



DOI:10.12404/j.issn.1671-1815.2403653

引用格式:余本功,陈明玥.基于加权关系卷积网络和辅助任务的方面级情感分析[J].科学技术与工程,2025,25(14):5967-5975.

Yu Bengong, Chen Mingyue. Aspect-level sentiment analysis based on weighted relational convolutional networks and auxiliary task[J]. Science Technology and Engineering, 2025, 25(14): 5967-5975.

基于加权关系卷积网络和辅助任务的方面级情感分析

余本功^{1,2}, 陈明玥¹

(1. 合肥工业大学管理学院, 合肥 230009; 2. 合肥工业大学过程优化与智能决策教育部重点实验室, 合肥 230009)

摘要 方面级情感分析从细粒度角度检测了给定方面术语的情感极性从挖掘文本方面的情感态度的角度给电商、消费者等群体提供决策辅助。现有方法对句法信息的不同依赖关系的平等建模,忽略了卷积网络中不同依赖关系和语义关系中的全局信息对方面词情感分类的影响。为了解决上述问题,考虑到图结构的灵活性、复杂性和辅助任务在基于全局语义信息捕捉方面情感、补全细粒度方面信息的优良表现,提出了包含加权关系卷积网络(weighted relational convolutional networks, WRCN)和辅助任务互补学习(complementary task, CL)两个任务的 WRCN-CL(weight relational convolutional networks and complementary task, WRCN-CL)模型。该模型利用 Bi-LSTM(bi-directional long short-term memory, Bi-LSTM)提取文本特征;之后分别输入 WRCN 和 CL 中;CL 从全局角度寻找方面相关的语义信息以达到增强知识的效果,WRCN 基于 CL 中的方面表示结合利用 GCN(graph convolutional networks)深入挖掘基于位置信息和关系类型的句法信息,最后将融合全局信息和局部信息的特征输入到池化层,得到全面的信息表征以提高模型的分类效果。实验结果表明:相较于其他基线模型,WRCN-CL 在三个公开数据集上准确率分别达到了 83.49%、78.19%、75.89%,从而证明了本文模型能有效地解决方面级情感分析分类任务。

关键词 方面级情感分析;关系类型;位置赋权;互补学习;图卷积网络

中图分类号 TP399; 文献标志码 A

Aspect-Level Sentiment Analysis Based on Weighted Relational Convolutional Networks and Auxiliary Task

YU Ben-gong^{1,2}, CHEN Ming-yue¹

(1. School of Management, Hefei University of Technology, Hefei 230009, China; 2. Key Laboratory of Process Optimization and Intelligent Decision Making, Ministry of Education, Hefei University of Technology, Hefei 230009, China)

[Abstract] Aspect-level sentiment analysis detects the sentiment polarity of given aspect terms from a fine-grained perspective, providing decision support for e-commerce, consumers, and other groups by mining textual aspect sentiment. Different syntactic dependencies were treated equally in existing methods resulting in the influence of relation types in convolutional networks and the global information from semantic perspective being overlooked. To address these issues, considering the flexibility and complexity of graph structures, the excellent performance of auxiliary tasks in capturing aspect sentiment based on global semantic information and completing fine-grained aspect information, the model WRCN-CL (weight relational convolutional networks and complementary task) which incorporates two tasks: WRCN (weighted relational convolutional networks) and CL (complementary learning) was proposed. Specifically, Bi-LSTM (bidirectional long short-term memory network) was used to extract textual features, which were entered into WRCN and CL tasks separately. Aspect-related semantic information was collected from a global perspective to enhance knowledge, while the aspect representations from CL combined with GCN (graph convolutional networks) to deeply explore syntactic information based on positional and type-aware relational information in WRCN. The fused global and local features were then input into a pooling layer to obtain comprehensive information representation for improved classification performance. Experimental results demonstrate significant improvements with the accuracy of 83.49%, 78.19%, 75.89% on three public datasets compared to baseline models, proving the effectiveness of the proposed model in aspect-level sentiment analysis classification task.

[Keywords] aspect-level sentiment analysis; type-aware; positional weighting; complementary learning; graph convolutional networks

收稿日期: 2024-05-17; 修订日期: 2025-02-13

基金项目: 国家自然科学基金(72071061)

第一作者: 余本功(1971—),男,汉族,安徽合肥人,博士,教授。研究方向:信息系统、自然语言处理。E-mail: bgyu@hfut.edu.cn。

投稿网址: www.stae.com.cn

基于方面的情感分析 (aspect-based sentiment analysis, ABSA) 旨在分析一个句子中包含多个方面且分别具有相同或不同情感极性, 并对方面的情感极性做出判断。随着移动网络的迅速发展, 基于用户反馈的个人数字社交平台得到广泛使用。用户评价、看法、态度等表达主观感受的用户评论数据也因此开始大量产生并得到广泛收集。其中的用户需求是最具价值的反馈信息, 其重要程度也在电商企业中逐年上升。方面级情感分析也因此被视为一项极具价值的研究任务。

目前的研究关注到语义信息和语法信息的交互必要性并提出了基于注意力机制、CHATGPT 等多种类型的方法应用在方面级情感分析任务中以更好地服务于社会管理^[1]。先前的研究利用注意力机制来获取基于上下文信息的方面词的观点。Phan 等^[2]结合词性、依赖关系和上下文文本信息利用自注意力机制来增强方面词提取并提高了分类精度。考虑到注意力机制聚焦于细粒度文本信息, 相关学者对方面相关的文本位置信息提出更近一步的研究方法。Liang 等^[3]为了精简计算, 在 BETR 预训练过程中利用逻辑向量寻找和方面空间距离相关的特征表示, 使得在降低计算复杂度的同时保证计算精度。因此, 方面词的相对距离信息可以作为一个有用的特征来强化重要信息, 减少噪声的影响。

