

郭荣荣, 高建瓴, 徐瑞涓, 等. 基于多通道图卷积网络的方面级情感分析[J]. 智能计算机与应用, 2024, 14(5): 36-43. DOI: 10.20169/j.issn.2095-2163.240505

## 基于多通道图卷积网络的方面级情感分析

郭荣荣, 高建瓴, 徐瑞涓, 戚玲珑

(贵州大学 大数据与信息工程学院, 贵阳 550025)

**摘要:** 方面情感三元组提取旨在通过提取方面词、观点词及情感极性的三元组来为方面级情感分析提供解决方案。以往的研究存在无法处理句子中方面词和观点词之间的一对多和多对一的关系, 以及在不同的子任务中存在错误传播的问题。本文提出一种基于多通道图卷积网络的情感分析方法, 通过设计多通道图卷积网络来捕获单词的词性组合信息和结构信息。使用图卷积网络在异构图上重复图卷积操作得到句法依存模块的输入, 同时使用双仿射注意力模块获取句子中单词间的关系概率分布。研究选取的数据集 Res14、Lap14、Res15 和 Res16 上的仿真实验表明, 在  $F1$  值上, 本文模型与现有的基线模型相比, 取得了更好结果。

**关键词:** 自然语言处理; 情感分析; 图卷积网络; 三元组提取

中图分类号: TP391.1

文献标志码: A

文章编号: 2095-2163(2024)05-0036-08

### Aspect-level sentiment analysis based on multi-channel graph convolutional network

GUO Rongrong, GAO Jianling, XU Ruijuan, QI Linglong

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

**Abstract:** Aspect sentiment triple extraction aims to provide solutions for aspect-level sentiment analysis by extracting the triples including aspect words, opinion words and sentiment polarity. Previous studies have been unable to deal with one-to-many and many-to-one relationship between aspect words and opinion words in sentences, as well as the problem of error propagation in different subtasks. This paper proposes a sentiment analysis method based on multi-channel graph convolutional network, which captures the part-of-speech combination information and structural semantic information of words by designing multi-channel graph convolutional network. The graph convolution network is used to repeat the graph convolution operation on the heterogeneous graph to obtain the input of the syntactic dependency module, and the double affine attention module is used to obtain the relationship probability distribution between words in the sentence. Experiments on data sets Res14, Lap14, Res15 and Res16 show that the proposed model achieves more significant results on  $F1$  values than the existing baseline model.

**Key words:** natural language processing; sentiment analysis; graph convolutional network; triple extraction

## 0 引言

近年来,随着网络社交媒体的快速发展,很多人都会在网络平台上发表对某些产品、服务和热点话题的看法。面对网络上海量的数据,如何利用自然语言处理(Natural Language Processing, NLP)技术挖掘文本中的关键信息是当前研究的热点,因此对这

些海量数据展开研究有着重要的理论价值和现实意义<sup>[1]</sup>。情感分析,也称为意见挖掘,是 NLP 领域的一项重要任务,旨在分析人类语言中的情感、观点、评价和态度<sup>[2]</sup>。方面级情感分析<sup>[3]</sup>(Aspect-based Sentiment Analysis, ABSA)则立足于分析人们对实体不同方面的观点,目前已经引起众多学者的广泛关注和浓厚兴趣。

现有的 ABSA 主要将单个子任务或 2 个子任务

**基金项目:** 国家自然科学基金(62166006)。

**作者简介:** 郭荣荣(1995-),男,硕士研究生,主要研究方向:自然语言处理;徐瑞涓(1997-),女,硕士研究生,主要研究方向:自然语言处理;戚玲珑(2000-),女,硕士研究生,主要研究方向:数据挖掘。

**通讯作者:** 高建瓴(1969-),女,副教授,硕士生导师,主要研究方向:数据库系统,数据挖掘。Email: 454965711@qq.com

收稿日期: 2023-04-18

组合,这些子任务包括方面词提取<sup>[4]</sup>(ATE)、观点词提取<sup>[5]</sup>(OTE)、方面观点对提取<sup>[6]</sup>(AOPE)和方面情感分类<sup>[7]</sup>(ASC)。方面情感三元组抽取<sup>[8-9]</sup>(Aspect Sentiment Triplet Extraction, ASTE)旨在提取由方面词、观点词和情感极性组成的三元组。方面词也称为意见目标,观点词是指用来表达观点的词,情感极性表示用户在方面词表达的情绪类型,如积极、中性和消极。在这类任务中,不仅要从句子中提取目标实体和对应的观点,还需要得到目标和观点间的情感依赖,形成三元组进行提取。

