

基于虚拟依存关系与知识增强的方面级情感分析

孔 博¹, 韩 虎¹, 陈景景², 白 雪¹, 邓 飞¹

(1. 兰州交通大学 电子与信息工程学院, 兰州 730070; 2. 兰州交通大学 数理学院, 兰州 730070)

摘 要: 借助句法依赖信息和外部知识的图神经网络近年来成为方面级情感分析领域的一个研究热点, 但是现有研究存在语法信息提取不充分和利用不合理等问题, 同时未考虑增强文本方面词与意见词等关键节点的背景知识。此外, 基于注意力机制的方法没有建立方面词与上下文词的语法信息交互, 导致方面词错误地关注到与其语法无关的上下文词信息。提出一种基于虚拟依存关系与双知识增强的多交互图卷积网络模型。对方面词内每个单词构建依赖树, 依据虚拟依存关系进行加权构造虚拟依存图, 依据外部情感知识构造情感依存图, 使用双通道图卷积神经网络处理虚拟依存图与情感依存图并进行融合, 对融合后特定方面的特征表示进行语义和语法双交互。利用概念知识增强特定方面后的特征表示与上下文表示并进行知识注意力交互, 对多种不同的增强表示进行融合从而实现不同表示间的共享与互补。实验结果表明, 相较于经典的图卷积网络模型 ASGCN, 该模型在 Rest15 和 Rest16 数据集上的 F1 值分别提升 4.71 和 8.57 个百分点, 具有较好的情感分类性能。

关键词: 方面级情感分析; 虚拟依存关系; 知识增强; 图卷积网络; 情感知识; 概念知识

开放科学(资源服务)标志码(OSID):



中文引用格式: 孔博, 韩虎, 陈景景, 等. 基于虚拟依存关系与知识增强的方面级情感分析[J]. 计算机工程, 2023, 49(10): 53-63.

英文引用格式: KONG B, HAN H, CHEN J J, et al. Aspect-based sentiment analysis through virtual dependency and knowledge enhancement[J]. Computer Engineering, 2023, 49(10): 53-63.

Aspect-Based Sentiment Analysis Through Virtual Dependency and Knowledge Enhancement

KONG Bo¹, HAN Hu¹, CHEN Jingjing², BAI Xue¹, DENG Fei¹

(1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2. School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou 730070, China)

[Abstract] Graph neural networks incorporating syntactic dependency information and external knowledge have recently gained significant attention in Aspect-Based Sentiment Analysis (ABSA). However, existing research exhibits limitations, such as inadequate extraction of grammatical information and suboptimal utilization. Moreover, the background knowledge pertaining to key nodes, such as context words and opinion words, has not been completely considered. Additionally, the attention mechanism-based method fails to establish a grammar information interaction between aspect words and context words, making aspect words mistakenly focus on irrelevant contextual information. A multi-interaction Graph Convolutional Network (GCN) model is proposed based on virtual dependencies and double knowledge enhancement. The model constructs a dependency tree for each word within an aspect word, generates a virtual dependency graph based on virtual relations, constructs an emotional dependency graph using external affective knowledge, and employs a dual channel graph convolutional neural network to process and fuse the virtual and affective dependency graphs. This approach enables semantic and grammatical double interaction on the feature representation of the fused specific aspect. Simultaneously, the feature and context representations of specific aspects, enhanced by conceptual knowledge, are used for knowledge-attention interaction. Finally, multiple enhanced representations are fused to facilitate sharing and complementation among different representations. Experimental results demonstrate that the proposed model achieves superior emotional classification performance compared to ASGCN. Specifically, the F1 values of the proposed model on the Rest15 and Rest16 datasets increase by 4.71 and 8.57 percentage points, respectively.

[Key words] Aspect-Based Sentiment Analysis (ABSA); virtual dependency; knowledge enhancement; Graph Convolutional

基金项目: 国家自然科学基金(62166024)。

作者简介: 孔 博(1997—), 男, 硕士研究生, 主研方向为自然语言处理; 韩 虎, 教授; 陈景景、白 雪、邓 飞, 硕士研究生。

收稿日期: 2022-10-12 **修回日期:** 2022-12-04 **E-mail:** hanhu_lzjtu@mail.lzjtu.cn

Network (GCN); emotional knowledge; conceptual knowledge
DOI: 10.19678/j.issn.1000-3428.0065982

0 概述

方面级情感分析 (Aspect-Based Sentiment Analysis, ABSA) 是自然语言处理 (Natural Language Processing, NLP) 领域的一个重要研究方向。ABSA 的目标是识别评论文本中给定方面的情感极性 (积极、消极或中性), 如句子 “I love Windows 7 which is a vast improvement over Vista!”, 其中方面 “Windows 7” 是积极的, 而方面 “Vista” 是消极的。

传统基于机器学习的方法依赖于所构建特征工程的质量且成本较高^[1]。TANG 等^[2]在方面和句子文本的词嵌入上利用长短期记忆 (Long Short-Term Memory, LSTM) 网络来获取句子级的特征信息, 从而获得情感极性。MA 等^[3]在 LSTM 模型的基础上, 对方面词与上下文的隐藏特征进行注意力交互。在之后的研究中, 使用注意力机制的混合神经网络模型在文本分类任务中得到了广泛应用。使用传统的注意力机制在方面词较多的句子中往往会忽略词语的句法表示, 导致权重分配错误。例如句子 “This French food tastes very well, but the restaurant has poor service.”, 在判断方面词 “restaurant” 的情感极性时, 由于缺乏依赖关系的表示, 否定副词 “but” 很容易被注意力机制忽略而将权重分配给 “well”。结合句法依赖树, 引入单词之间的依存关系进行交互, 能够让方面词更准确地关注到其对应的情感意见词。

近年来, 图卷积网络 (Graph Convolutional Network, GCN)^[4] 由于对句子句法依赖树等非结构化数据的处理优势而被广泛应用于方面情感分析中。HUANG 等^[5]提出图注意力网络 (Graph Attention Network, GAT), 将注意力机制引入图卷积网络中, 建立单词之间的依赖关系。ZHANG 等^[6]通过句子依赖树中的依存关系来构建图卷积网络的邻接矩阵, 但是其没有体现依赖树中的语法距离特征。现有研究表明, 语法距离特征有助于对方面情感进行分析^[7]。但是通过语法距离获得固定的系数或渐进的系数来增强不同单词之间的依存关系^[8-9]时, 都是在原始依存关系上进行加权, 未考虑有依存关系的节点对特定方面的影响。依赖树中包含丰富的语法信息, 但是通常不以目标方面作为依赖树的中轴, 而 ABSA 的重点是目标方面而不是树根节点。例如句子 “The battery gets so hot it is scary.”, 通过句

法解析器得到句法依赖树, 句子中的各单词之间在不同的依存关系, 在实际应用中只需要考虑 “gets” 与 “hot” 对方面词 “battery” 的影响, 而无须过多考虑 “gets” 与 “hot” 以及 “hot” 与 “is” 等非目标词之间的关系。WANG 等^[10]提出一种面向方面的树结构, 通过重塑和修剪普通的依赖树来关注目标方面。如果一个方面词包含多个单词, 以往是将其看作整体并作为根节点来重塑依赖树, 然而这种做法忽略了方面词内部的依存关系和上下文词对方面词内不同单词的语法距离的差异, 构造的依存关系对模型判断情感极性的重要程度不同。