随着 ABSA 上下文编码能力的进一步开发, 相关工作依赖 GCN 以学习对 ABSA 有帮助的词-词关系。Tian 等^[4]设计了结合了注意力机制和依赖树, 然后通过层级卷积全面学习句法信息。鲍小异等^[5]通过构建跨语言的 GCN 卷积图来预测属性级别的方面词情感。Jiang 等^[6]基于共现图和依赖图进行编码以提高分类准确性。可以发现先前的研究主要利用关系节点上基于方面词的依赖关系进行平等建模以挖掘词与词之间的全局关系。Chen 等^[7]设计了多通道图卷积网络模型包括语篇、句法依赖、方面词距离和相对位置距离多方面学习表示。Chen 等^[8]利用门控机制动态结合来自词依赖图和潜在图的信息从而提高产品和关系之间的关系精度。Zhang 等^[9]利用句法图和语义图构建了一个融合的多通道卷积网络来捕捉情感极性。Xing 等^[10]利用门控机制迭代从而动态捕捉局部信息, 并利用外部知识补全全局知识, 通过结合全局信息和局部信息完善句法结构以达到提高 ABSA 任务效果。Saifuddin 等^[11]从序列数据的复杂性和结构相似性出发, 采用了两级注意力机制和 GCN 设计了超图注意力网络模型。陈钰佳等^[12]从双通道的角度用 GCN 提高了分类效果。

部分学者注意到大多数研究在对句法依赖树和多重关系图进行分析时未能有效地将语义或其他细粒度信息整合在一起, 阻碍了多种信息元素的交流。而辅助任务通过共享表征捕捉辅助任务与主任务的相关信息提高了模型的效果与泛化能力。为了进一步区分句法依赖关系, Yuan 等^[13]提出了将依赖类型的外部知识也纳入了句法中并通过图注意力机制实现分类。Lu 等^[14]使用异构神经网络利用依赖树、先验情感字典等多通道编码句序信息, 以实现参数共享。Gu 等^[15]融合外部知识和外部情感词库为句子中构建情感分值矩阵以进行图卷积从而捕捉边缘标签的缺陷。Zhang 等^[16]设计了基于外部知识的依赖胶囊网络结构框架提高了模型性能。张涵等^[17]提出了利用持续学习在低资源背景下通过多段多任务的互补学习方式提高学习效果。Qin 等^[18]提出利用辅助任务通过构建解释器迫使模型进行两阶段训练从而使得训练目标聚焦于 ABSA 任务本身。Zhao 等^[19]将先验知识在微调阶段整合情感知识作为第一个辅助任务, 不同标签组合进行分类预测作为第二个辅助任务, 并在 ABSA 任务上获得优良表现。

尽管之前的研究都获得了很大的成功, 但是还存在两个方面值得深入研究。

(1) 关键节点权重偏差。基于依赖树的 GCN, 依赖树的质量很大程度上决定了最后的结果。然而, 不可避免的是在卷积过程中, 依赖树中一些不正确或可忽视的连接卷积后可能会给任务增加噪声。因此, 在卷积过程中对不同的语法结构和单词类型应该给予不同程度的重视。然而, 这对于常规的依赖树来说是一项具有挑战的任务。例如, 针对评论“这家服务并没有价格那样好。”中存在的两个方面词“服务”和“价格”, 方面词“价格”和观点词“好”是相关的。然而, 在依赖树中, “好”被视为是描述方面词“服务”。对于另一个词“价格”由于固定搭配得到的是观点词“没有”。噪声词“没有”导致了对其错误的情感预测。可以看出, 复杂的语义结构和常规依赖树解码的局限性会导致对方面词“服务”的错误情感预测。因此, 由于复杂的语义结构和解析树解码能力的限制会对分类结果产生负面影响, 所提出的模型构建了一个基于位置关系和关系类型以调整依赖树生成卷积图从而辅助 GCN 卷积过程来解决上述问题。

(2) 缺乏全局信息对局部信息的细粒度补充。GCN 的性能更侧重于局部信息。而全局信息是把握说话者的整体含义从而有效避免局部混淆带来的语义歧义等不良影响, 对情感的正确分类有着重

要的作用。因此,通过考虑全局信息有效弥补 GCN 聚焦于局部信息的不足,CL 被提出以帮助句子把握全局信息的语义信息的影响度从而获得全局信息对局部信息的细粒度补充。

针对上述问题,现提出基于加权关系卷积网络和辅助任务互补学习模型。采用 Bi-LSTM 提取文本特征;并输入 CL 和 GCN 中深入挖掘细粒度全局信息和局部信息,最终融合信息得到全面的信息表征以提高模型性能。在多个公开的方面及情感分析任务的数据上证明本文模型的性能超过现有的基准模型,具有优秀的情感挖掘、情感判断能力。通过利用这个模型,可以有效帮助企业判断用户对商品的态度以及用户间对企业情感倾向,从而帮助企业有效利用物力资源抢占市场份额,辅助用户管理决策。

1 WRCN-CL 模型构建

本文模型 WRCN-CL 如图 1 所示,它包含方面词位置和关系类型的加权 GCN 和辅助补充学习。模型进行词嵌入后利用 Bi-LSTM 提取文本特征,得到的特征分别输入基于依赖关系类型和方面词位置的依赖树模块和补充学习中,后将学习到的基于方面词的全局信息和句法信息共同输入到 GCN 中

提取包含交互信息下的隐层关系以得到最后的情感预测结果。

对任务的训练集 $\Omega = \{(s_c, \alpha, y_i)\}_{i=1}^N$, $s_c = \{s_1, \dots, s_{t+1}, \dots, s_{t+m}, \dots, s_k\}$ 表示原始句子, $\alpha = \{s_{t+1}, s_{t+2}, \dots, s_{t+m}\}$ 表示长度为 m 的单词的方面词在句中的位置, y_i 为方面词的情感极性。

