\leqslant a<t+1,t+m<a\leqslant n \tag{11}$$

其中，a 的取值范围指上下文中非方面术语的单词。方面特定掩码层的输出只保留了方面术语的特征，最终表示为 $\boldsymbol{H}_{\mathrm{mask}}^L=\{0,0,\cdots,\boldsymbol{h}_{t+1}^L,\cdots,\boldsymbol{h}_{t+m}^L,\cdots,0\}$。

经过修正动态权重层和方面特征掩码层，上下文特征表示结合了句法位置信息，方面术语特征表示同时结合了句法信息和位置信息，考虑了句法依赖和长距离的多词关系。

3.4 交互注意力模块

注意力机制能够使模型关注到句子中的重要部分，为了在上下文和方面术语之间进行交互建模，模型采用交互注意力机制关注上下文和方面术语中的重要部分。首先利用 3.3 节得到的上下文特征表示 $\boldsymbol{H}_m^C$ 和方面术语特征表示 $\boldsymbol{H}_{\mathrm{mask}}^L$ 的平均池化分别作为两者的初始表示，具体计算式如式(12)、式(13)所示：

$$\boldsymbol{c}_{\mathrm{avg}}=\sum_{i=1}^{n}\boldsymbol{h}_m^i/\boldsymbol{n} \tag{12}$$

$$\boldsymbol{a}_{\mathrm{avg}}=\sum_{i=1}^{n}\boldsymbol{h}_{\mathrm{mask}}^i/\boldsymbol{m} \tag{13}$$

其中，$\boldsymbol{h}_m^i$ 表示经过修正动态权重层输出的上下文特征表示，$\boldsymbol{n}$ 表示上下文的长度，$\boldsymbol{h}_{\mathrm{mask}}^i$ 表示经过方面特定掩码层输出的方面术语特征表示，$\boldsymbol{m}$ 表示给定方面术语的长度。

其次，为了获得方面术语对上下文的影响，将上下文和方面术语的初始特征 $\boldsymbol{c}_{\mathrm{avg}}$，$\boldsymbol{a}_{\mathrm{avg}}$ 作为输入，通过计算上下文各单词与方面术语的相关性，为上下文中各单词计算注意力权重 $\boldsymbol{\alpha}_i$，如式(14)和式(15)所示：

$$\gamma(\boldsymbol{h}_m^i,\boldsymbol{a}_{\mathrm{avg}})=\tanh(\boldsymbol{h}_m^i\cdot\boldsymbol{W}_a\cdot\boldsymbol{a}_{\mathrm{avg}}^{\mathrm{T}}+\boldsymbol{b}_a) \tag{14}$$

$$\boldsymbol{\alpha}_i=\frac{\exp(\boldsymbol{\gamma}(\boldsymbol{h}_m^i,\boldsymbol{a}_{\mathrm{avg}}))}{\sum_{j=1}^{n}\exp(\gamma(\boldsymbol{h}_m^i,\boldsymbol{a}_{\mathrm{avg}}))} \tag{15}$$

其中，tanh 是一个非线性函数，$\boldsymbol{\gamma}$ 是计算 $\boldsymbol{h}_m^i$ 在上下文中的重要性的数值函数；$\boldsymbol{W}_a$ 和 $\boldsymbol{b}_a$ 分别表示权重矩阵和偏置项；$\boldsymbol{a}_{\mathrm{avg}}^{\mathrm{T}}$ 表示方面术语初始特征 $\boldsymbol{a}_{\mathrm{avg}}$ 的转置矩阵。

同理，通过计算方面术语与上下文各单词的相关性，来为方面术语中的各单词计算注意力权重 β_i，具体计算式如式(16)、式(17)所示：

$$\boldsymbol{\gamma}(\boldsymbol{h}_{\mathrm{mask}}^i,\boldsymbol{c}_{\mathrm{avg}})=\tanh(\boldsymbol{h}_{\mathrm{mask}}^i\cdot\boldsymbol{W}_a\cdot\boldsymbol{c}_{\mathrm{avg}}^{\mathrm{T}}+\boldsymbol{b}_a) \tag{16}$$

$$\boldsymbol{\beta}_i=\frac{\exp(\boldsymbol{\gamma}(\boldsymbol{h}_{\mathrm{mask}}^i,\boldsymbol{c}_{\mathrm{avg}}))}{\sum_{j=1}^{n}\exp(\boldsymbol{\gamma}(\boldsymbol{h}_{\mathrm{mask}}^i,\boldsymbol{c}_{\mathrm{avg}}))} \tag{17}$$