近年来,研究人员提出一种多任务学习框架,联合提取方面词和观点词及其相关情感依赖的三元组算法。例如, Peng 等学者<sup>[10]</sup>提出一种两阶段方法实现 ASTE 任务。在第一阶段,从句子中提取所有潜在的意见目标和意见表达,并确定每个意见目标的情感极性。在第二阶段,将意见目标与其相应的意见表达相关联,以获得方面情感三元组。但这种方法没有充分利用方面词和观点词之间的交互关系,因此 Xu 等学者<sup>[11]</sup>提出一种位置感知标记方案,将意见表达和情感极性信息整合到标记意见目标的标签中,从而可以使用统一的标签序列来识别句子中的所有三元组。Zhang 等学者<sup>[12]</sup>提出一种多任务学习框架来联合提取方面词和意见词。在第一阶段,分别提取方面词和观点词;在第二阶段,使用情感依存解析器将情感分析任务转化为填表问题。

Chen 等学者<sup>[13]</sup>提出一种增强型多通道图卷积网络模型,通过将单词和邻接张量的关系视为节点和边,将句子转换为多通道图来学习关系感知节点表示,以充分利用单词之间的关系来提取方面情感三元组。由于现有的研究工作忽略了三元组中 3 个元素之间隐含的语法关系, Shi 等学者<sup>[14]</sup>提出了一

种基于依存图交互注意网络的解决方法,该方法通过从双向长短记忆神经网络中学习到的上下文特征和对应依存图中学习到的句法依存来引入单词之间的句法和语义关系。但由于现有的研究依赖先验假设,将每个观点词和方面词视为一组独立的词,因此在 ASTE 任务上表现不佳。Chen 等学者<sup>[15]</sup>在 2022 年针对 ASTE 任务提出了一种基于跨度的双解码器模型,这种结构可以有效地在三元组之间建立关联,融合句子中的语义信息,并在 2 个解码器间传递信息。

上述部分模型采用传统的编码方式对文本进行编码,这导致对上下文的特征提取不佳,影响分类效果。同时,大多数模型没有考虑词与词之间的关系,并且语义和句法信息的高度整合导致对句子的误解。本文贡献如下:

- (1) 本文提出一种多通道图卷积网络(BERT-Multi-Channel-GCN, BMGCN)模型来捕获单词的词性组合信息和结构信息,使用基于语言特征的方法来提高 GCN 网络的特征提取能力,这些方法包括词性组合、句法依存和基于树的距离。
- (2) 使用  $Bert_{Large}$  作为编码器,相比于  $Bert_{Base}$ ,更大的 Bert 模型可以捕捉更长距离的依赖,进一步提高语言的表征能力和特征提取能力。
- (3) 采用 Biaffine Attention 处理输入向量,获取单词间的关系概率分布。在数据集上的实验结果证明本文提出的 BMGCN 模型的有效性。