1.1 特征提取

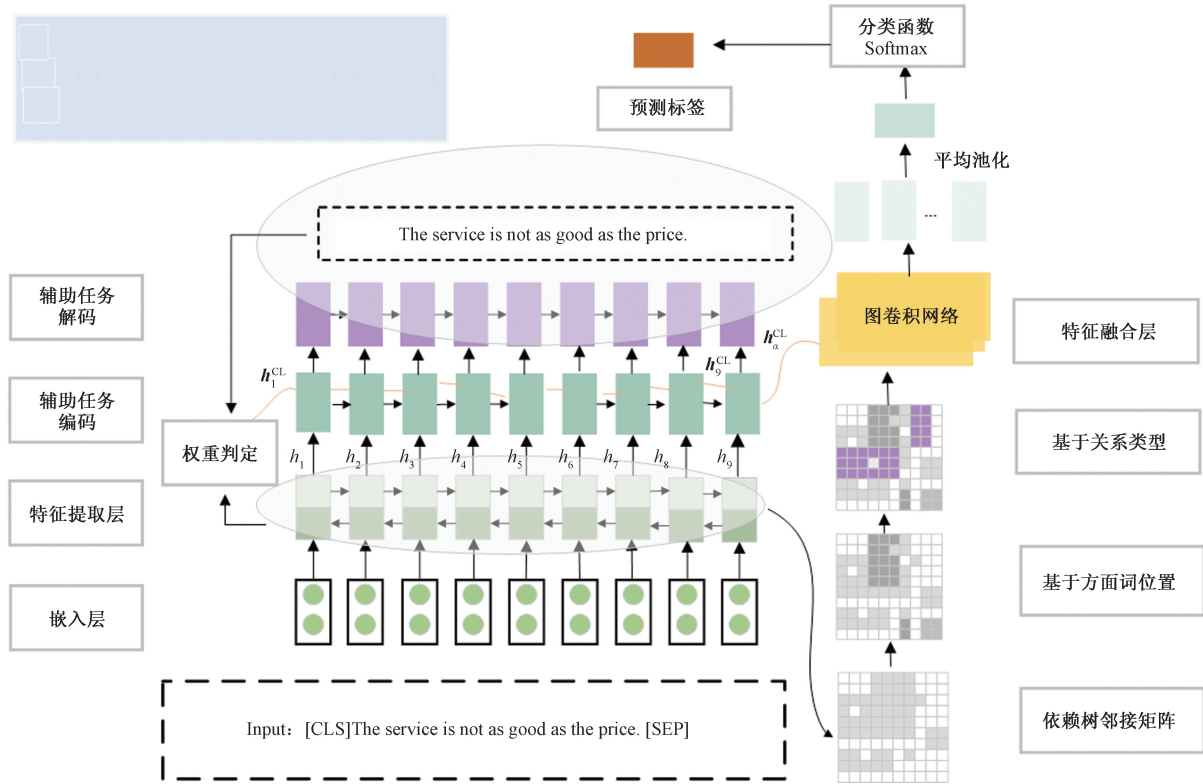
LSTM 在处理长文本时利用记忆门、遗忘门、更新门有选择地对远距离文本捕捉包含上下文文本的信息。Bi-LSTM 模型由前向和后向 LSTM 组成,通过考虑双向上下文信息提高语义信息。

因此,在利用预先训练模型 (BERT^[20] 或 GloVe^[21]) 得到词嵌入 $\hat{h}_i = \{\hat{h}_1, \hat{h}_2, \dots, \hat{h}_k\} \in \mathbf{R}^k \times D_{emb}$, 其中 D_{emb} 为词嵌入维度。随后,将每个词的词嵌入 \hat{h}_i 将输入到 Bi-LSTM 得到包含上下文文本信息的隐向量 $h_i = \text{Bi-LSTM}(\hat{h}_i; W_b) \in \mathbf{R}^{k \times 2D}$, 其中 W_b 为偏差项。

然后,这些文本特征将会输入两个不同的编码器得到相应的表示:①基于位置关系和依赖类型的依赖树;②基于补充学习的辅助任务。

1.2 辅助任务补充学习

这个模块针对学习方面词全局信息的学习,包



h_i 为上下文文本信息的隐向量; h_i^{CL} 为经过辅助任务编码器后得到的隐向量; h_i^{CL} 为训练后的方面词特征

图 1 模型整体框架

Fig. 1 The overall framework of the model

括全局语义信息学习(辅助任务编码和解码),辅助任务权重判定两个部分。

1.2.1 全局语义信息学习

首先将得到的包上下文文本信息的特征向量 $\mathbf{h}_i = \{h_1, h_2, \dots, h_k\}$ 输入 CL 编码器和解码器中。

$$\mathbf{h}_i^{\text{CL}} = \text{GRU}(s_i, \boldsymbol{\alpha}) \quad (1)$$

$$\hat{y}_i^{\text{CL}} = \text{GRU}(h_1^{\text{CL}}, h_2^{\text{CL}}, \dots, h_{t-1}^{\text{CL}}, s_1, \dots, s_{t-1}) \quad (2)$$

$$\mathbf{h}_\alpha^{\text{CL}} = \text{maxpooling}(\mathbf{h}_i^{\text{CL}} | \boldsymbol{\alpha}) \quad (3)$$

式中: \mathbf{h}_i^{CL} 为经过 GRU 后得到的隐向量; s_{t-1} 为在 $t-1$ 时刻的重构词; \hat{y}_i^{CL} 为在 t 时刻由解码器生成的单词; $\mathbf{h}_\alpha^{\text{CL}}$ 为训练后的方面词特征,并作为 GCN 的输入之一,用于指导在预测模型中对给定方面词的情感极性。