其中，$\boldsymbol{c}_{\mathrm{avg}}^{\mathrm{T}}$ 表示上下文初始特征 $\boldsymbol{c}_{\mathrm{avg}}$ 的转置矩阵。

最后，利用两次得到的注意力权重来重构上下文和方面术语的特征表示，即 $\boldsymbol{c}_r$，$\boldsymbol{a}_r$，如式(18)和式(19)所示：

$$\boldsymbol{c}_r=\sum_{i=1}^{n}\boldsymbol{\alpha}_i\ \boldsymbol{h}_m^i \tag{18}$$

$$\boldsymbol{a}_r=\sum_{i=1}^{n}\boldsymbol{\beta}_i\boldsymbol{h}_{\mathrm{mask}}^i \tag{19}$$

3.5 输出层

将式(18)、式(19)重构后的上下文和方面术语表示连接，得到最终的输入向量表示 $\boldsymbol{r}=\{\boldsymbol{c}_r,\boldsymbol{a}_r\}$，通过 softmax 函数得到情感极性分类，如式(20)所示：

$$\boldsymbol{p}=\mathrm{softmax}(\boldsymbol{W}_y\boldsymbol{r}+\boldsymbol{b}_y) \tag{20}$$

其中，$\boldsymbol{p}\in R^{d_y}$，$\boldsymbol{d}_y$ 与情感分类极性维数相同；$\boldsymbol{W}_y\in R^{d_y\times 4d_h}$ 表示学习的权重，$\boldsymbol{b}_y\in R^{d_y}$ 表示学习的偏置项。

3.6 模型训练

该模型采用交叉熵损失并使用 L2 正则化，其损失函数如式(21)所示，采用梯度下降算法进行训练。

$$Loss=-\sum_{(c,\hat{p})\in C}\log \boldsymbol{p}_{\hat{p}}+\lambda\|\Theta\|_2 \tag{21}$$

其中，C 表示训练数据集集合，c 表示训练数据，$\hat{p}$ 表示数据的真实标签，Θ 表示模型所有可训练参数，λ 表示 L2 正则化系数。

4 实验

4.1 实验数据

本文实验在 5 个标准数据集上进行训练，包括 Twitter[11]数据集、SemEval14[27]任务 4 中的 LAP14 和 REST14 数据集、SemEval15[28]任务 12 中的 REST15 数据集和 SemEval16[29]任务 5 中的 REST16 数据集。实验最终采用的在 5 个数据集上对训练集和测试集中情感标签的划分如表 1 所列。

表 1 数据集信息

Table 1 Dataset information

Dataset		#Pos	#Neu	#Neg
Twitter	Train	1 561	3 127	1 560
	Test	173	346	173
LAP14	Train	994	464	870
	Test	341	169	128
REST14	Train	2 164	637	807
	Test	728	196	196
REST15	Train	912	36	256
	Test	326	34	182
REST16	Train	1 240	69	439
	Test	469	30	117

4.2 参数设置

本次实验中，对于 IAGCN 模型，采用与对比模型 ASGCN[19]相同的参数：用 300 维预训练的 GloVe 向量来初始化词嵌入，隐藏状态向量设为 300，学习率设为 0.001，L2 正则化系数设为10^{-5}。对于 IAGCN-BERT 模型，嵌入维度设为 768，学习率设为 2×10^{-5}，正则化系数设为 0.01。在两个模型中，GCN 层数设为 2，根据数据集不同的特点，将 LAP14 和 REST14 的 SRD 阈值设为 3，将 REST15，REST16 和 Twitter 的 SRD 阈值设为 5，在 4.6.2 节中进行了对比实验的详细说明，情感极性分类维数设为 3。此外，以准确率和 F1 值作为评价指标对模型进行综合衡量，随机初始化 3 次最优运行结果的平均值，并将其作为最终的实验结果。