虽然基于深层神经网络模型能够很好地捕捉上下文词和方面词之间的语法和语义信息, 但是这样的语义信息不够丰富, 需要有效融合外部知识来增强评论文本的特征表示。因此, 一些学者利用外部知识等先验知识为模型提供监督信号, 增强文本的语义信息。外部情感知识通常被用作情感分析任务中增强情感特征表示的来源^[11]。LIANG 等^[12]将情感知识融入到图网络中, 为每个语句建立一个特定于方面的情感增强依存关系图, 以捕获上下文词与方面词之间的情感依赖关系。CHEN 等^[13]通过引入知识图谱来解决不同语境下的 “一词多义” 问题。BIAN 等^[14]使用多头注意力机制并融合从外部知识库中提取的概念知识来增强上下文词和方面词的特征表示。

本文提出一种基于虚拟依存关系与双知识增强的多交互图卷积网络 (Virtual Dependency and Two Knowledge Enhancement based Multi-Interaction Graph Convolutional Network, VKI-GCN) 模型, 并将其应用于方面级情感分析任务。本文的主要工作如下:

1) 提出一种面向方面的树结构, 并采用上下文到方面词的虚拟依存关系来丰富依赖树, 通过语法距离对虚拟依存关系赋予权重, 构建以方面为中心的邻接矩阵。

2) 提出引入外部情感知识与虚拟依存关系增强的图卷积网络, 采用双通道融合的方式来获取方面特征, 设计语义交互和语法交互模块, 并结合概念知识增强的方面特征表示, 充分融合语义语法、情感知识与概念知识。

3) 在 5 个公开数据集上进行实验, 测试本文模

型在不同数据集上的准确率(Accuracy)和宏 F1 (Macro F1)值,以验证该模型的有效性。

1 相关工作

1.1 方面级情感分析

方面级情感分析是情感分析中的一项细粒度分类任务,其目标是识别句子中一个特定方面的情感极性。如今大多数方面级情感分析的研究都是基于神经网络而展开的。KIM^[15]将卷积神经网络(Convolutional Neural Network, CNN)用于情感分类任务,将预训练的词嵌入作为输入,使用CNN网络层来获得文本的情感特征,其结果验证了CNN对文本具有良好的语义解析能力。MA等^[3]提出交互式注意神经网络(Interactive Attention Network, IAN)模型,首次在句子和方面词之间使用交互式学习机制,即分别对方面词和句子进行注意力操作。FAN等^[16]提出一种细粒度的注意力机制,用于表征方面和上下文词之间的词级交互,缓解了粗粒度注意力机制中的信息损失问题,但是其忽略了句法信息。KIPF等^[4]将GCN应用到文本分类任务中,对依存图信息进行处理,从而起到语法融合的作用。ZHANG等^[6]使用句法依赖树获取句子中的语法信息,并利用图卷积网络对文本语义语法信息进行处理,从而判断方面词所对应的情感极性。

在最近的研究中,部分学者对语义信息和语法信息进行融合,进一步提高了模型的性能。ZHANG等^[17]融合层次句法信息和局部词汇语义信息,通过层级融合模块和门控机制分别处理图中功能不同的关系类型。王汝言等^[8]提出一种多交互图卷积网络模型,利用语法距离特征对图卷积网络的邻接矩阵进行加权以及语义语法交互。在已有研究中,通常使用0、1来表示节点之间的关系,或者以固定的系数来表示与方面词之间存在依存关系的节点权重,但是不同单词之间的依存关系往往对模型判断情感极性的贡献度存在差异。齐嵩喆等^[9]提出一种基于渐进增强与协同融合的图卷积神经网络模型,利用基于单词位置关系的渐进增强算法对依存图中的边进行加权,弥补了图卷积在整合长距离语义信息方面的不足。

虽然上述模型在情感分析任务中有较好的表现,但是大多只用单一的句子信息而未考虑外部情感知识的影响,同时也未考虑单词在多种语境下存在的“一词多义”问题,极大地影响了情感分析任务的预测效果。

1.2 外部知识

1.2.1 情感知识

在方面级情感分析任务中,情感知识常用来增强评论语句的情感特征^[11]。SenticNet是一个公开的情感分析资源,其利用情感值描述知识库中的每个概念^[18]。在SenticNet中,情感值分布在-1~1之间,分别代表强消极到强积极的概念,如表1所示。本文从SenticNet 6^[19]中提取出39 891个单词及情感值作为情感知识库来补充依赖图,从而增强句子的情感表示。

表1 SenticNet中的部分情感词汇
Table 1 Some emotional vocabulary in SenticNet

词汇	SenticNet 值
Great	0.857
Fantastic	0.870
Bad	-0.800
Balefully	-0.810
Hard	-0.059

1.2.2 概念知识

知识图谱的本质是语义网络的知识库,它可以为句子提供丰富的背景信息,提高模型的语义解析能力^[20]。CAO等^[21]借助领域知识同时采取实体词替换机制,在提升模型分类性能的同时也明显提升了模型的训练速度。HU等^[22]提出一种基于知识库构造的异质图神经网络,利用知识库中的知识增强文本语义表示,提高短文本分类效果。

微软亚洲研究院发布的大型概念知识图谱系统Microsoft Concept Graph,利用海量的搜索日志和网页数据学习构建常识知识^[20],借助其大量的实体信息和关系信息将文本词条映射到不同的语义概念,并通过(实例,概念,关系)的三元组来表示,在实际应用中为计算机提供有助于文本理解的先验知识。相对于传统的词向量,这种概念化是机器和人都可以理解的,已被广泛应用于现实任务中。本文利用该知识图谱系统进行方面词语义消歧,以解决“一词多义”的问题。

2 模型建立

给定一个长度为 n 的句子 $s = \{w_1, w_2, \dots, w_{a+1}, w_{a+2}, \dots, w_{a+m}, \dots, w_{n-1}, w_n\}$, 其中包括长度为 m 的方面 $a = \{w_{a+1}, w_{a+2}, \dots, w_{a+m}\}$, 即方面词 a 是句子 s 的一个子序列。本文模型框架如图1所示,由6个部分构成,分别是嵌入层和BiLSTM层、图卷积网络层、特征融合层、特定遮蔽层、多交互层以及输出层。