值得注意的是,在解码过程中序列不断通过之前学习到的词语作为后一个生成词中,也就意味着生成句中包含丰富的全局语义信息。

因此,该部分的损失函数表示为

$$\ell_{\text{CL}} = - \sum_{t=1}^k \ln p(s_t | s_1 \dots s_{t-1}) \quad (4)$$

式(4)中: $p(s_t | s_1 \dots s_{t-1})$ 为对 ground-truth 的预测概率。

1.2.2 辅助任务权重判定

由于 CL 模块作为一个辅助任务,其主要目的是为了对预测任务提供全局语义信息的细粒度补充。

具体来说,通过将原句的原始向量与生成句的向量进行比对。如果 ground-truth 正确比重比较大,则说明生成句的准确性高,即全局信息的补充是正确且积极的,否则生成句比重需要降低以减少负面影响。

$$y_i^{\text{D}^*} = \begin{cases} 0, & \hat{y}_i^{\text{CL}} \neq s_i \\ 1, & \hat{y}_i^{\text{CL}} = s_i \end{cases} \quad (5)$$

$$o_i^{\text{D}} = \text{softmax}(\mathbf{W}_{\text{D}} \mathbf{h}_i + \mathbf{b}_{\text{D}}) \quad (6)$$

式中: y_i^{CL} 为 CL 解码器在 i 时刻生成的单词; o_i^{D} 为对在 i 时刻正确性的判定; \mathbf{W}_{D} 和 \mathbf{b}_{D} 分别为可训练的矩阵和偏差项。

因此,对于权重判定部分的损失函数为

$$\ell_{\text{D}} = - \sum_{i=1}^k \ln p(y_i^{\text{D}^*} | s_i) \quad (7)$$

1.3 基于依赖关系和方面词位置的依赖树

从 Bi-LSTM 提取的特征会因为不合理的断句产生干扰,而先前的研究证明方面词相关的文本信息和语义结构在正确预测情感中占有很大比重。因此,利用位置关系和依赖关系对依赖树语义结构进行调整能更好地提取原始特征。

1.3.1 构建依赖树

首先,通过现有的分析工具得到原始依赖树的邻接矩阵 $\mathbf{D}_{n,m} \in \mathbf{R}^{k \times k}$ 。

$$\mathbf{D}_{n,m} = \begin{cases} 1, & n = m, \text{ 或 } s_n, s_m \text{ 存在语义关系} \\ 0, & \text{其他} \end{cases} \quad (8)$$

2.3.2 基于依赖关系和方面词位置重构依赖树

原始输入文本方面词位置关系可以表示为 $\mathbf{s}_c = [s_1, s_2, \dots, s_{\tau+1-\lambda}, \dots, s_{\tau+1}, \dots, s_{\tau+m}, \dots, s_{\tau+m+\lambda}, \dots, s_k]$,通过判断单词到方面中心部分的距离来确定语义相关性(semantic relevance distance, SRD)。

$$\text{SRD}_k = \|k - P_a\| - \frac{m}{2} \quad (9)$$

式(9)中: k 和 P_a 分别为文本单词和方面词中心。

$$\mathbf{W}_{n,m} = \begin{cases} 1, & \text{SRD}_n \leq \lambda \cap \text{SRD}_m \geq \lambda \\ \frac{1 - \|n - P_a\| + \|m - P_a\|}{2k}, & \\ \text{其他} & \end{cases} \quad (10)$$

$$\mathbf{G}_{m,n} = \mathbf{D}_{m,n} \mathbf{W}_{m,n} \quad (11)$$

式中: $\mathbf{W}_{n,m}$ 为基于方面词中心距离构建的权重矩阵; $\mathbf{G}_{m,n}$ 为现在依赖树的权重矩阵。

值得注意的是,上述构建句子的依赖树对不同关系类型的重视程度相同,而词组等关系类型在日常交流中更容易影响情感取向的判断。因此,根据不同关系类型给定不同权重来调整依赖树结构是必要的。

对于 L 层的神经网络中,单词 s_m 和 s_n 构造的边在存在着一种依赖关系 $\mathbf{c}_{m,n}^r$,即: \mathbf{h}_m^{L-1} 和 \mathbf{h}_n^{L-1} 有 $\mathbf{c}_{m,n}^r$,这种关系类型对结果会产生影响。

$$\mathbf{g}_n^L = [\mathbf{h}_n^{L-1}; \mathbf{c}_{m,n}^r] \quad (12)$$

$$\mathbf{g}_m^L = [\mathbf{h}_m^{L-1}; \mathbf{c}_{m,n}^r] \quad (13)$$

$$\mathbf{w}_{m,n}^L = \frac{a_{m,n} \exp(\mathbf{g}_m^L \mathbf{g}_n^L)}{\sum_{z=1}^k a_{m,n} \exp(\mathbf{g}_m^L \mathbf{g}_z^L)} \quad (14)$$

式中: \mathbf{g}_n^L 为节点 n 在前一层 GCN 的输出结果后考虑位置信息和依赖类型后的隐含状态; \mathbf{g}_m^L 为节点 m 在前一层 GCN 的输出结果后考虑位置信息和依赖类型后的隐含状态, $[\cdot; \cdot]$ 向量的拼接,通过式(12)和式(13)可以得到基于依赖关系类型的特征表示; $a_{m,n}$ 为 s_m, s_n 的邻接矩阵的数值; z 为节点 m 与节点 n 之间的关系对数量; $\mathbf{w}_{m,n}^L$ 为考虑依赖关系类型后的权重。

