

基于句法增强与形变LSTM的图卷积网络方面级情感分析

张光明¹, 张弛¹, 高谦¹, 谈栋华¹, 张正悦¹, 彭菊红^{1,2}

(1. 湖北大学人工智能学院; 2. 湖北大学智能感知系统与安全教育部重点实验室, 湖北武汉430062)

摘要: 目前大多数图卷积网络模型均是在单一句法解析树上构建图结构, 当处理复杂长句时, 容易出现句法歧义, 导致句法分析结果不准确, 从而影响方面级分析的结果。为解决上述问题, 提出一种基于句法增强和形变LSTM的图卷积网络模型。首先, 通过多个外部句法解析器对输入数据进行句法解析, 得到多个句法解析树; 然后, 设计融合策略, 获得增强的依赖树; 最后, 构建一个双向的形变LSTM, 通过奇偶交替算法对输入信息与隐藏信息进行耦合迭代, 从而解决上下文信息未能充分利用的问题。模型在Restaurant、Laptop和Twitter三个数据集上的准确率和F1值分别达到了84.74%和78.44%、79.21%和75.76%、75.98%和74.38%。

关键词: 形变LSTM; 句法增强; 图卷积网络; 方面级情感分析

DOI: 10.11907/rjdk.241765

扫描二维码阅读全文:

中图分类号: TP391.1

文献标识码: A

文章编号: 1672-7800(2025)010-0097-07



Aspect-level Sentiment Analysis of Graph Convolutional Networks Based on Syntactic Enhancement and Mogrifier LSTM

ZHANG Guangming¹, ZHANG Chi¹, GAO Qian¹, TAN Donghua¹, ZHANG Zhengyue¹, PENG Juhong^{1,2}

(1. School of Artificial Intelligence, Hubei University;

2. Ministry of Education Key Laboratory of Intelligent Sensing System and Security, Hubei University, Wuhan 430062, China)

Abstract: Most graph convolutional network models were constructed on a single syntactic analytic tree, leading to inaccurate syntactic analysis results when dealing with complex long sentences. To address this issue, a graph convolutional network model based on syntactic enhancement and Mogrifier LSTM was proposed. Firstly, the input data was parsed by multiple external syntactic parsers to obtain multiple syntactic parser trees. Then, a fusion strategy was designed to obtain an enhanced dependency tree. Finally, a bidirectional Mogrifier LSTM (Mogrifier BiLSTM) was constructed, and the input information and hidden information were coupled and iterated using an odd-even alternation algorithm, which solved the problem of underutilized context information. The model accuracy and F1 values of the Restaurant, Laptop, and Twitter datasets reached 84.74% and 78.44%, 79.21% and 75.76%, and 75.98% and 74.38%, respectively.

Key Words: Mogrifier LSTM; syntactic enhancement; graph convolutional networks; aspect-level sentiment analysis

0 引言

方面级情感分析(aspect-based sentiment analysis, AB-SA)的目的是给定一个句子, 然后确定句子中不同方面词

的情感极性, 可以准确的识别出用户对于某一个方面的态度, 而不是简单的对一个句子进行情感判别^[1]。例如, 对于句子“The environment of the hotel is good, but the attitude of the waiter is not very good”, 其中就包含了“environment”和“attitude”2个不同的方面词, 这2个方面词所对应的观

收稿日期: 2024-09-05

基金项目: 湖北省重点研发计划项目(2023BAB068)

作者简介: 张光明(1999-), 男, 湖北大学人工智能学院硕士研究生, 研究方向为自然语言处理; 张弛(2000-), 男, 湖北大学人工智能学院硕士研究生, 研究方向为目标检测; 高谦(2001-), 女, 湖北大学人工智能学院硕士研究生, 研究方向为强化学习; 谈栋华(2001-), 男, 湖北大学人工智能学院硕士研究生, 研究方向为语义分割; 张正悦(2000-), 男, 湖北大学人工智能学院硕士研究生, 研究方向为自然语言处理; 彭菊红(1978-), 女, 博士, 湖北大学人工智能学院副教授, 研究方向为信号处理、电路建模与分析、人工智能算法。本文通讯作者: 彭菊红。

