

面向方面级情感分析的双通道图卷积网络

刘洋,黎茂锋,黄俊,陈立伟

(西南科技大学 计算机科学与技术学院,四川 绵阳 621010)

摘要:当前大多数方面级情感分析通过在句法依赖树上构建图神经网络,从句法依赖树上聚合方面项与上下文之间的情感信息,缺乏对文本语义相关性以及文本中情感知识的挖掘。为此,提出一种融合语义信息和文本自身情感信息的双通道图卷积神经网络(Dual-Channel Graph Convolutional Networks,DC-GCN)模型,该模型基于具有上下文语义相信息的句法依赖图(Sentiment-relationship Graph,SrG)和情感依赖图(Sentiment Graph,SeG)构建。具体来说:首先,使用注意力机制捕获文本的语义相关性,并将其融入句法依赖,设计了一个具有上下文语义信息的句法依赖图(SrG);其次,整合SenticNet词典中单词的情感信息构建了一个文本情感依赖图(SeG);最后,分别基于SrG和SeG构建单通道图卷积神经网络,再将两个单通道图卷积神经网络融合成双通道图卷积神经网络,以融合文本的具有语义信息的句法特征和文本自身的情感特征。在四个公开数据集(Twitter,Rest14,Lap14和Rest16)上取得了比对比模型更好的效果,其中中文模型在Marco-F1值上分别高出对比模型0.95百分点、1.24百分点、0.62百分点和2.75百分点。

关键词:方面级情感分析;句法依赖;自注意力机制;情感信息;图卷积网络

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2024)03-0049-08

doi:10.3969/j.issn.1673-629X.2024.03.008

Aspect-based Sentiment Analysis via Dual-channel Graph Convolutional Network

LIU Yang, LI Mao-feng, HUANG Jun, CHEN Li-wei

(School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China)

Abstract:At present, most aspect-level sentiment analysis aggregates the emotional information between aspects and context from syntactic dependency trees by constructing graph neural networks on syntactic dependency trees, and lacks the mining of semantic correlation of text and emotional knowledge in text. To this end, we propose a dual-channel graph convolutional network (DC-GCN) model that integrates semantic information and text self-emotional information, which is built by the SrG (Sentiment-relationship Graph) and the SeG (Sentiment Graph) with contextual semantic phase information. Specifically: firstly, the attention mechanism is used to capture the semantic relevance of the text and integrate it into the syntactic dependency, and the SrG with contextual semantic information is designed; secondly, the emotional information of words in SenticNet dictionary is integrated to construct the SeG; finally, a single-channel graph convolutional neural network is constructed based on SrG and SeG, and then the two single-channel graph convolutional networks are fused into a two-channel graph convolutional neural network to fuse the syntactic features of the text with semantic information and the emotional characteristics of the text itself. The four public datasets of Twitter, Rest14, Lap14 and Rest16 achieved better results than that of the comparison model, and the Marco-F1 values of the proposed model were 0.95%, 1.24%, 0.62% and 2.75% higher than that of the comparison model, respectively.

Key words: aspect-based sentiment analysis; syntactic dependency; self-attention mechanism; sentiment information; graph convolutional network

0 引言

在web2.0的时代背景下,互联网用户能够通过社交媒体浏览并发布大量带有情感极性的文本信息,所以,意见挖掘、情感分析^[1]逐渐成为热门研究话题。

而方面级情感分析(ABSA)是一种更加细粒度的情感分析任务,它能够对文本具体的方面实体所携带的情感极性进行预测,而不是简单地预测整条文本的情感极性。以“The food is good, but the service is

收稿日期:2023-04-26

修回日期:2023-08-29

基金项目:四川省自然科学基金项目(2022NSFSC0940)

作者简介:刘洋(1996-),男,硕士,CCF会员(K6883G),研究方向为自然语言处理;通信作者:黎茂锋(1977-),男,硕士,讲师,CCF会员(18367M),研究方向为软件架构。

terrible.”为例, food 和 service 是文本的两个方面项, 两者的意见项分别是 good 和 terrible, ABSA 任务的目标就是识别出两个方面项的情感极性分别是积极和消极。

ABSA 任务关键在于建立方面项与意见项之间的联系, 以此来引导方面项感知情感信息。近年来, 以句法依赖为基础的图神经网络在 ABSA 任务上取得了显著成果。然而, 这些方法^[2-3]忽略了上下文语义信息和单词自身所携带的情感信息对建立方面项与意见项联系的作用。而自注意力机制能有效地捕获上下文的语义相关性。受此启发, 该文利用多头自注意力机制增强句法依赖图, 设计了一个具有上下文语义信息的句法依赖图(SrG); 整合 SenticNet 词典中单词的情感信息构建文本情感依赖图(SeG); 最后, 基于 SrG 和 SeG 两个图构建了一种双通道图卷积网络(DC-GCN), 以同时利用文本的语义相关性和单词自身的情感信息引导方面项感知情感信息, 并将融合双通道图卷积网络特征用于情感分类。为了验证模型的有效性, 在四个公开数据集(Twitter, Rest14, Lap14 和 Rest16)上与多个较优的经典模型以及最新模型进行对比, 实验结果表明, 提出的模型取得了更好的效果。