1.4 特征融合

基于之前辅助特征和语法结构,辅助特征视为 $\mathbf{h}_0 = \mathbf{h}_\alpha^{\text{CL}}$ 并将基于关系类型的语义结构一同 GCN 的输入从而得到最后的结果。

$$\mathbf{h}_m^{(L-1)*} = \mathbf{h}_n^{L-1} + \mathbf{W}_r \mathbf{c}_{m,n}^r \quad (15)$$

$$\mathbf{h}_n^L = \sigma \left\{ \sum_{m=1}^k \mathbf{w}_{m,n}^L [\mathbf{W}^L \mathbf{h}_m^{(L-1)*} + \mathbf{b}^L] \right\} \quad (16)$$

式中: \mathbf{W}^L 和 \mathbf{b}^L 分别为图卷积中可训练矩阵和偏差项; $\sigma(\cdot)$ 为激活函数。

此处选取 GRU 作为激活函数,通过在不同层数的 GCN 中进行卷积,越重要的关系类型将会在卷积过程中权重更大,从而有效地提高情感分类结果。

通过关注关系类型和中心词位置来完善语法结构,减少冗余;利用辅助关系学习语义信息,从全局角度进行细粒度补充,从而得到最终的情感分类预测结果。

$$\mathbf{o}_{SA} = \text{softmax}(\mathbf{W}_{SA} \mathbf{h}^{SA} + \mathbf{b}_{SA}) \quad (17)$$

$$\mathbf{h}^{SA} = \text{average pooling}(\mathbf{h}_n^L) \quad (18)$$

式中: \mathbf{W}_{SA} 和 \mathbf{b}_{SA} 分别为模型预测情感时需要训练的参数矩阵和偏差项; \mathbf{o}_{SA} 为最后的输出结果。

1.5 损失函数

模型最后采用对数函数进行损失计算。

$$\ell_{SA} = -\ln p(y^* | \mathbf{h}_{i1} \cdots \mathbf{h}_{ik}) \quad (19)$$

$$\ell = \ell_{CL} \ell_D + \ell_{SA} \quad (20)$$

式中: ℓ_{SA} 为进行方面级情感分类主任务时的损失函数; ℓ_{CL} 为增强情感分类效果的辅助任务的损失函数; ℓ_D 为对辅助任务权重判定的损失函数; y^* 为最后的情感预测结果。

式(20)意味着 Bi-LSTM 的输出 \hat{h}_i 和重构句的判定 y_i^{D*} 都将被作为指标输入到第二阶段中以进行微调。当预测词 \hat{h}_i 和判定词 y_i^{D*} 都为 1,即 ℓ_{CL} 和 ℓ_D 都很小,意味着重构句子的语法正确,且在模型在前向传播时能够很好地学习到上下文的语义知识。但当两者都为 0 的时候,意味着重构句和原句的语义关系不大。因此, ℓ_{CL} 就会很大。当两者不等时,意味着并没有完全学习到上下文信息。因此, ℓ_D 会很大,从而提高辅助任务的比重。

2 实验方法

2.1 数据集

如表 1 所示,实验基于 3 个包含多个方面情感分析的公开数据集(即 Restaurant-14、Laptop、Twitter)与基准模型进行比较以评价模型性能。数据集的结构为:token、pos、head、deprel、aspects。

2.2 参数设置

本次实验基于 Pytorch 实现,编程语言 python3.9 并通过 NVIDIA RTX3090 来训练和测试模型。经过多次实验,最终选择 5 次结果平均值获得结果。分别使用 Glove 预训练 300 维词向量和 BERT 对文本数据

嵌入化,Bi-LSTM 的隐藏维度设置为 300。GCN 的层数为 2 层,隐藏层维度为 300。选择 Adam 优化器函数进行学习,设置模型的 dropout 为 0.3,单词距离 $\lambda=7$ 。句法依赖树是利用工具 SpaCy toolkit 获得。当使用 BERT 预训练时,batch size 和 epoch 分别设为 32 和 10。实验参数如表 2 所示。

表 1 数据集概况

Table 1 Dataset statistics

数据集		积极情感数量	中性情感数量	消极情感数量
Restaurant	训练集	2164	637	807
	测试集	728	196	196
Laptop	训练集	994	464	870
	测试集	341	169	128
Twitter	训练集	1 561	3 127	1 560
	测试集	173	346	173

表 2 实验参数

Table 2 Experimental hyper parameters

参数名称	参数值
Glove 词向量	300
Bi-LSTM 维度	300
BERT 词向量	768
最大句长	75
批次长度	32
丢失率	0.3
学习率	2×10^{-5}
数据迭代次数	10
优化器	Adam
GCN 层数	2
单词距离 λ	7