点却是相反的。

随着深度学习的兴起,逐渐出现很多基于注意力机制的方法,注意力机制可以对句子中的情感词给予很高的权重。但是注意力机制很容易受到句子噪声的影响,导致最终结果不够理想。传统的 LSTM^[2]难以充分利用上下文信息,于是出现了很多基于图卷积网络(Graph Convolutional Networks, GCN^[3])的方法。GCN 是根据句子的句法结构生成相对应的依赖树,能够充分考虑到句子本身的语法关系和句法关系,还能够很好的解决上述问题。在现有的图卷积网络(GCN)模型中,大多数方法仅使用一种句法解析器来构建句子的依存关系图,这在处理复杂长句时往往导致分析结果不准确,从而影响最终的情感分析性能。例如, Zhang 等^[4]通过单一解析器建立有向和无向依存图,再送入 GCN 层提取特征。Li 等^[5]则构建了一个双重图卷积网络,同时考虑句子的语法和语义结构。这些方法虽然在一定程度上提高了模型的性能,但由于依赖单一解析器,其准确性受限于解析器的精度。此外, Zhang 等^[6]融入方面感知注意力机制, Liang 等^[7]整合 SenticNet^[8]中的情感知识的方法,都未能实现对情感信息的深度挖掘。程帆等^[9]提出混合编码的双通道 GCN 和丁美荣等^[10]构建的句法依赖树、上下文聚焦机制,虽然引入了多元化信息特征,但仍存在对句法和语义信息提炼不充分的问题。张文豪等^[11]设计的多信息学习模型和廖列法等^[12]提出的增强图模型也面临类似的挑战,即未全面考虑到句子的语法语义信息。张吴波等^[13]提出的多通道句法门控图神经网络的句子级情感分析方法,杨春霞等^[14]提出的双图卷积与门控线性单元网络,以及穆一茹等^[15]提出的多层面信息解读,虽然都各有创新,但是均未能捕捉到局部细节,无法深层次的挖掘句子句法结构。上述方法均仅使用一种句法解析器进行句子句法结构分析,这导致了对句法解析器解析精度的过度依赖。然而,当面对复杂句法时,这些方法往往难以避免句法歧义问题,因此在对句子的句法结构、语法结构以及语义信息进行深度解析时存在一定的困难。

鉴于以上讨论,本研究提出了一种基于多句法信息融合和形变 LSTM 的图卷积网络模型,旨在解决现有方法在处理复杂句法结构时的局限性,从而有效缓解句法歧义问题。首先,利用融合策略融合多个句法解析树来解决单一解析器解析不准确的问题;其次,提出一种基于奇偶交替算法的双向形变 LSTM 模型,利用 Mogrifier 门控机制的奇偶交替迭代,赋予更丰富的超参数实现门控信息的深度交互,有效地耦合和迭代输入信息与隐藏状态^[16];再次,在注意力层后加入语义掩码,以获取更全面的句法结构信息。本文的主要工作与贡献如下:①提出一种基于句法增强和形变 LSTM 的图卷积网络模型(SMGCN),设计融合策略融合多个句法解析树实现句法增强,并结合双向的 MogrifierLSTM 来提取丰富的上下文信息,获得更加完整的上下文特征;②在注意力层后面加入语义掩码,获取更全面的

句法结构信息,帮助模型学习局部的细节和整体的语义信息;③本文提出的模型在 Restaurant、Laptop 和 Twitter 三个数据集上进行了实验,3 个数据集上准确率都有一定的提升,证明了本文方法的有效性。

1 SMGCN 模型

本文的模型命名为 SMGCN,其骨干网络主要由输入编码层、双向形变 LSTM(Mogrifier BiLSTM)、句法融合层、注意力层、语义掩码层和 GCN 层 6 个模块组成。输入编码层的作用是将输入句子转换为相应的字符序列,并将其表示为字符向量;Mogrifier BiLSTM 是一种融合了 Mogrifier 机制的双向 LSTM 模型,它通过 Mogrifier 机制独特的奇偶交替算法对当前输入信息与上一层的隐藏信息进行耦合迭;法融合层是利用融合策略融合多个句法解析树来解决单一解析器解析不准确的问题;注意力层的作用是通过多头自注意力机制对数据进行注意力计算,加强方面词和情感词之间的关联;语义掩码层是通过添加句法掩码矩阵,可以有选择地增强或削弱特定词之间的注意力权重,使得模型更专注于有实际语义关联的词对,提高模型的注意力得分矩阵的准确性;GCN 层的作用是进一步整合句法结构信息,增强模型对依存关系的建模能力。SMGCN 的模型的结构如图 1 所示。