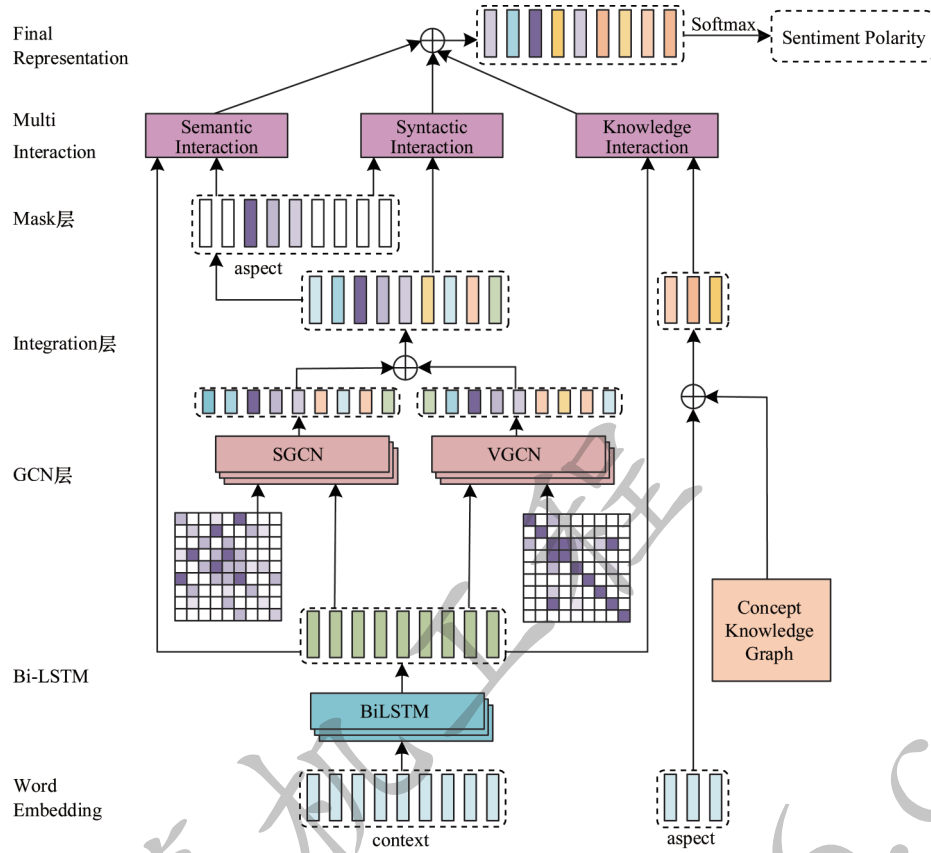


图1 VKI-GCN模型框架

Fig.1 VKI-GCN model framework

2.1 嵌入层与 Bi-LSTM 层

本文采用 GloVe 嵌入模型, 首先将每个单词映射到低维实值向量空间中得到低维的实数向量, 即词向量 $w' \in \mathbb{R}^n$; 然后将初始化的向量输入到 Bi-LSTM 中, 通过 Bi-LSTM 学习句子中单词的隐藏信息, 得到句子的隐藏状态 $H = \{h_1, h_2, \dots, h_n\}$ 。

2.2 图卷积层

2.2.1 基于虚拟依存关系增强的图卷积网络

本文提出一种 VGCN (Virtual Dependency Graph Convolutional Network) 模型, 构建面向方面的树结构, 通过重塑原始依赖树使其只聚焦于目标方面而丢弃与目标方面没有直接关联的依存句法关系。首先, 以方面词内部的每一个单词作为根节点构造依赖树, 并使用一个虚拟关系 $n: \text{con} (n: \text{connected})$ 表示方面词到每个节点的路径, 其中, n 表示 2 个节点之间的语法距离, 根据 n 来赋予不同的权重; 其次, 对由不同根节点词构造的依赖树进行融合。

图 2(a) 为原始依赖树, 图 2(b)、图 2(c) 是以方面词中的每个单词作为根节点所构造的依赖树, 图 2(d) 为融合后的依赖树。最后, 对融合后的依赖树构造邻接矩阵 $A_{i,j,v}$, 根据以往研究, $n_{\max}$ 一般取值

为 3。研究表明, 与目标方面距离越近的词, 其与目标方面间存在关联的可能性就越大, 引入基于相对距离的路径可以使树结构更加健壮^[23]。本文设计权重的思想是均分最大依赖距离, 距离越远, 权重越小, 逐级递减。 $A_{i,j,v}$ 计算公式如下:

$$A_{i,j,v} = \begin{cases} 1 - ((n-1)/n_{\max}), w_i, w_j \text{ 存在依存关系} \\ 1, i=j \\ 0, \text{其他} \end{cases} \quad (1)$$

将该邻接矩阵输入到 GCN 层, 根据其邻域节点的隐藏状态来更新第 l 层中每个节点的表示, 如式 (2)、式 (3) 所示, $\tilde{A}_{i,j,v}$ 是标准化后的邻接矩阵, 如式 (4) 所示。

$$h_i^l = \text{ReLU}(\tilde{A}_{i,j,v} g_i^{l-1} W^l + b^l) \quad (2)$$

$$g_i^{l-1} = F(h_i^{l-1}) \quad (3)$$

$$\tilde{A}_{i,j,v} = A_{i,j,v} / (E_{i,j,v} + 1) \quad (4)$$

其中: $E_{i,j,v} = \sum_{j=1}^n A_{i,j,v}$ 是单词节点的度; W^l 和 b^l 分别是权重矩阵和偏置项; ReLU 是非线性函数; g_i^{l-1} 是前一层卷积网络的输出结果; $F(\cdot)$ 是一种位置转换函数, 每一层的输出都通过位置编码进行加权。 L 层基于虚拟依存关系增强的图卷积网络的最终输出如式 (5) 所示:

$$h_v^L = \{h_1^L, h_2^L, \dots, h_n^L\} \quad (5)$$

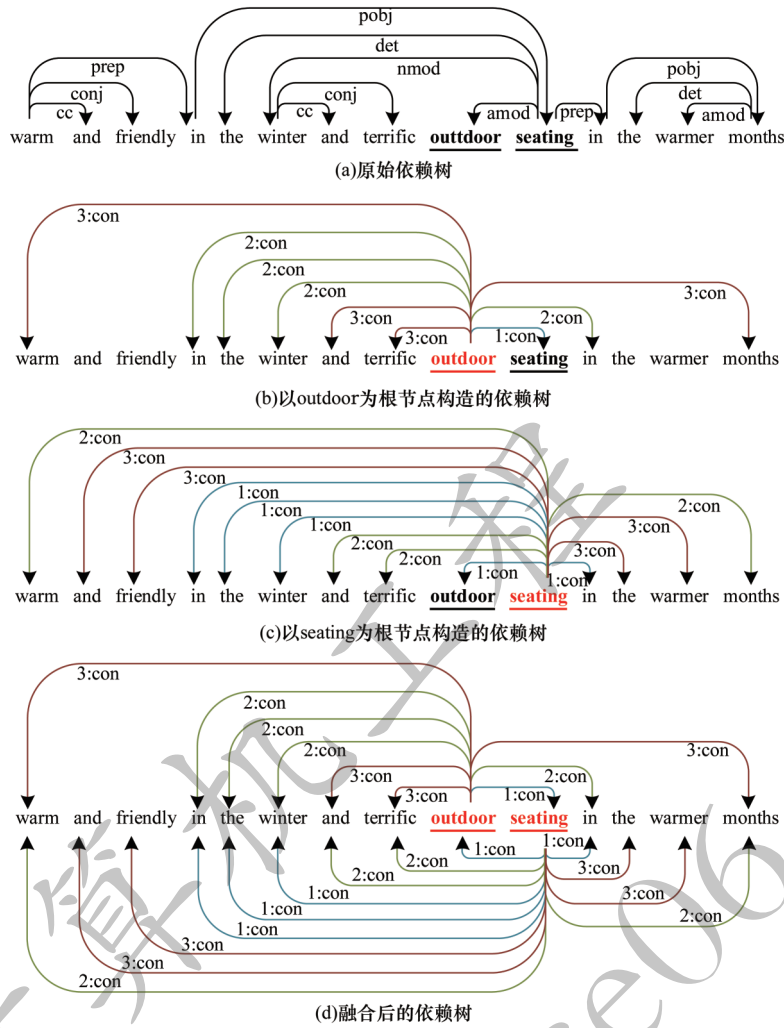


图2 普通依赖树以及面向方面的具有虚拟依存关系的依赖树

Fig.2 Ordinary dependency trees and aspect-oriented dependency trees with virtual dependency