4.3 模型对比

为了验证本文模型 IAGCN 和 IAGCN-BERT 的有效性，在 5 个数据集上与以下方面级情感分析模型进行比较。

(1)LSTM(Long Short-Term Memory)[12]采用目标依赖和目标关联的 LSTM，将目标词与语境词相结合，使用最后一次隐藏状态向量预测情感极性。

(2)AOA(Attention over Attention)[30]借用机器翻译领域的 Attention-over-Attention 的思想，以联合的方式对各个方面和句子进行建模，共同学习方面术语和上下文的表示，并自动关注上下文中的重要部分。

(3)IAN(Interactive Attention Network)[16]对方面术语单独建模，结合交互式注意力机制对方面术语和上下文之间的关系建模，学习两者的表示。

(4)ASGCN[19]采取图卷积网络对上下文进行建模，以此利用句法信息和单词之间的相互依赖关系进行情感极性分类。

(5)TD-GAT(Target-Dependent Graph Attention Network)[31]显示地利用单词之间的依赖关系，采用多层图注意力网络聚合与方面术语相关的单词信息，计算相关节点的特征表示。

(6)IGCN(Interactive Gated Convolutional Network)[32]是一种新颖的交互式门控卷积网络，利用特定领域的词嵌入，学习方面术语和上下文之间的相互关系，预测方面术语的情感极性。

(7)SKGCN(Syntax and Knowledge via Graph Convolutional Network)[23]结合了句法依存关系和常识知识图，利用 GCN 进行合并建模，增强了句子对给定方面术语的表示。

(8)BERT-SPC[17]将输入转化为“[CLS]+上下文+[SEP]+方面术语+[SEP]”，进行句子对分类。

(9)AEN-BERT(Attention Encoder Network)[17]结合预训练的 BERT 模型，采用基于注意力的编码器对语境和目标之间进行建模，并从词的嵌入中挖掘出丰富的交互式语义信息。

(10)AEGCN-BERT(Attentional Encoding Graph Convolution Networks)[33]采用预训练模型 BERT 作为嵌入层，结合 BiLSTM 和多头自注意力机制建模，并利用图卷积网络对句法依存树进行编码以获取丰富的语义信息。

4.4 实验结果与分析

本文在 4.1 节中的 5 个数据集上进行 3 分类实验，以准确率和 F1 值作为评价指标，实验结果如表 2 所列。从表中可以看出，与基线模型对比，本文提出的 IAGCN 模型在 5 个数据集上的准确率和 F1 值都有一定程度的提升。

从实验结果可以发现，IAGCN 模型的效果高于使用 LSTM 和注意力机制进行建模的一些模型(如 LSTM，AOA，IAN)，这表明注意力机制可能会忽略句法结构，针对上下文和方面术语相隔较远的情况，无法正确识别其情感描述词。使用 GCN 和 GAT 的一些模型(如 ASGCN，TD-GAT，IGCN，SKGCN 等)，使用图卷积网络或图注意力网络能够捕捉到上下文中长距离单词的依赖，但缺乏上下文和方面术语之间的交互联系以及上下文和方面术语各单词本身的语义信息。在大部分数据集上，IAGCN 模型的 F1 值都优于其他模型。IAGCN 充分利用了句法关系以及方面术语与上下文之间的交互关系，对最终的情感分类起到了关键性的作用。

具体地，在 5 个数据集上，本文提出的采用 BERT 作为嵌入层的 IAGCN-BERT 模型的整体性能优于 BERT-SPC 模型，这证明对于特定的下游任务设计网络是必要的。相比 AEN-BERT，AEGCN-BERT 和 SKGCN-BERT 模型，IAGCN-BERT 凭借交互注意力机制和句法依存树的优势，其整体性能

取得了提高。同样地，IAGCN-BERT 模型相比基于 GloVe 的 IAGCN 模型，其准确率和 F1 值分别提高了 1.47%～3.95%和 2.59%～7.55%，并且都优于表 2 中的非 BERT 模型，因此，与基于 GloVe 的模型相比，基于 BERT 的模型有足够的优势。