1.1 输入编码层

给定一个句子: $s = \{w_1, w_2, \dots, a_1, a_2, \dots, w_n\}$, 使用 Glove 词嵌入把每个词映射为低维词嵌入向量,即将句子 S 转化为向量 V 表示, $V = \{v_1, v_2, \dots, a_1, a_2, \dots, v_n\}$ 。其中, $v_i \in R_d$, d 表示单词嵌入空间的维度。随后将句子的词嵌入向量送入 MogrifierBiLSTM 层,得到句子的隐藏状态向量表示 $H = \{h_1, h_2, \dots, h_{a_1}, h_{a_2}, \dots, h_n\}$ 。其中, $h_i \in R_{2d}$, d 表示单向 LSTM 的隐藏层维度。

1.2 双向形变 LSTM(Mogrifier BiLSTM)

针对 LSTM 网络在处理长距离语义时可能遇到的梯度消失和信息遗忘问题,本文提出了一种改进方案。在 LSTM 网络中,当前的输入和隐藏状态是独立输入网络进行学习的,二者仅在门中进行信息交互,这种相对独立的关系很容易导致上下文中的重要信息丢失。为解决这一问题,本文在 LSTM 模型中引入了 Mogrifier 门控机制,实现了当前输入和之前的隐藏状态之间进行奇偶交互运算,使其可以充分交互。在交互的过程中得到新的输入和新的隐藏状态,随后再送入到 LSTM 网络中,进而显著提高模型对上下文信息的利用。图 2 为 Mogrifier LSTM 模型结构图,清晰的说明了其工作机制。Mogrifier BiLSTM 网络通过引入双向传播机制,实现对数据的充分理解, Mogrifier 模块增强了数据耦合关系,加强了对上下文信息的利用。这一改进不仅提升了 LSTM 网络的性能,也为处理复杂序列数据提供了更加有效的方法。

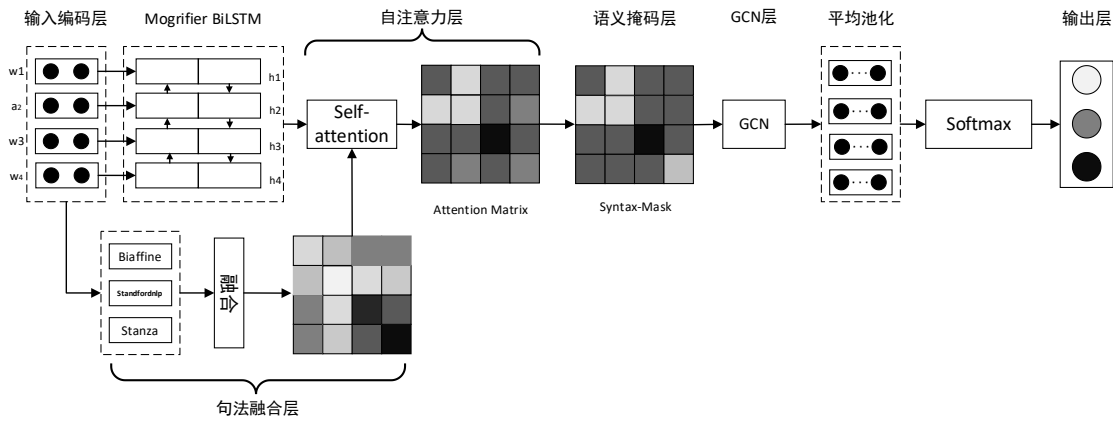


Fig. 1 SMGCN model structure

图1 SMGCN模型结构

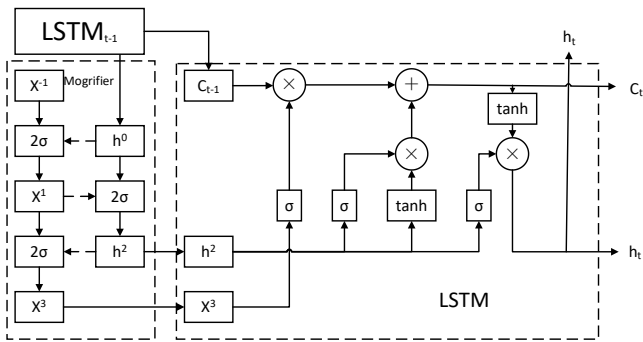


Fig. 2 Mogriifier LSTM structure

图2 Mogriifier LSTM结构

Mogriifier LSTM 的计算过程,如公式(1)-(2)所示:

$$x^i = 2\sigma(Q^i h^{i-1}) \odot x^{i-2} \text{ for odd } i \in [1, \dots, r] \quad (1)$$

$$h^i = 2\sigma(R^i x^{i-1}) \odot h^{i-2} \text{ for even } i \in [1, \dots, r] \quad (2)$$