2.2.2 基于情感知识增强的图卷积网络

情感知识的引入可以使模型学习到方面词与上下文词之间的情感先验知识。图1中的SGCN模块为融入情感知识SenticNet的图卷积网络。

为了利用句子中单词的依赖性,首先在依赖树上为每个输入句子构建邻接矩阵 $\mathbf{D} \in \mathbb{R}^{n \times n}$,其推导公式如式(6)所示:

$$\mathbf{D}_{i,j} = \begin{cases} 1, w_i, w_j \text{ 存在依赖关系或 } i=j \\ 0, \text{其他} \end{cases} \quad (6)$$

将2个节点的SenticNet情感分数进行叠加以增强邻接矩阵的表示,充分利用上下文和方面词之间的情感信息,如式(7)所示:

$$\mathbf{S}_{i,j} = \text{SenticNet}(w_i) + \text{SenticNet}(w_j) \quad (7)$$

其中: $\text{SenticNet}(w_i)$ 表示SenticNet中单词的情感得分,分布范围为 $[-1, 1]$,值取0表示单词是中性词或在SenticNet中不存在,取-1表示消极,取1表示积极。

此外,在构建情感知识的邻接矩阵时,还应该对特定方面给予重要关注,如式(8)所示:

$$\mathbf{T}_{i,j} = \begin{cases} 1, w_i \text{ 或 } w_j \text{ 是方面词} \\ 0, \text{其他} \end{cases} \quad (8)$$

由此得到情感知识增强后的邻接矩阵,如式(9)所示:

$$\mathbf{A}_{i,j} = \mathbf{D}_{i,j} \times (\mathbf{S}_{i,j} + \mathbf{T}_{i,j} + 1) \quad (9)$$

将该邻接矩阵输入到GCN层,根据其邻域节点的隐藏状态来更新第 l 层中每个节点的表示,如式(10)、式(11)所示:

$$\mathbf{h}_i^l = \text{ReLU}(\tilde{\mathbf{A}}_i \mathbf{g}_i^{l-1} \mathbf{W}^l + \mathbf{b}^l) \quad (10)$$

$$\mathbf{g}_i^{l-1} = F(\mathbf{h}_i^{l-1}) \quad (11)$$

$\tilde{\mathbf{A}}_i$ 是标准化后的邻接矩阵,如式(12)所示:

$$\tilde{\mathbf{A}}_{i,j} = \mathbf{A}_{i,j} / (\mathbf{E}_i + 1) \quad (12)$$

其中: $\mathbf{E}_i = \sum_{j=1}^n \mathbf{A}_{i,j}$ 是单词节点的度。

最后, L 层情感知识增强的图卷积网络的最终输出如式(13)所示:

$$\mathbf{h}_s^L = \{\mathbf{h}_1^L, \mathbf{h}_2^L, \dots, \mathbf{h}_n^L\} \quad (13)$$

2.3 融合层

2.3.1 双通道图卷积特征融合

本文采用双通道的形式,利用图卷积网络学习句法信息和情感知识。基于虚拟依存关系增强的图卷积网络与基于情感知识增强的图卷积网络在经过

L 层图卷积网络输出后进行拼接,得到更丰富的特征表示 $\tilde{H}$,如式(14)所示:

$$\tilde{H}=[h_v; h_s] \quad (14)$$

2.3.2 方面词概念知识融合

本文借助 Concept Graph 对方面词进行概念化,得到其概念集 $K=\{k_1, k_2, \dots, k_m\}$,将方面向量 $a=\{w_{a+1}, w_{a+2}, \dots, w_{a+m}\}$ 与其对应的概念向量 $K=\{k_1, k_2, \dots, k_m\}$ 进行拼接,得到方面的概念化向量 $a^K=[a; K]$ 。

2.4 遮蔽层

对融合层双通道图卷积拼接后的输出 $\tilde{H}$ 进行遮掩,能够屏蔽非方面词的隐藏状态向量,保留方面词向量。将方面词所对应的位置设置为1,非方面词所对应的位置设置为0,经过方面特定掩蔽层后的输出 $\tilde{H}_{\text{mask}}$ 如式(15)所示:

$$\tilde{H}_{\text{mask}}=\{0, \dots, \tilde{h}_{a+1}, \dots, \tilde{h}_{a+m}, \dots, 0\} \quad (15)$$

其中: $\tilde{H}_{\text{mask}}$ 是包含句子语法消息和语义信息的方面词隐藏特征。

2.5 多交互注意层

考虑到上下文词与融合概念知识后的方面词之间的关系,本文设计知识注意力交互机制,以提取融合概念知识后的特征消息。

1)知识交互。经过 Bi-LSTM 之后获取文本中隐含语义特征的上下文表示 H ,将其与融合概念知识后的方面词 a^K 进行交互,如式(16)~式(18)所示:

$$\beta_k = \sum_{i=1}^n H^T a_i^K \quad (16)$$

$$\alpha_k = \frac{\exp(\beta_k)}{\sum_{i=1}^n \exp(\beta_i)} \quad (17)$$

$$h_k = \sum_{j=1}^n \alpha_j^k H \quad (18)$$

为了同时关注方面词与上下文词之间的语义关系和语法关系,设计语义和语法2个分支的多交互注意力机制,同时提取语义特征信息和语法特征信息。

2)语义交互。经过 Bi-LSTM 之后获取文本中隐含语义特征的上下文表示 H ,将其与方面特征遮掩之后的方面词 $\tilde{H}_{\text{mask}}$ 进行交互。为了实现上下文与方面词的维度统一,将句子的隐藏状态进行复制拼接得到 $H_c=[H; H]$ 。语义交互如式(19)~式(21)所示:

$$\beta_{sc} = (H_c)^T \tilde{H}_{\text{mask}, i} \quad (19)$$

$$\alpha_{sc} = \frac{\exp(\beta_{sc})}{\sum_{i=1}^n \exp(\beta_{sc}^i)} \quad (20)$$

$$h_{sc} = \sum_{j=1}^n \alpha_{sc}^j H_c \quad (21)$$

3)语法交互。将双通道融合后隐含语法特征的上下文表示 $\tilde{H}$ 与方面特征遮掩之后的方面词 $\tilde{H}_{\text{mask}}$ 进行交互,如式(22)~式(24)所示:

$$\beta_{sy} = \sum_{i=1}^n \tilde{H}^T \tilde{H}_{\text{mask}, i} \quad (22)$$

$$\alpha_{sy} = \frac{\exp(\beta_{sy})}{\sum_{i=1}^n \exp(\beta_{sy}^i)} \quad (23)$$

$$h_{sy} = \sum_{j=1}^n \alpha_{sy}^j \tilde{H} \quad (24)$$

将多交互层得到的知识交互特征向量 h_k 、语义交互特征向量 h_{sc} 和语法交互特征向量 h_{sy} 相拼接,得到 $h_f=[h_k; h_{sc}; h_{sy}]$ 。