1 相关工作

方面级情感分析(ABSA)概念由 Thet 等人^[4]在 2010 年明确提出, 至今已有十多年的应用与发展。早期, 受限于硬件条件, 大多数研究采用传统机器学习结合特征工程的方法建模方面级情感分析任务。比如在最大熵、支持向量机等经典机器学习模型的基础上结合情感词典、依赖信息等对情感进行分类^[5], 这些方法需要投入较大的人力成本和时间成本。而随着技术的逐渐成熟与发展, 各种硬件设施越来越先进, 依托于先进的设备, 深度学习得到了良好的应用环境, 大大降低了 ABSA 研究的人力成本与时间成本, 所以深度学习渐渐代替了传统机器学习在方面级情感分析的研究地位。

循环神经网络(Recurrent Neural Network, RNN)对序列信息有良好记忆功能, 所以其对文本这种具有序列特征的信息具有良好的建模性能。长短时记忆网络(Long Short Term Memory, LSTM)和门控神经网络(Gated Recurrent Units, GRU)在传统 RNN 的基础上融入门控基础, 提高了网络对上下文数据的记忆性能, 从而改善了传统 RNN 存在的梯度爆炸和梯度弥散等问题。Tang 等人^[6]提出了两个改进模型(TD-LSTM 和 TC-LSTM)用于 ABSA 任务, 整合了方面项与上下文知识, 比直接使用 LSTM 对文本进行建模取得了更好的效果。注意力机制能根据文本的特征计算得到上

下文的语义相关性^[7], 所以, 研究人员结合注意力机制和 LSTM 网络在 ABSA 任务上开展了大量研究。Wang 等人^[8]基于方面项的词嵌入和注意力机制构建了 ATAE-LSTM 模型, 在 ABSA 任务上取得了良好的成果。Tan 等人^[9]针对文本中具有冲突意见的意见项, 提出了一种双重注意力网络, 以识别文本中方面项的冲突情感。另外, Devlin 等人^[10]提出的基于注意力机制的 BERT 预训练模型也一度刷新了 NLP 领域的各大任务成果。Sun 等人^[11]提出构造辅助句的方式微调 BERT 模型, 从句子对分类的角度考虑 ABSA 任务, 取得了良好的效果; Xu 等人^[12]针对 ABSA 任务提出了一种基于 BERT 的后训练微调方式, 提升了 BERT 预训练模型在微调 ABSA 任务上的性能。虽然以上方法使用循环神经网络对上下文序列信息建模, 并结合注意力机制在 ABSA 任务上取得了良好效果, 但是这些方法不能学习到文本的外部结构信息。

而利用文本的句法依赖知识有助于在方面项和文本上下文单词之间建立联系。遵循这一思路, 基于句法依赖知识的图神经网络(Graph Convolutional Networks, GCN)被广泛应用于 ABSA 任务。Lu 等人^[13]设计了一种星形图, 并提出基于节点和基于文档的分类方法, 在情感分析上取得了良好效果; Zhang 等人^[14]提出 ASGCN 模型, 首次使用文本句法依赖构建图卷积神经网络并将其用于 ABSA 任务; Liang 等人^[15]提出一种依赖关系嵌入式图卷积网络(DREGCN), 并且设计了一个简单有效的消息传递机制, 以充分利用语法知识和上下文语义知识, 在 ABSA 任务上取得了良好的效果; Pang 等人^[16]基于句法依赖树和多头注意力机制分别生成句法依赖图和语义图, 构建了动态多通道图卷积网络(DM-GCN)来有效地从生成的图中学习相关信息; Tang 等人^[17]利用依赖关系标签构建了一个潜在的文本依赖图, 并设计了一个情感细化策略引导网络捕捉方面项的情感线索; Liang 等人^[18]通过整合 SenticNet 词典的情感信息来增强句子的句法依赖, 在 ABSA 任务上取得了良好效果; Li 等人^[19]提出一种双图卷积神经网络同时对语法知识和语义知识进行学习, 并提出了正交正则化器和差分正则化器鼓励句法知识和语义知识相互感知学习; Xing 等人^[20]提出了 DigNet 模型, 设计了一种局部全局交互图, 将句法图和语义图结合在一起, 实现两个图中有益知识的相互交换。赵振等人^[21]提出基于关系交互的图注意力网络模型, 将关系特征融入到注意力机制中, 使用一个新的扩展注意力来引导信息从方面项的句法上下文传播到方面项本身; 阳影等人^[22]在句法依赖树上融合情感知识, 并与文本的语义相关性进行有效交互, 提出 SKDGCN 双通道 GCN 模型, 在多