3 实验结果与分析

3.1 基线模型

通过选择在同一任务下对相同数据集进行测试的基线模型进行对比以更好地评估模型性能。

3.1.1 基于 GloVe 的基线模型

IAN^[22]:该模型采用动态交互网络来分配方面词上下文信息的权重。

R-GAT^[23]:该模型根据特定方面重构和修剪依赖树以获得更好的依赖结构。

ASGCN^[24]:该模型从语义角度和语法角度双通道加强分类效果。

SKGCN^[25]:该模型通过 GCN 同时考虑依赖关系、语法和外部知识。

CRF-GCN^[26]:该模型基于条件随机场获得方面词的观点表示并将信息在 GCN 中聚合。

PSKE-GCN^[27]:该模型基于特定方面词裁剪依赖树并结合外部知识捕捉深层语法表示。

RMN-P^[28]:该模型从基于多任务学习区分方面

词和相近词间的语义关系从而提高模型分类效果。

Bi-GCN^[29]:该模型通过 GCN 学习特定方面的表征,解决长距离多词依赖问题。

KGCapsAN-LI^[16]:该模型将先验知识输入胶囊网络中从而得到基于特定方面词的特定表示。

3.1.2 基于 BERT 的基线模型

ACLT^[30]:该模型利用注意力机制对依赖树的边动态赋权并进行对裁剪依赖树。

RGAT-SPC-BERT^[31]:该模型使用 BERT 作为预训练模型获取其嵌入并利用双向 Transformer-like 编码器获取上下文的语义。

MOS-GCN-BERT^[32]:该模型基于注意力机制的裁剪依赖树 GCN 模型。

4.2 性能分析

表 3 列出了基于 3 个数据集的实验结果和比较。准确率(A_{cc})和 Marco- F_1 (F_1)被用作评价指标并根据评价指标确定所提出方法的性能。以 Twitter 数据集为例,所提出的模型基于 Glove 准确率提高 0.32%, F_1 提高 0.71%; 基于 BERT 准确率提高 1.37%, F_1 提高 1.06%。这说明:①通过关注方面词相关信息和句法结构能够有效地提高了该任务的分类能力;②由于 BERT 预训练语言模型对于分析句子结构,捕捉句间关系具有更好效果,从而针对句子结构的模型在其上具有更突出的性能表现。基于以上分析,可以得出在 3 个数据集上的评价指标准确率和 F_1 结果优于绝大多数基线方法,从而证明了所提出模型的有效性 with 优越性。

表 3 基线模型对比

Table 3 Comparison of baseline models

模型	Restaurant		Laptop		Twitter	
	$A_{cc}/\%$	$F_1/\%$	$A_{cc}/\%$	$F_1/\%$	$A_{cc}/\%$	$F_1/\%$
IAN	79.26	70.09	72.05	67.38	72.50	70.81
RGAT	83.30	76.08	77.42	73.76	75.57	73.82
ASGCN	81.73	73.10	72.62	66.72	71.05	69.45
SKGCN	80.36	70.43	73.20	69.18	71.97	70.22
CRF-GCN	82.71	73.87	75.83	74.78	—	—
PSKE-GCN	83.21	75.72	—	—	74.28	72.86
RMN-P	81.16	73.17	74.50	69.79	—	—
Bi-GCN	81.97	73.48	74.59	71.84	74.16	73.35
KGCapsAN-LI	82.49	74.21	77.02	72.97	74.57	72.74
Ours-BiLSTM	83.49	75.59	78.19	73.86	75.89	74.53
ACLT	85.71	78.44	79.68	75.83	75.48	74.51
RGAT-SPC-BERT	86.60	77.22	78.21	74.07	76.15	74.88
BMGCN-BERT	85.96	80.08	79.91	75.66	76.16	74.59
Ours-BERT	85.44	81.29	80.29	76.59	77.53	75.94

注:加粗数值表示在该评价指标下的对比模型中的最优数值。

3.3 消融实验与分析

进一步探讨模型单一模块的影响,如表 4 所示。其中,符号 w/o 表示没有该部分。第一列是抹去每个部分的分类,每一行代表基于第一列所列模块的结果。w/o WRCN 表示删除基于位置权重和关系类型的 GCN 后模型效果,w/o SRD 表示 GCN 的邻接矩阵未考虑位置关系对模型的影响,w/o 关系类型表示在 WRCN 中将构图部分删除关系类型的影响,w/o CL 表示如果不考虑全局信息的细粒度补充条件下模型效果。

通过分析每个模块的变化程度,可以发现每个模块的各自发挥着重要的作用。以 Restaurant 数据集为例,可以得出如下结果。

(1) 去掉 WRCN 模块后模型准确率由原来的 83.49% 下降至 79.26%, 下降了 4.23 个百分点; F_1 由原来的 75.59% 下降至 74.09%, 下降了 1.5 个百分点。这说明对于 ABSA 任务来说,以位置权重和关系类型的 GCN 发挥着重要的作用,对模型准确率影响极大。

(2) 为了进一步分析 GCN 中各模块的作用,设计了删除 SRD 和删除关系类型的消融实验,可以发现这两个消融实验都会导致模型的准确率和 F_1 下降,但是去掉 SRD 的消融实验的准确率和 F_1 比去掉关系类型下降更快。这说明虽然位置关系和关系类型在模型中都发挥着重要的作用,但是 SRD 在 GCN 模块中发挥着更重要的作用且对句中的单词位置权重进行调整对后续基于关系类型的 GCN 来说是必要的。

(3) 去掉 CL 模块后模型的准确率由原来的 83.49% 降至 82.86%, 下降了 0.63 个百分点。这说明全局信息和局部信息的整合是有效的,通过挖掘方面词相关的局部信息,从语义结构的角度捕捉方面词情感为 ABSA 任务提供辅助支撑能有效地提升模型效果。

表 4 消融实验结果

Table 4 Results of ablation experiments

模型	Restaurant		Laptop		Twitter	
	$A_{cc}/\%$	$F_1/\%$	$A_{cc}/\%$	$F_1/\%$	$A_{cc}/\%$	$F_1/\%$
WRCN-CL	83.49	75.59	78.19	73.86	75.89	74.53
w/o WRCN	79.26	74.09	74.05	72.38	73.50	71.81
w/o SRD	81.87	74.64	76.23	73.82	76.21	72.88
w/o 关系类型	82.82	75.49	76.36	73.83	76.53	73.11
w/o CL	82.86	75.33	76.80	74.10	76.79	72.81

3.4 灵敏度分析

对 SRD 模块中参数 λ 对模型的性能影响分析和 GCN 层数 L 对模型性能的影响分析。

(1) SRD 模块中参数 λ 对模型的影响分析:基

于不同数据集数据特征不同,局部上下文范围 λ 需要不断调整来获得适合的参数结果。因此,在保证只有温度系数改变的情况下,在 1~10 对模型在 3 个数据集上分别进行实验以确定参数。