表 2　各模型准确率和 F1 值的对比

Table 2　Comparison of accuracies and F1 scores of different models

Model		Twitter		LAP14		REST14		REST15		REST16	
		Acc	*F1*	*Acc*	*F1*	Acc	F1	*Acc*	*F1*	*Acc*	*F1*
基线模型	LSTM	69.56	67.70	69.28	63.09	78.13	67.47	77.37	55.17	86.80	63.88
	AOA	72.30	70.20	72.62	67.52	79.97	70.42	78.17	57.02	87.50	66.21
	IAN	72.50	70.81	72.50	67.38	79.26	70.09	78.54	52.65	84.74	55.21
	ASGCN	72.15	70.40	75.55	71.05	80.77	72.02	**79.89**	61.89	**88.99**	67.48
	TD-GAT	72.20	70.45	75.63	70.74	81.32	71.12	80.38	60.50	87.71	67.87
	IGCN	—	—	75.24	—	81.34	—	—	—	—	—
	SKGCN	71.97	70.22	73.20	69.18	80.36	70.43	80.12	60.70	85.17	68.08
BERT 模型	BERT-SPC	73.55	72.14	78.99	75.03	84.46	76.98	—	—	—	—
	AEN-BERT	74.71	73.13	**79.93**	**76.31**	83.12	73.76	—	—	—	—
	AEGCN-BERT	75.04	73.68	78.73	74.22	82.58	73.40	82.71	69.00	89.61	73.93
	SKGCN-BERT	75.00	73.01	79.00	75.93	83.48	75.19	**83.20**	66.78	87.19	72.02
本文模型	IAGCN	**73.41**	**71.90**	**75.97**	**72.00**	**81.49**	**73.65**	79.34	**64.39**	88.31	**69.69**
	IAGCN-BERT	**75.14**	**73.76**	78.84	75.93	**83.48**	**75.96**	82.47	**69.25**	**89.61**	**74.04**

4.5　注意力可视化

为了更直观地感受到交互注意力的工作原理，本文将上下文和方面术语的权重进行可视化，以图 2 中的句子“the environment is great but the service attitude is poor”为例作交互注意力可视化说明，如图 3 所示，其中区域块颜色越深，表示注意力权重越大，能获得更高的关注。图 3(a)给出了两个上下文的注意力权重，图 3(b)给出了两个方面术语的注意力权重。在计算上下文各单词的权重时，当“environment”作为方面术语时，模型更关注“great”描述词，同时也发现模型的注意力权重更集中在“environment”的相关描述中，很少关注“service attitude”。同样地，当“service attitude”作为方面术语时，模型更关注“service attitude”的相关描述，这说明注意力机制能够关注到与方面术语相关的单词。

图 3　上下文和方面术语交互注意力权重

Fig. 3　Interactive attention weights for context and aspect terms

特别地，当方面术语由多个单词构成时，计算方面术语中各单词的权重，如“service attitude”，其中“attitude”是主词，而“service”用于修饰“attitude”，因此“attitude”对表达方面术语应该更重要。从图 3(b)中也可看出，模型的确更加关注“attitude”，这也表明本文模型能够识别多方面词中更重要的单词信息。因此，通过交互注意力不仅能够关注上下文和方面术语的交互，还能关注上下文和方面术语本身的含义，为方面术语情感极性的识别提供更多的信息。

4.6　消融实验

4.6.1　各组件对 IAGCN 整体性能的影响

为了研究 IAGCN 模型中各个组件对模型整体性能的影响，本文对 IAGCN 模型进行了消融实验，结果如表 3 所列。具体实验如下：

(1) IAGCN w/o mdw。去除修正动态权重，模型在 Twitter，Rest15 和 REST16 这 3 个数据集上的 F1 值均有所下降，而在 REST14 数据集上的准确率和 F1 值均有所提高，这表明对于句法关系不明显的数据集，运用修正动态权重不会有明显的作用。

(2)IAGCN w/o pw。去掉位置权重，模型在 5 个数据集上的效果均有所下降，这说明利用位置权重能帮助减少依存句法分析过程中产生的误差，学习更有效的信息。

(3)IAGCN w/o ian。利用单个注意力机制替代交互注意力机制，除了 REST16 数据集以外，模型在其他的数据集上的效果均下降，这说明方面术语和上下文之间的交互对情感分类起到了明显的作用。