个公开数据集上表现优异;杨春霞等人^[23]提出基于注意力与双通道的模型 TCCM,实现了通道内信息的残差互补,有效结合了句法与语义关联信息。以上基于图卷积网络的方法使用句法依赖、情感知识增强句法依赖来聚合文本的句法结构特征;使用注意力得分矩阵、句法解析标签聚合文本的上下文特征,这些方法在 ABSA 任务上取得了良好的效果。

而文中方法通过将上下文语义信息与句法依赖进行结合,从而将上下文语义信息融进句法依赖中,以此来引导方面项聚合具有上下文语义性的句法信息;并且使用文本自身的情感知识来引导方面项聚合上下文的情感信息;最后使用双通道的图卷积神经网络将具

有上下文语义性的句法信息和上下文情感信息进行融合,用于最终的情感分类。

2 模型

该文提出一种融合具有上下文语义信息的句法依赖知识和文本情感知识的双通道图卷积网络模型(DC-GCN),如图1所示。该模型主要由以下三个模块构成:(1)词嵌入与 BiLSTM 模块;(2)具有上下文语义信息的句法依赖图(SrG),整合 SenticNet 词典情感信息构建的一个文本情感依赖图(SeG)构造模块;(3)双通道图卷积网络(DC-GCN)特征融合与情感分类模块。

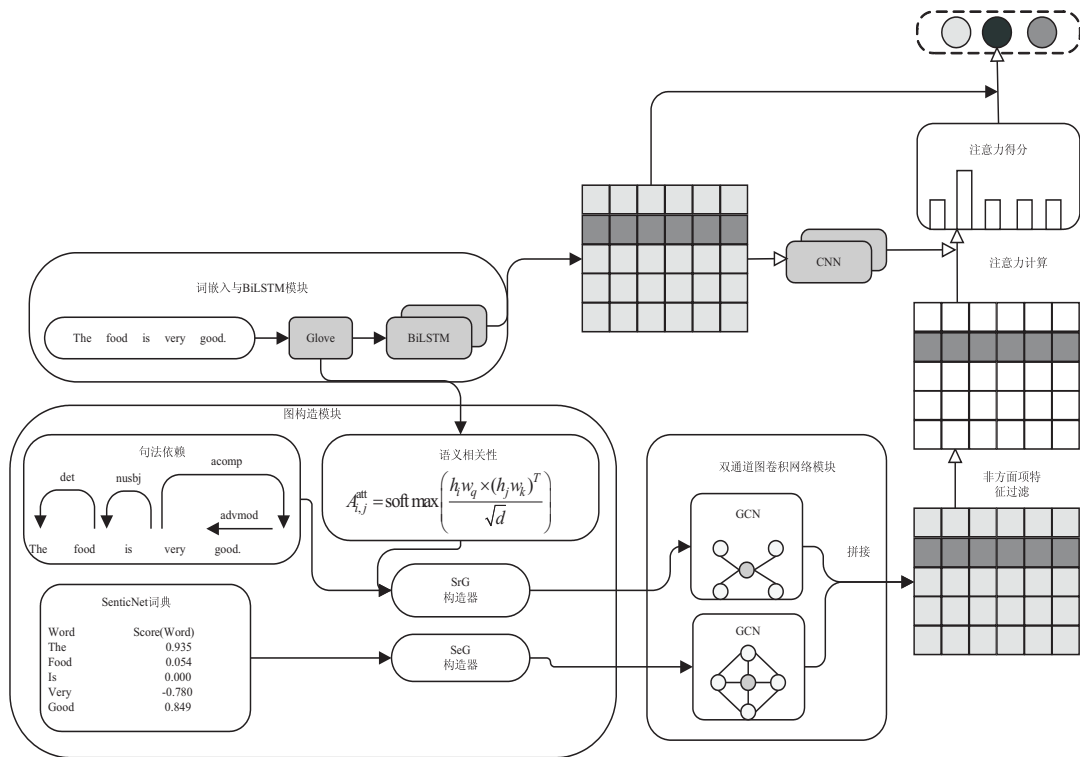


图1 DC-GCN 模型

2.1 词嵌入与 BiLSTM

使用 GloVe 将文本 $S = \{w_1, w_2, \dots, w_n\}$ 的每个单词都映射成一个三百维的词向量,然后使用 BiLSTM 中对其进行训练,生成了包含丰富上下文信息的文本词向量矩阵 $E = \{e_1, e_2, \dots, e_n\}$ 。

2.2 图构造模块

2.2.1 具有上下文语义信息的句法依赖图(SrG)构造

首先,使用 Spacy 依赖解析器提取文本的句法依赖树,以文本“The food is very good.”为例,将有向句法依赖转换成无向句法依赖,并考虑单词的自循环,示例如图2所示。