### 1 基于多通道图卷积的情感分析

BMGCN 模型主要包括 BERT ( Bert - Large - Uncased) 编码模块、双仿射注意力模块、图卷积网络模块和分类模块,如图 1 所示。

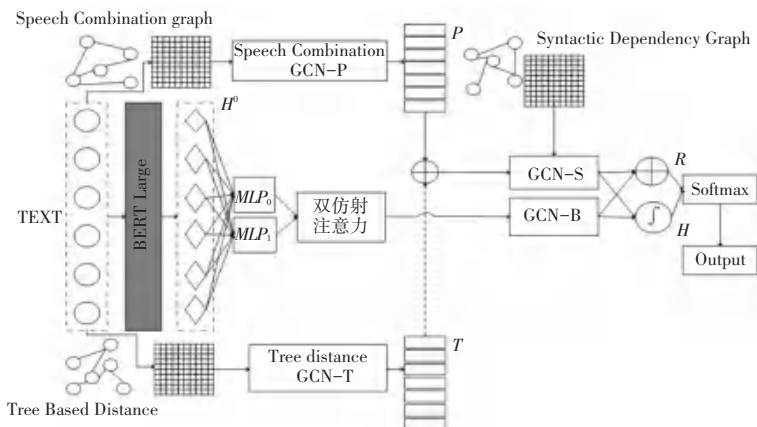


图 1 BMGCN 模型结构

Fig. 1 BMGCN model structure

### 1.1 任务定义

文章首先介绍 ASTE 任务的定义。ASTE 任务的目的是从给定输入一个具有  $n$  个单词的句子  $X = \{W_1, W_2, \dots, W_n\}$  中提取观点三元组  $T = \{(a, o, s)_k\}_{k=1}^{ml}$ , 其中  $a$  表示方面词,  $o$  表示观点词,  $s$  表示情感极性。在文章中, 将 ASTE 任务分解为以下 3 个子任务:

(1) 方面词标注。为句子中的方面词分配一个标签  $y_a \in \{A, A-F, A-T\}$ , 其中  $A$  表示词对  $(W_i, W_j)$  属于同一方面词,  $A-F$  表示句子中的第一个方面词,  $A-T$  表示句子中的第二个方面词。

(2) 观点词标注。为句子中的观点词分配一个标签  $y_o \in \{O, O-F, O-T\}$ , 其中  $O$  表示词对  $(W_i, W_j)$  属于同一观点词,  $O-F$  表示第一个观点词,  $O-T$  表示第二个观点词。

(3) 情感极性标注。为词对  $(W_i, W_j)$  分配情感极性标签  $y_s \in \{POS, NEG, NEU, N\}$ , 其中  $POS$ 、 $NEG$  和  $NEU$  表示与潜在意见目标相关的积极、消极和中性的情感极性,  $N$  表示其它。

### 1.2 Transformer 模型

本文使用 BERT<sup>[16]</sup> (Bert-Large-Uncased) 来编码上下文信息。BERT 是基于 Google 提出的多层双向 Transformer<sup>[17]</sup> 的编码器模型, 该编码器能够充分

提取上下文和句子间的一般语义特征。BERT 由多层双向的 Transformer 编码单元组成的, 整体结构如图 2 所示。

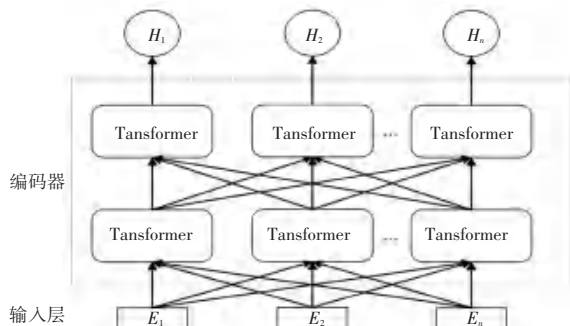


图 2 BERT 输出表示

Fig. 2 BERT output representation

### 1.3 输入输出表示

预训练的 BERT 可以生成序列的词向量并作为下游任务的输入。对于基于 BERT 的编码器, 在输入序列的开头放置标记 CLS, 在不同句子间放置标记 SEP 进行分割, 将特殊标记 CLS 和 SEP 分别表示开始标记和段标记。通过将词向量、句子向量和位置向量叠加, 然后将叠加后的向量序列送入 Transformer 层以获得上下文信息。BERT 的输入表示如图 3 所示。