2.6 输出层及模型训练

将多交互注意层的输出作为全连接层的输入,通过 Softmax 函数输出最终的情感极性,如式(25)所示:

$$y = \text{Softmax}(W_o h_f + b_o) \quad (25)$$

其中: h_f 为多交互注意层的输出; W_o 为全连接层的权重项; b_o 为全连接层的偏置项。本文通过梯度下降算法训练模型来完成分类任务:

$$\zeta = - \sum_{i=1}^S \sum_{j=1}^C y_i^j \cdot \log_a \hat{y}_i^j + \lambda \|\Theta\| \quad (26)$$

其中: S 是训练样本数; C 是类别数; y_i^j 和 $\hat{y}_i^j$ 分别是训练集的真实标签和预测标签; Θ 代表所有的可训练参数; λ 是L2正则化系数。

3 实验验证

3.1 数据集

本文在5个公开数据集上进行实验,其中, Twitter数据集来自于 TWITTER^[24], Lap14、Rest14数据集来自于 SemEval-2014^[25], Rest15数据集来自于 SemEval-2015^[26], Rest16数据集来自于 SemEval-2016^[27]。各数据集的统计信息如表2所示。

表2 数据集的统计信息

Table 2 Statistical information of datasets

数据集	积极		中性		消极	
	训练集	测试集	训练集	测试集	训练集	测试集
Twitter	1 561	173	3 127	346	1 560	173
Lap14	994	341	464	169	870	128
Rest14	2 164	728	637	196	807	196
Rest15	912	326	36	34	256	182
Rest16	1 240	469	69	30	439	117

3.2 评价指标

模型性能采用准确率(简称 Acc)和宏 F1(简称 F1)值作为评价指标,2个指标的计算公式分别如下:

$$A_{\text{Acc}} = \frac{T}{N} \quad (27)$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \quad (28)$$

其中: T 表示正确预测的样本数量; N 表示样本总数。F1值是分类问题的衡量指标,是精确率与召回率的调和平均数; P 表示预测为正例的样本中预测为真正例的概率; R 表示真正例样本中预测为真正例的概率。

3.3 参数设置

在实验过程中, 本文利用预训练 GloVe 作为初始化词的嵌入向量, 超参数设置如表 3 所示。

表 3 实验参数设置
Table 3 Experimental parameters setting

参数	设置
词向量维度	300
隐藏层节点维度	300
学习率 Learning rate	0.002
训练迭代次数 epoch/次	100
批量训练样本数 Batch size/个	16
丢失率 Dropout rate	0.3
L_2 正则化系数	0.000 01

3.4 对比模型

将本文提出的 VKI-GCN 模型与以下方面级情感分析模型进行比较:

- 1) LSTM。文献[2]利用 LSTM 编码上下文信息, 将最后一层的输出作为文本特征表示输入到分类器中, 从而预测情感极性。
- 2) IAN。文献[3]提出的 IAN 模型利用注意力机制对上下文和目标进行交互学习, 得到最终的表示。
- 3) ASGCN。文献[6]提出一种基于图卷积的网络模型 ASGCN, 利用句法依存树获取句法信息图, 使用注意力机制对方面词与上下文语义信息进行交互。
- 4) KumaGCN。文献[28]提出一种门控机制来动态地结合单词依赖图和自我注意网络学习的潜在图中的信息, 用潜在的语义依赖来补充受监督的句法特征。
- 5) DGCN。文献[29]结合词语的句法结构和语法距离, 提出一种根据句法依存树获得的语法距离

权重, 结合句子的上下文语义信息、语法距离权重向量和邻接矩阵, 通过图卷积网络提取方面词的情感特征。

6) SK-GCN。文献[30]提出一种对句法依存树和常识知识进行编码的图卷积模型, 以丰富句子对特定方面的表示。

7) AEGCN。文献[31]以双通道的形式分别利用多头注意力和基于依赖树注意力改进的 GCN 来对文本表示进行编码, 并利用通道间的交互注意力进一步增强表示。

8) MIGCN。文献[8]提出一种多交互模型, 设计语义交互和语法交互过程, 分别对单词之间的语义信息和语法信息进行学习。

3.5 结果分析

本文在 Twitter、Lap14、Rest14、Rest15、Rest16 数据集上进行 8 组模型对比实验, 结果如表 4 所示, 最优结果加粗标注。从表 4 可以看出: 本文 VKI-GCN 模型与最新模型 MIGCN 相比, 在 5 个数据集上 Acc 值分别提升 1.55、0.68、0.72、1.29 和 0.76 个百分点, F1 值分别提升 0.83、0.81、0.03、2.39 和 4.08 个百分点, 验证了本文所提模型的优越性; 相比使用基础 GCN 的模型, 本文 VKI-GCN 模型性能均有所提升, 这是因为评论语句较短, 缺乏足够的语义信息, 而 VKI-GCN 模型可以有效融合方面词的概念知识和意见词的情感知识等先验知识来丰富语义表示, 尤其在 Twitter 数据集上, VKI-GCN 较 ASGCN 的 Acc 值提升 2.71 个百分点, F1 值提升 2.55 个百分点, 在 Rest15 和 Rest16 数据集上 VKI-GCN 较 ASGCN 的 F1 值分别提升 4.71 和 8.57 个百分点。由实验结果可看出, 外部知识对精确表示文本语义信息具有重要性, 本文模型可以通过外部知识同时增强方面词的语义信息和意见词的情感信息, 加强评论文本表示, 提高模型分类能力。

表 4 对比实验结果
Table 4 Results of comparative experiments

模型	Twitter		Lap14		Rest14		Rest15		Rest16	
	Acc	F_1	Acc	F_1	Acc	F_1	Acc	F_1	Acc	F_1
LSTM	69.56	67.70	69.28	63.09	78.13	67.47	77.37	55.17	86.80	63.88
IAN	72.50	70.81	72.05	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
KumaGCN	72.45	70.77	76.12	72.42	81.43	73.64	80.69	65.99	89.39	73.19
DGCN	72.87	71.21	74.63	71.03	81.24	72.84	—	—	—	—
SK-GCN	71.97	70.22	73.20	69.18	80.36	70.43	80.12	60.70	85.17	68.08
AEGCN	73.86	71.59	75.91	71.88	81.43	73.66	80.85	63.96	88.76	68.73
MIGCN	73.31	72.12	76.59	72.44	82.32	74.31	80.81	64.21	89.50	71.97
VKI-GCN	74.86	72.95	77.27	73.25	83.04	74.34	82.10	66.60	90.26	76.05