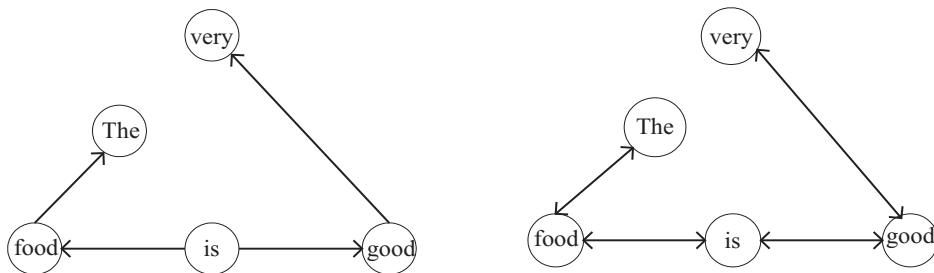


图2 有向句法依赖(左)与无向句法依赖(右)

根据文本的有向句法依赖图导出文本的邻接矩阵 $\mathbf{A}_{i,j}^{\text{syn}} \in R^{n \times n}$:

$$\mathbf{A}_{i,j}^{\text{syn}} = \begin{cases} 1 & \text{单词 } i \text{ 和 } j \text{ 存在句法依赖} \\ 0 & \text{其他} \end{cases} \quad (1)$$

其中, i, j 分别表示某一天评论数据中的第 i, j 个单词。

在有向句法依赖导出的邻接矩阵 $\mathbf{A}_{i,j}^{\text{syn}} \in R^{n \times n}$ 基础上, 加上 $\mathbf{A}_{i,j}^{\text{syn}} \in R^{n \times n}$ 的转置矩阵以及一个单位矩阵 \mathbf{I} 即可得到无向句法依赖的邻接矩阵 $\hat{\mathbf{A}}_{i,j}^{\text{syn}} \in R^{n \times n}$:

$$\hat{\mathbf{A}}_{i,j}^{\text{syn}} = \mathbf{A}_{i,j}^{\text{syn}} + (\mathbf{A}_{i,j}^{\text{syn}})^T + \mathbf{I} \quad (2)$$

其次, 使用多头自注意力机制求得文本上下文相关矩阵 $\mathbf{A}_{i,j}^{\text{att}} \in R^{n \times n}$:

$$\mathbf{A}_{i,j}^{\text{att}} = \text{softmax}\left(\frac{\mathbf{h}_i \mathbf{w}_q \times (\mathbf{h}_j \mathbf{w}_k)^T}{\sqrt{d}}\right) \quad (3)$$

其中, \mathbf{h}_i 和 \mathbf{h}_j 分别表示数据本经过 BiLSTM 层的第 i 和第 j 个单词的词向量, \mathbf{w}_q 和 \mathbf{w}_k 是两个可训练的参数矩阵, d 表示词向量的维度。

最终使用上下文相关矩阵对无向句法依赖邻接矩阵进行增强, 生成一个具有上下文相关性的句法依赖邻接矩阵 $\mathbf{D}_{i,j}^{\text{SrG}} \in R^{n \times n}$:

$$\mathbf{D}_{i,j}^{\text{SrG}} = \hat{\mathbf{A}}_{i,j}^{\text{syn}} \times \mathbf{A}_{i,j}^{\text{att}} \quad (4)$$

2.2.2 整合 SenticNet 词典情感信息构建的一个文本情感依赖图 (SeG) 构造

该文基于 SenticNet 词典构建了潜在的文本情感依赖图 (SeG), 其中 SenticNet 词典是由单词及其单词的情感得分组成, 示例如表 1 所示。

表 1 SentiNet 词典示例

word	Score(word)
Great	0.857
Fantastic	0.870
Food	0.170
Bad	-0.800
Hard	-0.059

首先, 基于 SenticNet 词典中单词的情感得分构建文本情感依赖图 (SeG) 并得到情感得分矩阵 $\mathbf{A}_{i,j}^{\text{Seg}} \in R^{n \times n}$:

$$\mathbf{A}_{i,j}^{\text{Seg}} = \text{Score}(w_i) + \text{Score}(w_j) \quad (5)$$

另外, 为了使网络对方面项的关注度增加, 增大了方面项的情感得分:

$$\mathbf{T}_{i,j}^{\text{asp}} = \begin{cases} 1 & w_i w_j \text{ is aspect} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

最终, 得到了基于文本情感知识的文本情感依赖图邻接矩阵 $\mathbf{D}_{i,j}^{\text{SeG}} \in R^{n \times n}$:

$$\mathbf{D}_{i,j}^{\text{SeG}} = \mathbf{A}_{i,j}^{\text{Seg}} + \mathbf{T}_{i,j}^{\text{asp}} + 1 \quad (7)$$