其中, $x = x^i$ 表示当前的输入; $h^i = h^0$ 表示隐藏状态; 上标 i 表示在进行第 i 次交互时所对应的变量; r 表示当前输入 x 和隐藏状态之 h 间的交互次数。当 $r=0$ 时, 所表示的就是一个基本的 LSTM 网络。

1.3 句法融合层

为了解决单一句法解析器可能存在解析不准确的问题, 本文提出利用多个句法解析器进行融合, 使用 3 个解析器对一个句子进行依赖解析, 从而获得 3 个解析数组。随后, 利用融合算法对 3 个解析数组进行融合, 获得最佳解析数组。

算法 1 句法融合算法

输入: 一个句子 S

输出: 融合后的一个解析数组

1. 将一个句子分别送入三个不同的解析器中, 获得三个解析数组 $array1$, $array2$, $array3$
2. 获取这三个解析数组的长度 $length$
3. 新建一个空的数组 $empty = []$
4. 遍历每个位置 i (从 0 到 $length-1$)

5. 尝试执行以下操作

6. 取出每个解析数组在第 i 个位置的元素 $elements = [array1[i], array2[i], array3[i]]$

7. 统计 $elements$ 中每个元素出现的次数 $counter = \{\}$

8. for element in elements

9. if element in counter

10. $counter[element] += 1$

11. else

12. $counter[element] = 1$

13. 找到出现次数最多的元素及其出现次数 $a, b = \max(counter.items(), key=\lambda x: x[1])$

14. if $b > \text{len}(arrays) / 2$

15. $empty.append(a)$

16. else

17. $parser_votes = [0, 0, 0]$

18. for $parser_index, element$ in $\text{enumerate}(elements)$

19. if element in counter

20. $parser_votes[parser_index] += counter[element]$

21. $\max_vote_parser = \text{parser_votes.index}(\max(parser_votes))$

22. $empty.append(arrays[\max_vote_parser][i])$

23. $parser_votes.clear()$

24. 如果发生 IndexError 异常

25. 打印错误信息并退出循环

26. 返回结果数组 $empty$

图 3 展示了句法融合算法的例子。从图中可以看到, 对于 $array1=[1, 0, 1]$, 经过句法融合算法后输出的结果是 1; 对于 $array8=[8, 8, 8]$, 经过句法融合算法后输出结果是 8。

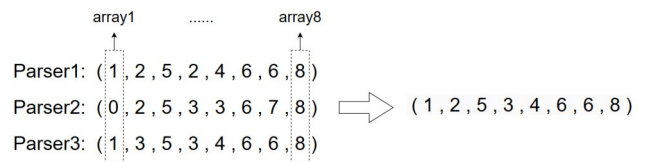


Fig. 3 Syntax fusion example

图3 句法融合示例

1.4 注意力层

本文采用的注意力是基于 p-head 自注意构建的,通过这种方式捕获句子中 2 个任意单词之间的相互作用信息。计算公式如式(3)所示:

$$A = \frac{QW^Q \times (KW^K)^T}{\sqrt{d}} \quad (3)$$

其中, Q 和 K 等于编码层产生的 H 。 $W^Q \in R^{d \times d}$, $W^K \in R^{d \times d}$ 是可学习的权重。

1.5 语义掩码层

在注意力机制中,每个词都会对其他所有词进行注意力计算,其中也包括部分毫不相关的词。为此,本文通过引入句法掩码矩阵,可以选择性地增强或削弱特定词对之间的注意力权重,使模型更加专注于具有实际语义关联的词对,从而提高注意力得分矩阵的准确性。通过引入句法掩码矩阵可以获得更全面的句法结构信息,帮助模型学习到局部的细节和整体的语义信息。对于每个句法解析树,根据不同的语法距离对解析树中的一些节点进行掩码处理。首先,假设有 2 个节点 s_i 和 s_j ,他们之间的距离表示为 $d(s_i, s_j)$,最短距离表示为 $D(i, j) = \min d(s_i, s_j)$;然后,根据句法距离构造 P 个句法掩码矩阵;最后,再将注意力得分矩阵和句法掩码矩阵结合,利用公式(4)得到最终的句法掩码矩阵。

$$A_{mask}^i = \text{softmax}(A^i + M^i) \quad (4)$$

其中, A_{mask}^i 表示第 i 个注意力得分矩阵 A^i 与第 i 个句法掩码矩阵 M^i 结合得到的句法掩码矩阵。

1.6 GCN 层

根据前文所述,本文构建了 p 个不同的句法掩码矩阵,因此需要对 A_{mask} 进行 p 次卷积操作。GCN 层中的每个节点的更新过程如公式(5)所示。

$$h_i^l = \sigma \left(\sum_{j=1}^n A_{ij} W^l h_j^{l-1} + b^l \right) \quad (5)$$