参数 λ 的可视化结果如图 2 所示。随着 λ 发生变化,模型的总体性能呈现着先上升再下降的趋势。同时,从图 2(a) 中可以直观地看出,针对 F_1 , 模型实验效果在 $\lambda = 7$ 时在 Twitter 和 Laptop 数据集中达到最佳,在 Restaurant 数据集上 $\lambda = 6$ 的效果最好。这意味着在 $\lambda = 6$ 或 7 时模型能够在这个范围中获得更多的有效信息从而更好识别方面词从而提升模型效果。接着,从图 2(b) 中可以观察到 3 个数据集都在 $\lambda = 7$ 时实验效果达到顶峰。并且在 $\lambda = 6$ 到 $\lambda = 7$ 的过程中,Restaurant 数据集上模型评价指标 F_1 差距较小,而准确率差距较大,即在 $\lambda = 7$ 时模型效果得到明显提升。

综合以上分析,本文模型选取 $\lambda = 7$ 以得到最好的分类效果。

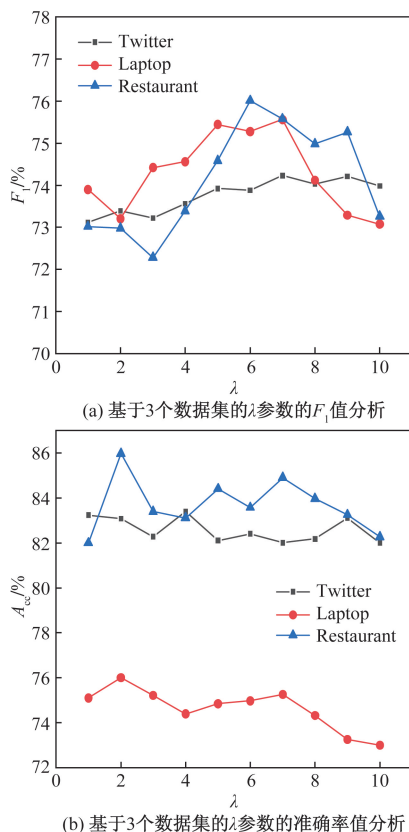


图2 SRD 中 λ 的灵敏度分析

Fig. 2 The sensitivity analysis of λ in SRD

(2) GCN 层数讨论: 保证其他参数不变的条件 下,观察随着 GCN 层数 L 在 1~10 层中增加,模型 的变化效果。

关于不同数据集上层数变化对模型性能的影响结果如图 3 所示。可以看出 3 个数据集中,模型

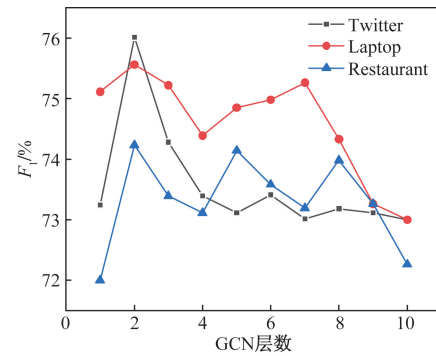


图3 GCN 层数分析

Fig. 3 The analysis of the layer of GCN

的整体性能呈现着先上升后下降的一个趋势,并且在层数为 1~5 时这种变化尤为明显。其次,在 3 个数据集中都是当 GCN 层数为 2 时,模型效果最佳。这说明此时模型能很好地聚合附近的语义关系,且并不会引入噪声,而过多的局部信息在数据量有限的背景下会降低分类效果。

4 结论

提出了一种新的模型 WRCN-CL,通过将补充学习与类型感知的方面位置依赖树相结合利用 GCN 来判断句中的方面词情感,并解决了依赖树中关系类型平等建模和缺乏全局信息的问题。具体来说:首先,设计了基于位置关系和关系类型的依赖树结构,得到加权关系卷积网络。旨在考虑到不同距离的信息差异和依赖关系重要性差异,利用相对位置距离和依赖关系类型,从语法结构上调整常规依赖树结构。这样不仅从语义角度充分捕捉方面词信息,还能通过纠正错误的依赖关系获得有效地细粒度信息,从而达到更好的分类效果。其次,构建了基于全局信息的 CL 任务以通过知识增强的手段捕获方面词相关的语法信息。这样不仅能理解-重构句子学习全局信息并利用判别器获得最终情感融合的权重。最后,将其和辅助任务一起输入 GCN 中完成最后的分类,得到方面词的情感倾向,并在实验中取得了不错的效果。在未来的工作中,多模态情感分析将会进一步发展,从而提高分类的准确性并应用到实际中。