2.3 双通道图卷积模块

2.3.1 特征提取

该文基于 SrG 和 SeG 两个文本依赖图构建了双通道的图卷积神经网络, 以同时利用文本上下文单词的语义相关性和单词自身的情感信息来提取增强后的句法特征和情感知识特征。

$$\hat{h}_{\text{SrG}}^l = \sum_{j=0}^k \mathbf{D}_{i,j}^{\text{SrG}} \mathbf{W}_1^l h_{\text{SrG}}^{l-1} \quad (8)$$

$$h_{\text{SrG}}^l = \text{ReLU}(\hat{h}_{\text{SrG}}^l / \sum_{j=1}^n \mathbf{D}_{i,j}^{\text{SrG}} + b_1^l) \quad (9)$$

$$\hat{h}_{\text{SeG}}^l = \sum_{j=0}^k \mathbf{D}_{i,j}^{\text{SeG}} \mathbf{W}_2^l h_{\text{SeG}}^{l-1} \quad (10)$$

$$h_{\text{SeG}}^l = \text{ReLU}(\hat{h}_{\text{SeG}}^l / \sum_{j=1}^n \mathbf{D}_{i,j}^{\text{SeG}} + b_2^l) \quad (11)$$

其中, h_{SrG}^{l-1} 和 h_{SeG}^{l-1} 分别表示上一层图卷积层输出的文本特征; \mathbf{W}_1^l 和 \mathbf{W}_2^l 是两个可训练的参数矩阵; b_1^l 和 b_2^l 是两个可训练偏置项。

2.3.2 特征融合

首先使用拼接的方式, 将双通道图卷积网络输出的文本特征进行融合:

$$h_{\text{GCN}}^l = \text{cat}(h_{\text{SrG}}^l; h_{\text{SeG}}^l) \quad (12)$$

再使用平均池化提取特征的重要信息并对特征进行降维:

$$h^l = \text{avgPooling}(h_{\text{GCN}}^l) \quad (13)$$

2.4 非方面项特征过滤

为了获取到融合了具有语义相关性的句法信息和聚合了情感知识的方面项特征, 便于接下来进行方面项的权重计算, 将双通道图卷积网络 (DC-GCN) 输出的非方面项特征过滤, 即将公式 13 输出的 h^l 的非方面项特征置为零向量, 而方面项特征保持不变, 如下所示:

$$h_t^l = 0 \quad 1 \leq t < \gamma + 1, \gamma + m < t \leq n \quad (14)$$

$$H_{\text{GCN}}^l = \{0, 0, \dots, h_{\gamma+1}^l, \dots, h_{\gamma+m}^l, \dots, 0, 0\} \quad (15)$$

2.5 获取文本最终特征表示

2.5.1 方面项局部特征提取

使用卷积神经网络提取方面项的具局部特征, 并过滤掉文本中非方面项特征, 实现的方法如下:

$$h_{\text{CNN}}^l = \text{Conv1D}(h_{\text{GCN}}^{l-1}) \quad (16)$$

$$h_{t-\text{CNN}}^l = 0 \quad 1 \leq t < \gamma + 1, \gamma + m < t \leq n \quad (17)$$

$$H_{\text{CNN}}^l = \{0, 0, \dots, h_{\gamma+1}^l, \dots, h_{\gamma+m}^l, \dots, 0, 0\} \quad (18)$$

2.5.2 方面项权重计算

首先运用双通道图卷积网络输出的方面项特征和卷积神经网络输出的方面项特征进行计算, 并在特征维度上做加法, 得到最终方面项及非方面项的权重:

$$\beta_t = \sum_{i=1}^n h_i^l (h_{t-\text{CNN}}^l)^T = \sum_{i=\gamma+1}^{\gamma+m} h_i^l (h_{t-\text{CNN}}^l)^T \quad (19)$$

$$\alpha_i = \frac{\exp(\beta_i)}{\sum_{i=1}^n \exp(\beta_i)} \quad (20)$$

2.5.3 最终特征表示

最后将特征权重反馈到 BiLSTM 中,使得富含多种依赖信息的方面项特征感知到上下文的情感信息,以获取最终的文本特征表示:

$$r = \sum_{i=1}^n a_i e_i \quad (21)$$

其中, e_i 是 BiLSTM 输出的文本第 i 个单词的向量表示。

2.6 情感分类

在获取到文本的最终特征表示后,将其馈送进一个全连接层并使用 softmax 进行归一化,得到情感极性的概率分布 $p \in R^{(d_p \times 2d_p)}$:

$$p = \text{softmax}(W_p r + b_p) \quad (22)$$

其中, d_p 是情感类别的维度, d_n 是向量的维度, W_p 和 b_p 分别是一个可训练的参数矩阵和偏置。

文中模型选择具有交叉熵损失和 L2 正则化的标准梯度下降算法进行训练:

$$\text{Loss} = - \sum_{(c,t) \in D} \log p_t + \lambda \|\theta\|_2 \quad (23)$$

其中, D 表示数据集, t 表示标签, p_t 表示标签 t 的概率, θ 表示所有可以训练的参数, λ 表示 L2 正则化的系数。

3 实验

3.1 数据集与实验环境

3.1.1 实验数据集

在四个公开的数据集上进行了实验, Twitter 数据集是从 Twitter 上收集整理的一些评论数据; Rest14 数据集和 Lap14 数据集来自于 SemEval-2014 方面级情感分析大赛; Rest16 数据集是 SenEval-2016 方面级情感分析大赛提供的公开数据集。每个数据集都提供了方面项及其相应的情感极性, 数据集详细统计如表 2 所示。