其中, W^l 和 b^l 分别表示线性变换权重值和偏置项; σ 表示非线性函数。

GCN 的最终输出如公式(6)所示。

$$H^l = \{h_1^l, h_2^l, \dots, h_n^l\} \quad (6)$$

其中, h^{l-1} 为第 l 层的输入状态; h^l 为第 l 层的输出状态。

对 SMGCN 各层的节点表示进行聚合后,得到最终的特征表示。本文对 GCN 层学习到的输出表示中的非方面词进行屏蔽,以获得方面词表示。除此之外,平均池化保留了方面项表示中的大部分信息 h_a^l ,具体如公式(7)所示。

$$h_a^l = f(h_{a1}^l, h_{a2}^l, \dots, h_{an}^l) \quad (7)$$

其中, $f(\cdot)$ 是 GCN 层在增强的方面表示上应用的均值池化函数。随后,将 h_a^l 输入线性层,并使用 softmax 函数生成极性决策空间上的概率分布,如公式(8)所示:

$$p(a) = \text{softmax}(W_p h_a^l + b_p) \quad (8)$$

其中, w_p 是学习的权重; b_p 是可学习的偏差。

2 实验结果分析

本文实验设备为 NVIDIA GeForce RTX3060ti,显存为 12GB,实验所选择的深度学习框架为 PyTorch transformers 3.4.0,选择的优化器为 Adam,部分参数 Batch size 设置为 16, Embed_dim=300, Epoch=50 等。

2.1 数据集介绍

为了验证模型的性能,在 3 个公开的数据集进行训练评估,包括 SemEval2014Task4^[17] 的 Restaurant、Laptop 评论数据和 Twitter^[18] 数据集。3 个数据集的具体划分如表 1 所示。

Table 1 Dataset description

表 1 数据集描述

数据集	划分	积极	消极	中性
Restaurant	Train	2164	807	637
	Test	727	196	196
Laptop	Train	976	851	455
	Test	337	128	167
Twitter	Train	1507	1528	3016
	Test	172	169	336

2.2 评价指标

实验采用的评价指标为准确率 (Accuracy, Acc) 和 F1。Acc 表示分类正确的样本数占比,计算公式如式(9)所示。F₁ 是精确率 (P) 和召回率 (R) 的调和平均值,计算公式如式(10)-(12)所示。

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (9)$$

$$F_1 = \frac{2PR}{P + R} \quad (10)$$

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

其中, TP 为预测结果是正,实际也是正; FP 为预测结果是正,实际是负; TN 为预测结果是负,实际也是负; FN 为预测结果是负,实际为正。

2.3 对比实验

为了验证本文模型的性能,将其与以下基线方法进行对比实验。

(1) ASGCN^[4]。构造两种句法依存图,并输入 GCN 网络学习 ABSA 的句法信息和词依赖关系。

(2) CDT^[19]。在句子依存关系树上利用 GCN 模型提取上下文词和方面词之间的依赖关系。

(3) BiGCN^[20]。构建了一种双向 GCN,通过句法图和词法图来进行情感预测。

(4) kumaGCN^[21]。结合依赖图和潜在图的信息来学习语法特征。

(5)R-GAT^[22]。研究一种关系图注意力网络,并对基于方面的树结构进行编码,用于情感预测。

(6)DGEDT^[23]。提出一种双向的 transformer 结构的网络,通过依赖图增强 transformer 网络学习平面表示和基于图的表示。

(7)DualGCN^[5]。设计一个具有正交和微分正则器的 SynGCN 模块和 SemGCN 模块。

(8)MILFST^[11]。设计一种融合句法树多信息学习的模型,能够根据依存句法树的树形结构更新序列信息,并利用卷积神经网络和图卷积网络学习语义信息和语法信息。