本文 VKI-GCN 模型虽然在 5 个数据集上相较于以往模型分类性能均有提升, 但在不同数据集上的提升效果存在一定差异, 在 Rest14 数据集上提升较

小, 观察该数据集可以发现, 其包含大量相对较长的评论语句, 而本文 VKI-GCN 模型是将方面词作为根节点构建依赖树, 并忽略语法距离大于等于 4 的虚

拟依存关系,使得本文模型在处理较长评论语句时容易将过远的意见词忽略掉,因此,在Rest14等数据集上无法发挥出模型的优势。

3.6 消融实验

为了明确 VKI-GCN 中各部分对模型性能的影响,本文设计消融实验,对 VGCN 与 ASGCN、VGCN 上是否融合情感知识和概念知识以及 VKI-GCN 上是否进

行语义交互和语法交互这3组对比实验进行可视化分析。

如表5所示,W/O SGCN+CK+Sy表示仅保留VGCN模块并进行语义交互,W/O SGCN表示仅去掉融入情感知识 SenticNet的图卷积网络分支,W/O CK表示仅去掉引入概念知识的分支,W/O Sy表示仅去掉语法交互,W/O Se表示仅去掉语义交互。从表5可以看出,各组件均对本文模型的性能有提升效果。

表5 消融实验结果

Table 5 Results of ablation experiment

%

模型	Twitter		Lap14		Rest14		Rest15		Rest16	
	Acc	F_1	Acc	F_1	Acc	F_1	Acc	F_1	Acc	F_1
W/O SGCN+CK+Sy	73.55	71.92	76.80	72.88	82.68	75.36	81.37	64.26	89.61	71.48
W/O SGCN	73.84	71.71	77.27	73.38	82.69	74.21	80.81	67.32	89.77	72.94
W/O CK	74.28	73.14	77.12	72.96	82.95	74.44	80.81	65.11	89.61	72.31
W/O Sy	73.12	71.03	75.71	72.31	82.32	73.72	81.37	63.54	89.12	69.40
W/O Se	72.11	69.95	76.96	72.75	80.63	72.43	79.70	64.64	88.47	66.64
VKI-GCN	74.86	72.95	77.27	73.25	83.04	74.34	82.10	66.60	90.26	76.05

3.6.1 VGCN 与 ASGCN 的对比实验

为了验证面向方面构建虚拟关系来增强依赖树的语法信息对于模型性能的提升效果,本文设置对比实验,对比模型包括基础模型 ASGCN 和本文所提 VGCN 模型,即将 ASGCN 的 0、1 矩阵替换为由新构建的依赖树所获得的邻接矩阵,其余参数条件和数据集保持不变,实验结果如图3、图4所示。

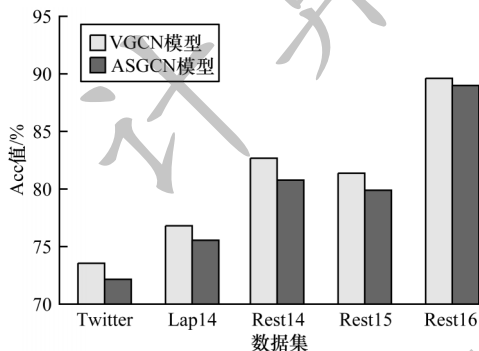


图3 VGCN 与 ASGCN 的 Acc 值对比

Fig.3 Comparison of Acc values between VGCN and ASGCN

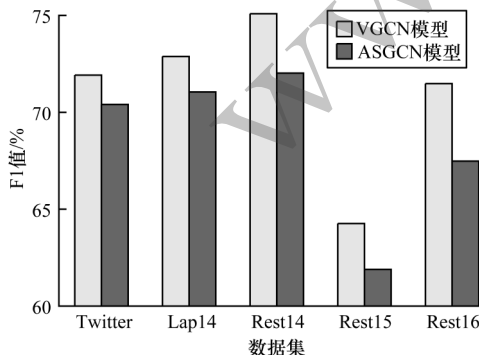


图4 VGCN 与 ASGCN 的 F1 值对比

Fig.4 Comparison of F1 values between VGCN and ASGCN

3.6.2 VGCN 与双知识融合的对比实验

观察 SenticNet 情感知识和 Concept 概念知识可知, SenticNet 是对情感词赋予情感得分,而评论语句中情感词多为意见词, Concept 概念知识是针对方面词进行概念化。为了验证 2 种外部知识对模型性能的影响,本文设计对比实验,分别采用融合 SenticNet 情感知识、融合 Concept 概念知识以及双知识融合的方式,实验结果如图5、图6所示。

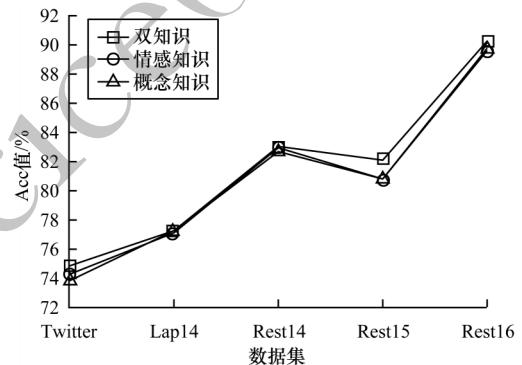


图5 VGCN 融合外部知识时的 Acc 值

Fig.5 Acc values when VGCN integrates external knowledge

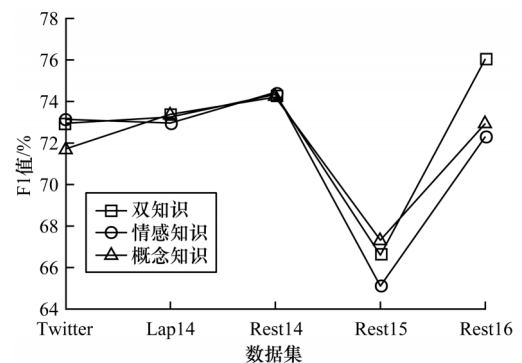


图6 VGCN 融合外部知识时的 F1 值

Fig.6 F1 values when VGCN integrates external knowledge

3.6.3 单交互与双交互的对比实验

为了验证语义交互与语法交互对本文模型的影响, 针对 VKI-GCN 模型的多交互层分别设计语义交互、语法交互和双交互的对比实验, 其他实验条件与模型架构保持不变, 实验结果如图 7、图 8 所示。