表 2 数据集

数据集	分布	积极	中性	消极
Twitter	训练集	1 561	3 127	1 560
	测试集	173	246	173
Rest14	训练集	2 164	637	807
	测试集	726	196	196
Lap14	训练集	994	464	870
	测试集	341	169	128
Rest16	训练集	1 620	88	709
	测试集	597	38	190

3.1.2 实验环境

实验环境如表 3 所示。

表 3 实验环境

实验设备	具体信息
操作系统	Windows10
CPU	i5-11400F
显卡	RTX2060
内存	16G
开发语言	Python3.9
开发平台	Pythoch 深度学习框架
开发工具	Pycharm

3.1.3 实验评价指标

为了便于与对比模型作比较,以体现文中模型的有效性,选择 ABSA 任务模型性能的通用评价指标—准确率 Acc 和 Macro-F1, 两者的计算过程如下所示:

$$P = \frac{TP}{TP + FP} \quad (24)$$

$$R = \frac{TP}{TP + FN} \quad (25)$$

$$\text{Acc} = \frac{TP + TN}{TP + FP + FN + TN} \quad (26)$$

$$F1 = \frac{2PR}{P + R} \quad (27)$$

$$\text{Macro-F1} = \frac{1}{C} \sum_{v=1}^C F1^v \quad (28)$$

其中, TP 表示成功预测的正样本数, FP 表示错误预测的负样本数, FN 是错误预测的正样本数, TN 表示成功预测的负样本数, P 为精确率, R 是召回率, F1 为精确率与召回率的调和平均, C 是情感类别数, $F1^v$ 为第 v 个类别的 F1 值。

3.2 基线模型及结果分析

ATAE-LSTM^[8]: 利用方面嵌入和注意力机制进行方面级情感分类。

IAN^[24]: 在 BiLSTM 的基础上利用交互注意力机制, 实现方面项与上下文信息的交互。

RAM^[25]: 使用多层注意力和记忆网络学习句子表达, 用于方面级情感分析。

MGAN^[26]: 结合粗粒度和细粒度的注意力机制, 设计了一个方面对齐损失实现不同方面项与上下文之间的词级交互。

TNet^[27]: 将 BiLSTM 词嵌入转换为特定方面项嵌入, 并用 CNN 提取用于分类的特征。

ASGCN^[14]: 首次提出基于句法依赖树的图卷积神经网络用于方面项情感分析。

CDT^[28]: 利用基于句法依赖树的 GCN 来学习文本特征。

BiGCN^[29]:提出基于词共现信息和句法依赖树的层次图卷积网络。

kumaGCN^[30]:使用潜在的图结构信息来补充句法特征。

CPA-SA^[31]:设计了两个不对称的上下文权重函数来调整方面项特征的权重。

IAGCN^[32]:使用 BiLSTM 和修正动态权重对上下

文进行建模,GCN 对句法信息编码,再利用交互注意力机制学习上下文和方面项的注意力,将上下文信息和方面项进行重构。

该文选取的对比模型涵盖了方面级情感分析从 2016 年到 2022 年期间在方面级情感分析上取得良好效果的经典模型以及最新的模型。对比实验结果如表 4 所示。

表 4 对比实验结果 %

Model	Twitter	Rest14	Lap14	Rest16
	Accuracy/Macro-F1	Accuracy/Macro-F1	Accuracy/Macro-F1	Accuracy/Macro-F1
ATAE-LSTM	-----/-----	77.20/-----	68.70/-----	-----/-----
IAN	-----/-----	78.60/-----	72.10/-----	-----/-----
RAM	69.36/67.30	80.23/70.80	74.49/71.35	-----/-----
MGAN	72.54/70.81	81.25/71.94	75.39/72.47	-----/-----
TNet	74.90/73.60	80.69/71.27	76.54/71.75	-----/-----
ASGCN	72.15/70.40	80.77/72.02	75.55/71.05	88.99/67.48
CDT	74.66/73.66	82.30/74.02	77.19/72.99	85.58/69.93
BiGCN	74.16/73.35	81.97/73.48	74.59/71.84	88.96/70.84
kumaGCN	72.45/70.77	81.43/73.64	76.12/72.42	89.39/73.19
CPA-SA	-----/-----	82.64/73.38	75.18/71.50	89.02/72.43
IAGCN	73.41/71.90	81.49/73.65	75.97/72.00	88.31/69.69
DC-GCN	75.58/74.61	82.95/75.26	77.12/73.61	90.42/75.94