(4)IAGCN w/o gcn。去除 GCN 后，模型在所有数据集上的效果都有明显的降低，这说明利用图卷积分析句法依存信息对情感识别有重要的作用。

表 3　消融实验结果对比

Table 3　Comparison of ablation experimental results

Model	Twitter		LAP14		REST14		REST15		REST16	
	Acc	*F1*	*Acc*	*F1*	*Acc*	*F1*	*Acc*	*F1*	*Acc*	*F1*
IAGCN	**73.41**	**71.90**	**75.97**	**72.00**	81.49	73.65	79.34	**64.39**	88.31	69.69
IAGCN w/o mdw	72.92	71.52	75.55	72.01	82.56	75.01	79.64	63.82	88.31	69.36
IAGCN w/o pw	72.16	70.48	75.08	70.91	80.33	71.47	78.78	61.68	87.45	66.46
IAGCN w/o ian	71.87	70.08	75.39	71.14	81.40	72.41	79.46	63.60	88.80	70.05
IAGCN w/o gcn	72.25	70.53	74.29	70.15	79.55	70.60	78.97	61.51	87.55	67.87

因此，在 IAGCN 中，交互注意力模块和结合位置权重的图卷积网络层对整个模型的贡献很大，修正动态权重层也有一定的作用。修正动态权重层通过激活函数和 SRD 距离修正上下文中的单词与方面术语之间的权重比例，图卷积网络和交互注意力模块的结合不仅充分利用句法关系捕捉到了上下文的依赖关系，而且能够捕捉方面术语和上下文之间的交互关系。然而，从实验结果中也能发现，在句法不敏感的 REST14 数据集上，模型 IAGCN 没有达到预期效果。

4.6.2 SRD 阈值 α 分析

修正动态权重层是 IAGCN 模型中的一个重要组件，由于动态权重涉及到 SRD 阈值 α，因此本文研究了阈值 α 对模型最终效果的影响。具体地，根据数据集的特点，α 从集合 {1,2,3,4,5,6} 中取值进行实验，在 5 个数据集上对比模型的 F1 值，实验结果如图 4 所示。

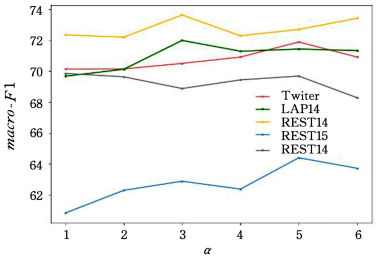

图 4 α 的取值对模型效果的影响

Fig. 4 Influence of the value of α

综合 5 个数据集上的 F1 值，当 $\alpha=3$ 时，LAP14 和 REST14 数据集取得了较好的效果；当 $\alpha=5$ 时，REST15，REST16 和 Twitter 这 3 个数据集取得了较好的效果。

结束语 在方面级情感分析任务中，本文提出了一种交互注意力和 GCN 相结合的模型 IAGCN，该模型同时考虑了上下文和方面术语之间的联系以及上下文句子之间的句法依存信息，并将预训练语言模型 BERT 应用于此任务，进一步提高了模型的效果。实验结果验证了该模型在方面级情感分析任务中的有效性。图注意网络(Graph Attention Network，GAT)在 GCN 的基础上加入了注意力机制，考虑了相邻节点的重要性。在后续的工作中，将考虑利用句法依存树生成边有权图，并结合 GAT 对节点和边同时进行考虑，以提取更多的语义信息。

参 考 文 献

[1] PANG B, LEE L. Opinion mining and sentiment analysis[M]. Now Publishers Inc, 2008.

[2] LIU B. Sentiment analysis and opinion mining [J]. Synthesis Lectures on Human Language Technologies, 2012, 5(1): 1-167.

[3] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv: 1409.0473, 2014.

[4] SUN X W, WANG L, WANG X, et al. Aspect-based sentiment analysis model based on dual-attention networks[J]. Journal of Computer Research and Development, 2019, 56(11): 2384-2395.

[5] MA Y, CHENG C L. Joint left and right attention mechanism for aspect-level text sentiment analysis[J]. Application Research of Computers, 2021, 38(6): 1753-1758.

[6] MARCHEGGIANI D, TITOV I. Encoding sentences with graph convolutional networks for semantic role labeling[C]// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1506-1515.

[7] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[J]. arXiv: 1609.02907, 2016.