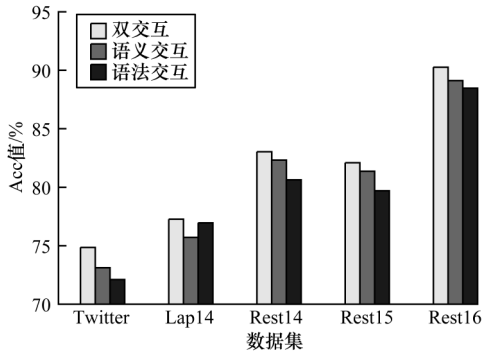


图 7 单交互与双交互的 Acc 值

Fig.7 Acc values for single and double interactions

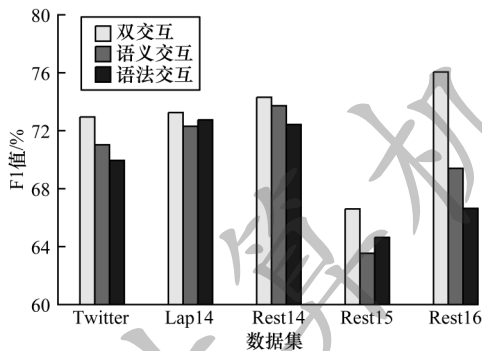


图 8 单交互与双交互的 F1 值

Fig.8 F1 values for single and double interactions

通过表 5 及上述 3 组对比实验可以发现: 面向方面构建虚拟依存关系的依赖树, 对虚拟依存关系赋权后所得的邻接矩阵能够更好地融合句法信息, 在多组数据集上均能获得大于 1 个百分点的性能提升; 引入外部知识能够有效增强评论语句的特征表示, SenticNet 情感知识和 Concept 概念知识对评论语句增强的着力点不同, 在 5 个数据集上性能提升各不相同, 但是整体而言在 ABSA 任务上结合使用外部知识的效果优于单独使用外部知识; 同时进行语义交互和语法交互能够有效提升模型性能。

3.7 GCN 层数分析

本次实验将 GCN 层数分别设置为 $L=\{1, 2, 3, \dots, 8\}$, 在 5 个公开数据集上进行实验的 Acc 值和 F1 值分别如图 9、图 10 所示。从中可以看出: 由于 Rest15 数据集的语句简单且数量少, 即在第 1 层取得了最优值; 从 5 个数据集总体效果来看, VKI-GCN 在 GCN 层数为 2 时达到了最优性能, 随着网络深度的增加, 模型性能发生波动, 但整体性能呈现下降趋势, 当层数为 8 时, 相较于最佳性能, 准确率平均下降 2.03%, F1 值平均下降 3.42%, 由于网络层数增加, 模型引入了过多参数, 产生了过拟合现象。

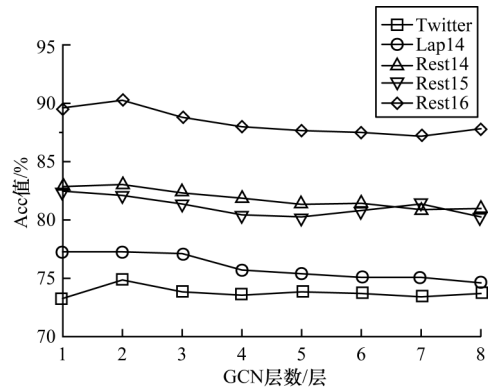


图 9 GCN 网络深度与准确率的关系

Fig.9 Relationship between GCN network depth and accuracy

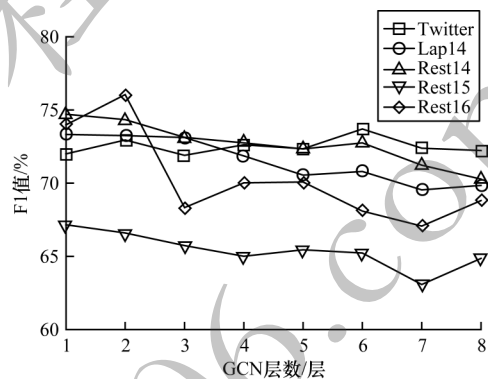


图 10 GCN 网络深度与 F1 值的关系

Fig.10 Relationship between GCN network depth and F1 values

3.8 样例分析

为了直观理解基于方面构建虚拟依存关系依赖树所生成的邻接矩阵, 本文针对上述邻接矩阵进行可视化对比。

从 Rest16 数据集中选取一条评论语句 “my friend got the mushroom pizza which tasted better.” 进行句法依存分析, 其邻接关系如图 11 所示, 其中, 下划线为方面词, 区域颜色越深代表单词之间的关系权重越大。因为普通的句法依存图的邻接矩阵被定义为二进制, 2 个节点词之间存在句法依存关系时用 1 表示, 否则用 0 表示, 所以图中有邻接关系的区域颜色相同。

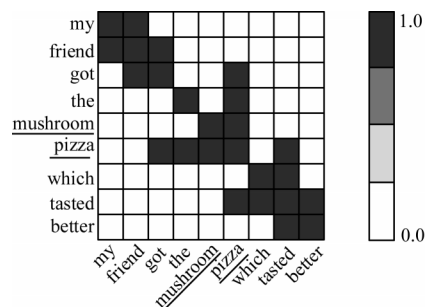


图 11 句法依存邻接关系

Fig.11 Syntactic dependency adjacency

基于方面重构依赖树,对不同的虚拟依存关系分配不同的权重,虚拟关系 $n:con$ 的 n 值越小,句法依存关系越紧密,权重值越大。虚拟关系如图 12 所示,其中,下划线为方面词,区域颜色越深代表虚拟关系权重越大。

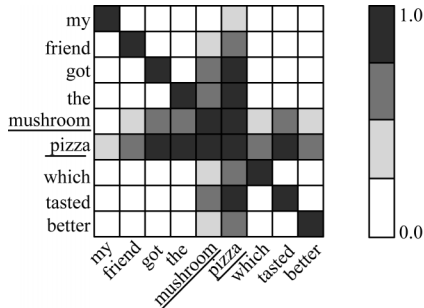


图 12 虚拟依存邻接关系

Fig.12 Virtual dependency adjacency

就评论句“my friend got the mushroom pizza which tasted better.”而言,通过本文设计的 VGCN 模型对注意力分数进行可视化分析,如图 13 所示。

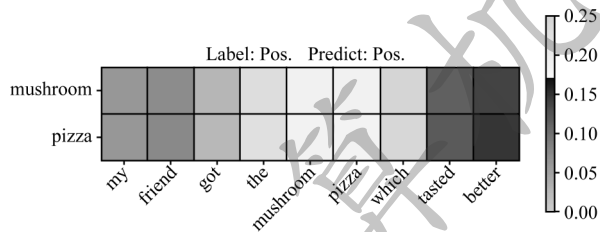


图 13 VGCN 模型的可视化结果

Fig.13 Visual results of VGCN model

在图 13 中,第 1 行是方面词中第 1 个单词的交互注意力分数,第 2 行是第 2 个单词的交互注意力分数,区域颜色越深,注意力权重越大。方面词“mushroom pizza”中每一个单词都可以对情感词“better”正确分配出较高的权重,同时“pizza”相较于“mushroom”在情感词“better”上又能获得较高的权重,体现了方面词内部之间的关联,同时方面词中的每个单词对目标任务的重要程度得以区分。从图 13 可看出,注意力分数整体呈现以方面为中心向周围扩散并着重趋于情感词的分布,这正体现出本文模型构建以方面词为中心的邻接矩阵,能够针对特定方面进行重点关注,从而突出目标任务。