分析表 4 结果发现,与基线模型中取得最好效果的模型作对比,总体上文模型取得了更优的效果,在四个公开数据集(Twitter, Rest14, Lap14 和 Rest16)上 Macro-F1 分别比对比模型中最好的结果高出 0.95 百分点、1.24 百分点、0.62 百分点、2.75 百分点;在 Twitter, Rest14, Rest16 上 Accuracy 分别比对比模型中最好的结果高出 0.68 百分点、0.31 百分点、1.03 百分点,但是在 Lap14 上 Accuracy 比 CDT 模型低 0.07。对比了 Lap14 和 Twitter, Rest14, Rest16 的差别,发现 Lap14 中的数据对方面项隐式表达情感的样本相较于其他三个数据集占比更高,这会导致模型更难根据文本依赖图引导方面项感知上下文情感信息,从而影响模型的性能。

另外,在训练过程中,文中模型在 BiLSTM 和 GCN 网络中均使用正则化来防止模型出现过拟合现象。为探究不同的正则化系数对实验结果的影响,以 Lap14 为例,设置了 5 组正则化系数(在 BiLSTM 和 GCN 的正则化系数分别为 0.7 与 0.4, 0.7 与 0.5, 0.7 与 0.6, 0.6 与 0.5, 0.8 与 0.5)展开对比实验,结果如图 3 所示。

图 3 中横坐标 B_0.7/G_0.4 表示 BiLSTM 和 GCN 中的正则化参数分别设置为 0.7 和 0.4,其余四个横坐标的含义以此类推;纵坐标表示准确率和 F1 值;经图 3 可知,正则化参数对模型最终的性能有影响,在 BiLSTM 和 GCN 网络中的正则化参数选择取得最好效果的 0.7 和 0.5。

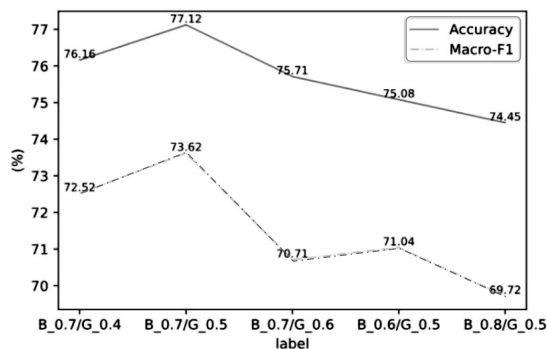


图 3 不同正则化系数下的模型效果

3.3 消融实验

为验证具有上下文语义信息的句法依赖图、基于情感知识构建的文本情感依赖图对提升模型性能的有效性,以及使用 CNN 提取方面项局部特征对更新方面

项权重的有效性,在保持参数一致的情况下开展消融实验,结果如表5所示。

表5 消融实验结果 %

Model	Twitter	Rest14	Lap14	Rest16
	Accuracy/Macro-F1	Accuracy/Macro-F1	Accuracy/Macro-F1	Accuracy/Macro-F1
DC-GCN	75.58/74.61	82.95/75.26	77.12/73.61	90.42/75.94
DC-GCN w/o SrG	74.57/73.21	82.14/74.58	76.49/72.20	88.96/70.30
DC-GCN w/o SeG	74.86/73.33	82.41/74.96	75.70/71.47	88.31/69.82
DC-GCN (SrG w/o Att)	74.28/73.12	82.59/74.09	75.55/71.88	89.45/71.73
DC-GCN w/o CNN	73.40/72.75	81.34/71.94	75.08/70.83	89.45/68.49

DC-GCN:表示该文提出的完整的模型,使用了具有上下文语义信息的句法依赖图(SrG)、情感知识构建的文本情感依赖图,并使用CNN提取的方面项局部特征用于更新方面项权重。

DC-GCN w/o SrG:表示在完整的模型中移除了具有上下文语义信息的句法依赖图(SrG)。

DC-GCN w/o SeG:表示在完整的模型中移除了基于情感知识构建的潜在的文本依赖图(SeG)。

DC-GCN (SrG w/o Att):表示不使用自注意力机制对文本的句法依赖进行增强。

DC-GCN w/o CNN:表示在计算方面项权重的过程中不使用CNN提取方面项的局部特征参与方面项权重更新过程。

根据表5中消融实验的结果证明:在移除具有上下文语义信息的SrG、移除基于情感知识构建的SeG、移除SrG中的上下文信息、不使用CNN提取方面项的局部特征参与方面项权重更新过程,均会导致模型性能的下降。由此证明了所提方法在方面级情感分析上的有效性。

4 结束语

设计了具有上下文相关性的SrG和SeG,并基于两个图构建了一个双通道的图卷积神经网络(DC-GCN),实现了对文本特征的优化以及多种文本特征的有效融合,提升了模型在ABSA任务上的性能。通过大量实验证明该方法相比于基线模型有更好的效果。

虽然提出的模型在ABSA任务上取得了不错的效果,但模型仍有进一步改进的空间。一方面,可以设计更有效的特征融合方法以提升模型的性能;另一方面,需要将模型扩展到更多的数据集上进行训练,以增强模型的鲁棒性。

参考文献:

[1] LIU B, ZHANG L. A survey of opinion mining and sentiment analysis[M]//Mining text data. Boston; Springer, 2012: 415-463.
[2] 李帅,徐彬,韩玮珂,等. SS-GCN:情感增强和句法增

强的方面级情感分析模型[J]. 计算机科学, 2023, 50(3): 3-11.