参 考 文 献

- [1] 曾桢,王擎宇. 融合 BERT 中间隐藏层的方面级情感分析模型 [J]. 科学技术与工程, 2023, 23(12): 5161-5169.
Zeng Zhen, Wang Qingyu. Aspect-based sentiment analysis model incorporating the BERT intermediate hidden layer [J]. Science Technology and Engineering, 2023, 23(12): 5161-5169.
- [2] 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. New Orleans: ACL, 2020: 3211-3220.
- [3] 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. DOI: 10.1016/j.knosys.2021.107643.
- [4] Tian Y, Chen G, Song Y. Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble [C]// Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. New York: NAACL, 2021: 2910-2922.
- [5] 鲍小异, 姜晓彤, 王中卿, 等. 基于跨语言图神经网络模型的属性级情感分类 [J]. 软件学报, 2023, 34(2): 676-689.
Bao Xiaoyi, Jiang Xiaotong, Wang Zhongqing, et al. Cross-lingual aspect-level sentiment classification with graph neural network [J]. Journal of Software, 2023, 34(2): 676-689.
- [6] Jiang B, Xu G, Liu P. Aspect-level sentiment classification *via* location enhanced aspect-merged graph convolutional networks [J]. The Journal of Supercomputing, 2023, 79(9): 9666-9691.
- [7] Chen H, Zhai Z, Feng F, et al. Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction [C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Dublin: ACL, 2022: 2974-2985.
- [8] Chenhua Chen, Zhiyang Teng, Yue Zhang. Inducing target-specific latent structures for aspect sentiment classification [C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Seattle: Association for Computational Linguistics, 2020: 5596-5607.
- [9] Zhang Z, Zhou Z, Wang Y. SSEGCN: 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. Washington: NAACL, 2022: 4916-4925.
- [10] Xing B, Tsang I W. Understand me, if you refer to aspect knowledge: Knowledge-aware gated recurrent memory network [J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2022, 6(5): 1092-1102.
- [11] Saifuddin K M, May C, Tanvir F, et al. Seq-Hygan: sequence classification *via* hypergraph attention network [C]// Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. Birmingham: CIKM, 2023: 2167-2177.
- [12] 陈钰佳, 郑更生, 肖伟. 基于 RoBERTa 与句法信息的中文影评情感分析 [J]. 科学技术与工程, 2023, 23(18): 7844-7851.
Chen Yujia, Zheng Gengsheng, Xiao Wei. Sentiment analysis of Chinese film reviews based on RoBERTa and syntactic information [J]. Science Technology and Engineering, 2023, 23(18): 7844-7851.
- [13] Yuan L, Wang J, Yu L C, et al. Syntactic graph attention network for aspect-level sentiment analysis [J]. IEEE Transactions on Artificial Intelligence, 2022, 5(1): 140-153.
- [14] Lu G, Li J, Wei J. Aspect sentiment analysis with heterogeneous graph neural networks [J]. Information Processing & Management, 2022, 59(4): 102953.
- [15] Gu T, Zhao H, He Z, et al. Integrating external knowledge into aspect-based sentiment analysis using graph neuralnetwork [J]. Knowledge-Based Systems, 2023, 259: 110025.
- [16] Zhang B, Li X, Xu X, et al. Knowledge guided capsule attention network for aspect-based sentiment analysis [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 2538-2551.
- [17] 张涵, 王晶晶, 罗佳敏, 等. 针对低资源场景下连续情感分析任务的持续注意力建模 [J]. 软件学报, 2024, 35(12): 1-17.
Zhang Han, Jiang Xiaotong, Wang Jingjing, et al. Continual attention modeling for successive sentiment analysis in low-resource scenarios [J]. Journal of Software, 2024, 35(12): 1-17.
- [18] Qin H, Tian Y, Xia F, et al. Complementary learning of aspect terms for aspect-based sentiment analysis [C]// Proceedings of the Thirteenth Language Resources and Evaluation Conference. Marseille: European Language Resources Association, 2022: 7029-7039.
- [19] Zhao Q, Ma S, Ren S. KESA: a knowledge enhanced approach to sentiment analysis [C]// Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Dublin: Association for Computational Linguistics, 2022: 766-776.
- [20] Kenton J D M W C, Toutanova L K. Bert: pre-training of deep bidirectional transformers for language understanding [C]// Proceedings of NAACL-HLT. Minneapolis: Association for Computational Linguistics, 2019: 1-2.
- [21] Pennington J, Socher R, Manning C D. Glove: global vectors for word representation [C]// Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). Minneapolis: Association for Computational Linguistics, 2014: 1532-1543.
- [22] Ma D, Li S, Zhang X, et al. Interactive attention networks for aspect-level sentiment classification [C]// Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization. Melbourne: AI Journal, 2017: 4068-4074.
- [23] Wang K, Shen W, Yang 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. Seattle: Association for Computational Linguistics, 2020: 3229-3238.
- [24] 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). Hong Kong: Association for Computational Linguistics, 2019: 4568-4578.
- [25] 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]. Knowledge-Based Systems, 2020, 205. DOI: 10.1016/j.knosys.2020.106292.
- [26] Huang B, Zhang J, Ju J, et al. CRF-GCN: an effective syntactic dependency model for aspect-level sentiment analysis [J]. Knowledge-Based Systems, 2023, 260: 110125.
- [27] 韩虎, 孔博, 何勇禧, 等. 基于剪枝策略的知识增强方面级情感分析 [J]. 华中科技大学学报 (自然科学版), 2024, 52(11): 1-8.

- Han Hu, Kong Bo, He Yongxi, et al. Aspect based sentiment analysis based on knowledge enhancement of pruning strategy[J]. Journal of Huazhong University of Science and Technology(Natural Science Edition), 2024, 52(1): 1-8.
- [28] Zeng J, Liu T, Jia W, et al. Relation construction for aspect-level sentiment classification[J]. Information Sciences, 2022, 586: 209-223.
- [29] 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). Seattle: Association for Computational Linguistics, 2020; 3540-3549.
- [30] Zhou Y, Liao L, Gao Y, et al. To be closer: learning to link up aspects with opinions[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana: Association for Computational Linguistics, 2021; 3899-3909.
- [31] Xiao Z, Wu J, Chen Q, et al. BERT4GCN: using BERT intermediate layers to augment GCN for aspect-based sentiment classification[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana: Association for Computational Linguistics, 2021; 9193-9200.
- [32] 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]. (2024-07-29) [2025-02-15]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20240726.1350.004.html>.
- Mu Yiru, Han Hu, Kong Bo. Aspect-based sentiment analysis based on multiple ordinal scales of graph convolutional network[J/OL]. (2024-7-29) [2025-02-15]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20240726.1350.004.html>.