4 结束语

针对方面级情感分析任务,本文提出一种 VKI-GCN 模型。将方面词的每个单词分别作为根节点构建依赖树,为方面词与节点词创建虚拟关系并赋予渐进权重,增强依存树的特征表示,获得针对方面词的更丰富的语法信息。同时,考虑到外部知识对方面级情感分析任务的影响,VKI-GCN 模型同时融合情感知识和概念知识,丰富评论语句的语义信息。实验结果表明,本文模型设计的虚拟依存关

系与双知识增强方式在方面级情感分析任务中能够发挥积极作用。由于本文模型融合了双外部知识,带来了过大噪声,因此下一步将解决由外部知识带来的噪声问题,学习更有利的外部知识同时设计更协调的融合方式,以提高本文模型的情感分析效果。

参考文献

- [1] JIANG L, YU M, ZHOU M, et al. Target-dependent Twitter sentiment classification [C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics; Human Language Technologies. [S. l.]: Association for Computational Linguistics, 2011: 151-160.
- [2] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification[EB/OL]. [2022-09-05]. <https://arxiv.org/abs/1512.01100>.
- [3] MA D H, LI S J, ZHANG X D, et al. Interactive attention networks for aspect-level sentiment classification [C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Washington D. C., USA: IEEE Press, 2017: 4068-4074.
- [4] KIPF T N, WELLMING M. Semi-supervised classification with graph convolutional networks[EB/OL]. [2022-09-05]. <https://arxiv.org/abs/1609.02907>.
- [5] HUANG B, CARLEY K M. Syntax-aware aspect level sentiment classification with graph attention networks [EB/OL]. [2022-09-05]. <https://aclanthology.org/D19-1549.pdf>.
- [6] ZHANG C, LI Q C, SONG D W. Aspect-based sentiment classification with aspect-specific graph convolutional networks [C]//Proceedings of 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. [S. l.]: Association for Computational Linguistics, 2019: 4568-4578.
- [7] LIANG B, YIN R D, GUI L, et al. Jointly learning aspect-focused and inter-aspect relations with graph convolutional networks for aspect sentiment analysis [C]//Proceedings of the 28th International Conference on Computational Linguistics. Washington D. C., USA: IEEE Press, 2020: 150-161.
- [8] 王汝言,陶中原,赵容剑,等. 多交互图卷积网络用于方面情感分析[J]. 电子与信息学报, 2022, 44(3): 1111-1118. WANG R Y, TAO Z Y, ZHAO R J, et al. Multi-interaction graph convolutional networks for aspect-level sentiment analysis [J]. Journal of Electronics & Information Technology, 2022, 44(3): 1111-1118. (in Chinese)
- [9] 齐嵩喆,黄贤英,孙海栋,等. 基于渐进增强与图卷积的方面级情感分析模型[J]. 计算机应用研究, 2022, 39(7): 2037-2042. QI S Z, HUANG X Y, SUN H D, et al. Aspect based sentiment analysis with progressive enhancement and graph convolution [J]. Application Research of Computers, 2022, 39(7): 2037-2042. (in Chinese)
- [10] WANG K, SHEN W Z, YANG Y Y, et al. Relational graph attention network for aspect-based sentiment analysis [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. [S. l.]: Association for Computational Linguistics, 2020: 3229-3238.

- [11] MA Y K, PENG H Y, KHAN T, et al. Sentic LSTM: a hybrid network for targeted aspect-based sentiment analysis [J]. *Cognitive Computation*, 2018, 10(4): 639-650.
- [12] LIANG B, SU H, GUI L, et al. Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks [J]. *Knowledge-Based Systems*, 2022, 235: 107643.
- [13] CHEN J D, HU Y Z, LIU J P, et al. Deep short text classification with knowledge powered attention [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 6252-6259.
- [14] BIAN X M, FENG C, AHMAD A, et al. Targeted sentiment classification with knowledge powered attention network [C]//*Proceedings of the 31st IEEE International Conference on Tools with Artificial Intelligence*. Washington D. C. , USA: IEEE Press, 2020: 1073-1080.
- [15] KIM Y. Convolutional neural networks for sentence classification [C]//*Proceedings of 2014 Conference on Empirical Methods in Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2014: 1746-1751.
- [16] FAN F F, FENG Y S, ZHAO D Y. Multi-grained attention network for aspect-level sentiment classification [C]//*Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2018: 3433-3442.
- [17] ZHANG M, QIAN T Y. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis [C]//*Proceedings of 2020 Conference on Empirical Methods in Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2020: 3540-3549.
- [18] CAMBRIA E, SPEER R, HAVASI C, et al. SenticNet: a publicly available semantic resource for opinion mining [EB/OL]. [2022-09-05]. <http://sentic.net/senticnet.pdf>.
- [19] CAMBRIA E, LI Y, XING F Z, et al. SenticNet 6: ensemble application of symbolic and subsymbolic AI for sentiment analysis [C]//*Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. New York, USA: ACM Press, 2020: 105-114.
- [20] JI L, WANG Y J, SHI B T, et al. Microsoft concept graph: mining semantic concepts for short text understanding [J]. *Data Intelligence*, 2019, 1(3): 238-270.
- [21] CAO S L, QIAN B Y, YIN C C, et al. Knowledge guided short-text classification for healthcare applications [C]//*Proceedings of IEEE International Conference on Data Mining*. Washington D. C. , USA: IEEE Press, 2017: 31-40.
- [22] HU L M, YANG T C, SHI C, et al. Heterogeneous graph attention networks for semi-supervised short text classification [C]//*Proceedings of 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2019: 4821-4830.
- [23] CHEN P, SUN Z Q, BING L D, et al. Recurrent attention network on memory for aspect sentiment analysis [C]//*Proceedings of 2017 Conference on Empirical Methods in Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2017: 452-461.
- [24] DONG L, WEI F R, TAN C Q, 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*. [S. l.]: Association for Computational Linguistics, 2014: 49-54.
- [25] 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*. [S. l.]: Association for Computational Linguistics, 2014: 27-35.
- [26] 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*. [S. l.]: Association for Computational Linguistics, 2015: 486-495.
- [27] 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*. [S. l.]: Association for Computational Linguistics, 2016: 19-30.
- [28] CHEN C H, TENG Z Y, ZHANG Y. Inducing target-specific latent structures for aspect sentiment classification [C]//*Proceedings of 2020 Conference on Empirical Methods in Natural Language Processing*. [S. l.]: Association for Computational Linguistics, 2020: 5596-5607.
- [29] 巫浩盛, 缪裕青, 张万桢, 等. 基于距离与图卷积网络的方面级情感分析 [J]. *计算机应用研究*, 2021, 38(11): 3274-3278, 3321.
- [30] WU H S, MIAO Y Q, ZHANG W Z, et al. Aspect level sentiment analysis based on distance and graph convolution network [J]. *Application Research of Computers*, 2021, 38(11): 3274-3278, 3321. (in Chinese)
- [31] HOU 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]. *Knowledge-Based Systems*, 2020, 205: 106292.
- [32] XU G T, LIU P Y, ZHU Z F, et al. Attention-enhanced graph convolutional networks for aspect-based sentiment classification with multi-head attention [J]. *Applied Sciences*, 2021, 11(8): 3640-3653.