- [3] YU C, WU T, LI J, et al. Syngen: a syntactic plug-and-play module for generative aspect-based sentiment analysis[C]//ICASSP 2023-2023 IEEE international conference on acoustics, speech and signal processing (ICASSP). Hong Kong: Association for Computational Linguistics, 2023: 1-5.
[4] THET T T, NA J C, KHOO C S G. Aspect-based sentiment analysis of movie reviews on discussion boards[J]. Journal of Information Science, 2010, 36(6): 823-848.
[5] WIEBE J, WILSON T, CARDIE C. Annotating expressions of opinions and emotions in language[J]. Language Resources and Evaluation, 2005, 39: 165-210.
[6] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification[J]. arXiv: 1512. 01100, 2015.
[7] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
[8] 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. Austin: Association for Computational Linguistics, 2016: 606-615.
[9] TAN X, CAI Y, ZHU C. Recognizing conflict opinions in aspect-level sentiment classification with dual 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). Hong Kong: Association for Computational Linguistics, 2019: 3426-3431.
[10] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding[J]. arXiv: 1810. 04805, 2018.
[11] SUN C, HUANG L, QIU X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence[J]. arXiv: 1903. 09588, 2019.
[12] XU H, LIU B, SHU L, et al. BERT post-training for review reading comprehension and aspect-based sentiment analysis[J]. arXiv: 1904. 02232, 2019.

- [13] LU H, HUANG S H, YE T, et al. Graph star net for generalized multi-task learning[J]. arXiv:1906.12330, 2019.
- [14] ZHANG C, LI Q, SONG D. Aspect-based sentiment classification with aspect-specific graph convolutional networks[J]. arXiv:1909.03477, 2019.
- [15] LIANG Y, MENG F, ZHANG J, et al. A dependency syntactic knowledge augmented interactive architecture for end-to-end aspect-based sentiment analysis[J]. Neurocomputing, 2021, 454:291-302.
- [16] PANG S, XUE Y, YAN Z, et al. Dynamic and multi-channel graph convolutional networks for aspect-based sentiment analysis[C]//Findings of the association for computational linguistics; ACL-IJCNLP 2021. [s. l.]: Association for Computational Linguistics, 2021:2627-2636.
- [17] TANG S, CHAI H, YAO Z, et al. Affective knowledge enhanced multiple-graph fusion networks for aspect-based sentiment analysis[C]//Proceedings of the 2022 conference on empirical methods in natural language processing. Abu Dhabi: Association for Computational Linguistics, 2022:5352-5362.
- [18] 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.
- [19] LI R, CHEN H, FENG F, et al. Dual graph convolutional networks for aspect-based sentiment analysis[C]//Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: long papers). [s. l.]: Association for Computational Linguistics, 2021:6319-6329.
- [20] XING B, TSANG I. Dignet: digging clues from local-global interactive graph for aspect-level sentiment classification[J]. arXiv:2201.00989, 2022.
- [21] 赵振, 朱振方, 王文玲. 基于关系特征交互的方面级情感分类方法[J]. 计算机技术与发展, 2023, 33(3):187-193.
- [22] 阳影, 张凡, 李天瑞. 基于情感知识的双通道图卷积网络的方面级情感分析[J]. 计算机科学, 2023, 50(5):230-237.
- [23] 杨春霞, 徐奔, 桂强, 等. 基于注意力与双通道网络的方面级情感分析[J]. 南京信息工程大学学报:自然科学版, 2023, 15(1):42-50.
- [24] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification[J]. arXiv:1709.00893, 2017.
- [25] CHEN P, SUN Z, BING L, et al. Recurrent attention network on memory for aspect sentiment analysis[C]//Proceedings of the 2017 conference on empirical methods in natural language processing. Copenhagen: Association for Computational Linguistics, 2017:452-461.
- [26] FAN F, FENG Y, ZHAO D. Multi-grained attention network for aspect-level sentiment classification[C]//Proceedings of the 2018 conference on empirical methods in natural language processing. Brussels: Association for Computational Linguistics, 2018:3433-3442.
- [27] LI X, BING L, LAM W, et al. Transformation networks for target-oriented sentiment classification[C]//Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers). Melbourne: Association for Computational Linguistics, 2018:946-956.
- [28] 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). Hong Kong: Association for Computational Linguistics, 2019:5679-5688.
- [29] ZHANG M, QIAN T. 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). [s. l.]: Association for Computational Linguistics, 2020:3540-3549.
- [30] CHEN C, TENG Z, 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). [s. l.]: Association for Computational Linguistics, 2020:5596-5607.
- [31] HUANG B, GUO R, ZHU Y, et al. Aspect-level sentiment analysis with aspect-specific context position information[J]. Knowledge-Based Systems, 2022, 243:108473.
- [32] 王娅丽, 张凡, 余增, 等. 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.