(9)SDEG^[12]。通过在原始句法依赖图上引入外部情感知识和词性信息,增强情感词的权重。

(10)AEDGCN^[13]。设计一个情感增强双图卷积网络,通过引入外部情感知识来增强图结构。

(11)BGCCN^[14]。设计一个双图卷积与门控线性单元网络。

(12)R-GAT+BERT^[22]。基于预训练 BERT 的 RGAT 模型。

(13)DGEDT+BERT^[23]。基于预训练 BERT 的 DGEDT 模型。

(14)BERT^[24]。采用“[CLS]句子[SEP]方面[SEP]”作为输入。

(15)BERT4GCN^[25]。整合 BERT 的 PLM 语法顺序特征和依赖图的语法知识。

2.4 实验结果

实验结果如表 2 所示。相较于基线模型 ASGCN、CDT、BiGCN、KumaGCN、R-GAT 和 DGEDT,本文模型的实验结果表现出色。现有模型主要是基于句法树的 GCN 模型,未能充分利用句子的语法和语义信息。这表明,在利用 GCN 模型进行学习时,不仅要考虑句法结构,还需结合句子语法语义信息来帮助模型充分理解上下文结构。然而,DualGCN、BGCCN 和 AEDGCN 模型通过结合句法和语义信息,在实验结果上明显优于前面仅考虑句法信息的模型。这表明,利用句法信息和语义信息进行互相补充,可以有效弥补邻接矩阵中包含信息不足的问题。SMGCN 在大多数数据集上均取得了最优性能,仅在 Twitter 数据集上的表现略低于 AEDGCN。这表明,通过融合句法信息可以有效解决单一句法解析器解析错误的问题,同时引入 Mogrifier 机制的 LSTM 可以有效解决上下文信息不能充分利用的问题。此外,研究发现采用注意力机制的模型普遍比未采用注意力机制的模型表现更优,表明注意力机制能有效帮助模型学习上下文语义信息,并关注与方面词相关的词汇,从而提升模型分类性能。

本文将模型与预训练模型 BERT 相结合,在 Restaurant、Laptop 和 Twitter 三个数据集上的准确率和 F1 值分别达到了 87.42% 和 81.46%、80.54% 和 77.26%、77.85% 和 76.76%。

Table 2 Comparative experiment results

表 2 对比实验结果

Models	Restaurant		Laptop		Twitter	
	Acc	F1	Acc	F1	Acc	F1
ASGCN ^[4]	80.77	72.02	75.55	71.05	72.15	70.40
CDT ^[19]	82.30	74.02	77.19	72.99	74.66	73.66
BiGCN ^[20]	81.97	73.48	74.59	71.84	74.16	73.35
KumaGCN ^[21]	81.43	73.64	76.12	72.42	72.45	70.77
R-GAT ^[22]	83.30	76.08	77.42	73.76	75.57	73.82
DGEDT ^[23]	83.90	75.10	76.80	72.30	74.80	73.40
DualGCN ^[5]	84.27	78.08	78.48	74.74	75.92	74.29
MILFST ^[11]	82.50	74.54	77.74	74.27	74.27	73.14
SDEG ^[12]	82.41	74.26	77.12	73.44	72.54	71.46
AEDGCN ^[13]	84.32	77.54	78.20	74.82	76.12	74.44
BGCCN ^[14]	-	-	77.61	73.52	74.82	72.97
Our SMGCN	84.74	78.44	79.21	75.76	75.98	74.38
R-GAT+BERT ^[22]	86.60	81.35	78.21	74.07	76.15	74.88
DGEDT+BERT ^[23]	86.30	80.00	79.80	75.60	77.90	75.40
BERT ^[24]	85.97	80.09	79.91	76.00	75.92	75.18
BERT4GCN ^[25]	84.75	77.11	77.49	73.01	74.73	73.76
Our SMGCN+BERT	87.42	81.46	80.54	77.26	77.85	76.76