[8] HU M, LIU B. Mining opinion features in customer reviews [C]// Proceedings of the 19th National Conference on Artificial intelligence. 2004: 755-760.

[9] DING X, LIU B, YU P S. A holistic lexicon-based approach to opinion mining[C]// Proceedings of the 2008 International Conference on Web Search and Data Mining. 2008: 231-240.

[10] LIU Q, LIANG B, XU J, et al. A deep hierarchical neural network model for aspect-based sentiment analysis[J]. Chinese Journal of Computers, 2018, 41(12): 3-18.

[11] DONG L, WEI F, TAN C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C]// Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. 2014: 49-54.

[12] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification[C]// Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers. 2016: 3298-3307.

[13] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 606-615.

[14] ZHANG Z L, LI L C, ZHU X Q, et al. Aspect sentiment analysis combining on-lstm and self-attention mechanism[J]. Journal of Chinese Computer Systems, 2020, 41(9): 1839-1844.

[15] TANG D, QIN B, LIU T. Aspect level sentiment classification with deep memory network[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 214-224.

[16] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification [C] // Proceedings of the 26th International Joint Conference on Artificial Intelligence. 2017: 4068-4074.

[17] SONG Y, WANG J, JIANG T, et al. Attentional encoder network for targeted sentiment classification [J]. arXiv: 1902.09314, 2019.

[18] FAN C, GAO Q, DU J, et al. Convolution-based memory network for aspect-based sentiment analysis[C]// Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018: 1161-1164.

[19] ZHANG C, LI Q, SONG D. 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 (EMNLP-IJCNLP). 2019: 4568-4578.

[20] SUN K, ZHANG R, MENSAH S, et al. Aspect-level sentiment analysis via convolution over dependency tree[C]// Proceedings

of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 5679-5688.

[21] ZHAO P, HOU L, WU O. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification[J]. arXiv: 1906. 04501, 2019.

[22] QI S Z, HUANG X Y, ZHU X F. Aspect-based sentiment analysis model based on weight enhancement[J/OL]. Journal of Chinese Computer Systems, 2021: 1-8. https://kns. cnki. net/kcms/detail/21. 1106. TP. 20210319. 1014. 012. html.

[23] ZHOU J, HUANG J X, HU Q V, et al. SK-GCN: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification [J/OL]. Knowledge-Based Systems, 2020, 205: 106292. https://doi. org/10. 1016/j. knosys. 2020. 106292.

[24] PENNINGTON J, SOCHER R, MANNINGG C D. Glove: Global vectors for word representation [C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014: 1532-1543.

[25] KENTON J D M W C, TOUTANOVA L K. BERT: Pre-training of deep bidirectional transformers for language understanding [C] // Proceedings of Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019: 4171-4186.

[26] 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. 2020: 3211-3220.

[27] 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. 2014: 27-35.

[28] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. Semeval-2015 task 12: Aspect based sentiment analysis [C] // Proceedings of the 9th International Workshop on Semantic Evaluation. 2015: 486-495.

[29] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. SemEval-2016 Task 5: Aspect based sentiment analysis[C] // Proceedings of the 10th International Workshop on Semantic Evaluation. 2016: 19-30.

[30] HUANG B, OU Y, CARLEY K M. Aspect levelsentiment classification with attention-over-attention neural networks [C] // Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Cham: Springer, 2018: 197-206.

[31] HUANG B, CARLEY K M. Syntax-aware aspect level sentiment classification with graph attention networks[C] // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 5469-5477.

[32] KUMAR A, NARAPAREDDY V T, SRIKANTHrikanth V A, et al. Aspect-based sentiment classification using interactive gated convolutional network [J]. IEEE Access, 2020, 8: 22445-22453.

[33] XIAO L, HU X, CHEN Y, et al. Targeted sentiment classification based on attentional encoding and graph convolutional networks[J/OL]. Applied Sciences, 2020, 10(3): 957https://doi. org/10. 3390/app10030957.

WANG Yali, born in 1997, postgraduate, is a member of China Computer Federation. Her main research interests include sentiment analysis and natural language processing.

LI Tianrui, born in 1969, Ph.D, professor, Ph.D supervisor, is a distinguished member of China Computer Federation. His main research interests include big data intelligence, rough sets, granular computing.

(责任编辑:喻藜)