2.5 消融实验

为检验不同模块在本文模型中的重要性,进行了消融实验,实验结果如表 3 所示。w/o DTF 代表移除 SMGCN 网络的句法依赖树融合模块;w/o mogrifier 代表移除 SMGCN 网络的 mogrifier 模块,使用简单的 BiLSTM;w/o self-attention 代表移除 SMGCN 网络的 self-attention 模块;w/o syntactic-mask 代表移除 SMGCN 网络的语义掩码模块。首先,实验结果表明在去除依赖树融合模块后,模型的性能确实有所下降,这一结果验证了融合策略的有效性。通过融合多个句法解析树的结果来获得增强的依赖树,可以有效解决单一句法解析器在处理复杂长句时容易出现句法歧义的问题。其次,实验表明在没有加入 Mogrifier 机制的情况下,模型的表现并不令人满意,其缺乏对上下文信息的充分利用,导致在 Restaurant、Laptop 和 Twitter 上的准确率分别降低了 0.7%、0.6% 和 0.5%。这表明,Mogrifier 机制可以实现当前输入和隐藏状态的相互交互,从而解决上下文信息不能充分利用的问题。此外,实验表明在去除自注意力模块后,模型的效果有所降低。该结果表明,模型在捕捉方面相关语义特征的能力上存在不足。因此,可以得出结论,自注意力是获取相关语义信息的必要条件。最后,实验在去除句法掩码矩阵后,观察到在 Restaurant、Laptop 和 Twitter 上的准确率分别降低了 1.09%、0.78% 和 1.09%。这表明,句法掩码矩阵可以帮助 GCN 更好地学习原始依赖树中的句法结构信息,从而帮助模型学习到局部的细节和整体的语义信息。综上所述,消融实验结果表明,依赖树融合模块、Mogrifier 机制、自注意力模块以及句法掩码矩阵等核心组件在模型性能的提升中都发挥了重要作用。这些组件共同协作,使得模型能够更有效地处理复杂长句并减

少句法歧义,同时充分利用上下文信息并捕获相关语义信息。

Table 3 Ablation experiment results

表3 消融实验结果

Models	Restaurant		Laptop		Twitter	
	Acc	F1	Acc	F1	Acc	F1
SMGCN	84.74	78.44	79.21	75.76	75.98	74.38
w/o DTF	83.65	76.58	78.01	74.61	74.89	73.76
w/o mogrifier	84.04	77.44	78.61	75.06	75.48	74.08
w/o self-attention	83.23	77.03	77.96	74.04	74.34	73.02
w/o syntactic-mask	83.65	77.23	78.43	74.81	74.89	73.12

2.6 句法融合实验

本文对不同的解析器进行了消融实验,使用单个解析器的模型性能明显低于使用多个解析器融合的模式。实验结果表明,多个句法解析结果融合可以充分考虑句法结构,进而实现句法信息增强。本文使用的是目前主流的 Deep Biaffine Parser、Stanford NLP 和 Stanza 三个句法解析器。表4中展示了使用单个解析器和本文的多个解析器融合模型的实验数据对比情况。

Table 4 Syntactic fusion experiment results

表4 句法融合实验结果

Models	Restaurant		Laptop		Twitter	
	Acc	F1	Acc	F1	Acc	F1
Biaffine	83.20	75.95	77.53	73.45	74.74	73.6
Stanfordnlp	83.65	76.58	78.01	74.61	74.89	73.76
Stanza	83.82	76.61	78.16	74.64	75.33	73.83
SMGCN	84.74	78.44	79.21	75.76	75.98	74.38

3 结语

本文提出了一种基于句法增强和形变 LSTM 的图卷积网络方面级情感分析方法,旨在解决现有方法在处理复杂句法结构时的局限性,以及由此导致的句法歧义问题。首先,本文针对单一句法解析器的解析不准确性问题,提出了一种融合多个句法解析树的策略,通过对多个句法解析结果进行融合,充分考虑了句子的句法结构,进而实现了句法信息的增强;其次,提出一种基于奇偶交替算法的双向形变 LSTM 模型,利用 Mogrifier 门控机制的奇偶交替迭代,赋予更丰富的超参数实现门控信息的深度交互,有效地耦合和迭代输入信息与隐藏状态;再次,本文还在注意力层后加入语义掩码,以获取更全面的句法结构信息,并引入了多头自注意力机制,进一步优化了模型性能;最后,根据不同的句法距离构建句子的句法掩码矩阵,学习句法结构信息,以实现句子的语法语义信息相互融合。在未来的工作中,计划将不同类型的知识图谱等引入模型,以此

来实现背景知识的增强,进一步加强意见词和方面词之间的关联。

参考文献:

- [1] TANG H, JI D H, LI C L, et al. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 6578-6588.
- [2] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [3] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[DB/OL]. <https://arxiv.org/abs/1609.02907>.
- [4] 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, 2019: 4568-4578.
- [5] LI R, CHEN H, FENG F, et al. Dual graph convolutional networks for aspect-based sentiment analysis[C]// The 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 6319-6329.
- [6] ZHENG Z, ZHOU Z L, WANG Y N. SSEGNC: syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis[C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2022: 4916-4925.
- [7] 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.
- [8] CAMBRIA E, PORIA S, HAZARIKA D, et al. SenticNet 5: discovering conceptual primitives for sentiment analysis by means of context embeddings[C]//Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, 2018: 1795-1802.
- [9] CHENG F, WANG F, HUANG S C. Aspect-level sentiment analysis based on hybrid-coding and dual-channel GCN[J]. Software Guide, 2023, 22(7): 15-20.
程帆,王芳,黄树成. 基于混合编码与双通道 GCN 的方面级情感分析[J]. 软件导刊, 2023, 22(7): 15-20.
- [10] DING M R, LAI J Q, ZENG B Q, et al. Aspect sentiment analysis based on local-global context guidance[J]. Software Guide, 2024, 23(1): 190-196.
丁美荣,赖锦钱,曾碧卿,等. 基于局部全局上下文引导的方面级情感分析[J]. 软件导刊, 2024, 23(1): 190-196.
- [11] ZHANG W H, LIAO L F, WANG R X. Aspect-level sentiment analysis of multi-information learning with fusion syntactic tree[J]. Computer Engineering, 2023, 49(10): 72-79.
张文豪,廖列法,王茹霞. 融合句法树多信息学习方面级情感分析[J]. 计算机工程, 2023, 49(10): 72-79.
- [12] LIAO L F, XIA W H, YANG Y G. Aspect-level sentiment analysis based on syntactic dependency enhancement graph[J]. Computer Engineering and Design, 2024, 45(6): 1857-1864.

- 廖列法,夏卫欢,杨翌毓. 基于句法依赖增强图的方面级情感分析[J]. 计算机工程与设计, 2024, 45(6): 1857-1864.
- [13] ZHANG W B, ZOU W, XIONG L, et al. Multi-channel syntactic gated graph neural network for sentence-level sentiment analysis[J/OL]. Computer Engineering and Applications, 2024-05-06. <https://link.cnki.net/urlid/11.2127.TP.20240430.1145.008>.
- 张吴波,邹旺,熊黎,等. 多通道句法门控图神经网络用于句子级情感分析[J/OL]. 计算机工程与应用, 2024-05-06. <https://link.cnki.net/urlid/11.2127.TP.20240430.1145.008>.
- [14] YANG C X, WU Y L, YAN H, et al. Aspect-level sentiment analysis model integrating double-graph convolution and gated linear elements[J]. Computer Engineering, 2024, 50(4): 141-149.
- 杨春霞,吴亚雷,闫晗,等. 融合双图卷积与门控线性单元的方面级情感分析模型[J]. 计算机工程, 2024, 50(4): 141-149.
- [15] MU Y R, HAN H, KONG B. Aspect-level sentiment analysis of graph convolutional network based on multi-ordered scale[J/OL]. Computer Engineering and Applications, 2024-07-29. <https://link.cnki.net/urlid/11.2127.tp.20240726.1350.004>.
- 穆一茹,韩虎,孔博. 基于多定序尺度的图卷积网络方面级情感分析[J/OL]. 计算机工程与应用, 2024-07-29. <https://link.cnki.net/urlid/11.2127.tp.20240726.1350.004>.
- [16] MELIS G, KOISK T, BLUNSOM P. Mogrifier LSTM[DB/OL]. <https://arxiv.org/abs/1909.01792>.
- [17] 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.
- [18] 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.
- [19] SUN K, ZHANG R C, SAMUEL M, 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.
- [20] 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 (EMNLP), 2020: 3540-3549.
- [21] CHEN C H, TENG Z Y, ZHANG Y. Inducing target-specific latent structures for aspect sentiment classification [C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020: 5596-5607.
- [22] 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, 2020: 3229-3238.
- [23] TANG H, JI D H, LI C L, et al. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification [C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 6578-6588.
- [24] DEVLIN J, CHANG M Y, LEE K, et al. BERT: pre-training of deep bi-directional transformers for language understanding[DB/OL]. <https://arxiv.org/abs/1810.04805>.
- [25] XIAO Z G, WU J R, CHEN Q L, et al. BERT4GCN: using BERT intermediate layers to augment GCN for aspect-based sentiment classification [DB/OL]. <https://arxiv.org/abs/2110.00171>.

(责任编辑:陈维捷)