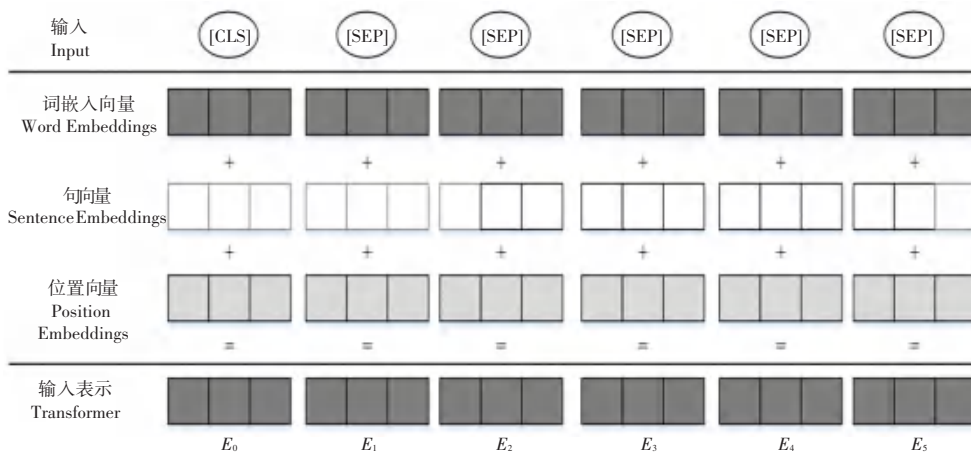


图 3 BERT 输入表示

Fig. 3 BERT input representation

### 1.4 双仿射注意力模块

本文使用双仿射注意力<sup>[18]</sup> (Biaffine Attention) 代替传统的 MLP 注意力机制或双线性机制。Biaffine Attention 使用双线性层, 比传统双层线性层和一个非线性激活单元的 MLP 网络更简单。本文使用 MLP 网络对 BERT 的输出向量  $r_i$  降维后作为

双仿射注意力模块的输入, 最终得到单词  $W_i$  自身向量对句中其它任意单词向量的注意力分数。研究中需用到的公式可写为:

$$h_0^{(\text{arc-dep})} = MLP^{(\text{arc-dep})}(r_i) \quad (1)$$

$$h_1^{(\text{arc-dep})} = MLP^{(\text{arc-dep})}(r_j) \quad (2)$$

$$s^{ij} = h_i^T U^{(0)} h_j + (h_j \oplus h_i)^T U^{(1)} + b \quad (3)$$

### 1.5 多通道图构建

多通道图构建是为了更好地将单词间的句法依存关系、词性组合关系和位置信息整合到图卷积网络模型中,以便学习到更准确的句子表示,增强模型的特征提取能力。Zhang 等学者<sup>[19]</sup>提出一种使用 Spacy 工具包生成输入句子依存树的方法,并在此基础上构建句法依存图。通过具有依存关系的词来

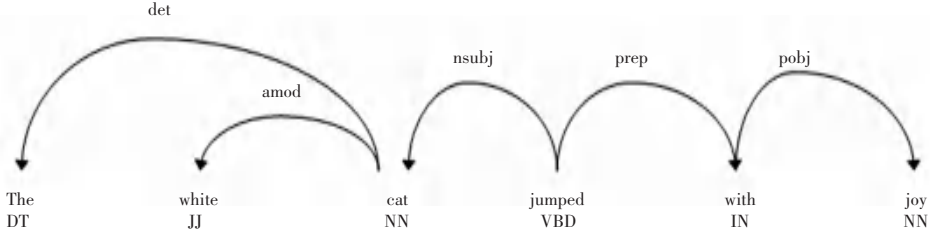


图 4 句法依存和词性组合树

Fig. 4 Syntactic dependency and part-of-speech combination tree

与一般语义图不同,基于树的距离的图侧重于句子内结构关系局部词的相关性。为了捕获句子序列中的结构关系,本文采用一种基于树的距离关系来构建特征图。通过计算句子中词对之间的距离关系得到邻接矩阵  $A^t$ , 并用于 GCN-T 模型的特征矩阵。关系定义见表 1。

表 1 距离关系示例

Table 1 Distance relationship example

The	white	cat	jumped	with	joy
1	2	3	4	6	5
2	1	2	3	5	4
3	2	1	2	4	3
4	3	2	1	3	2
6	5	4	3	1	2
5	4	3	2	2	1

### 1.6 多通道图卷积神经网络模块

图卷积神经网络<sup>[20]</sup> (GCN) 是传统卷积神经网络的变体。对于图结构数据,GCN 可以在直连节点上应用多层图卷积来编码局部信息。根据以往的研究<sup>[21]</sup>,以增强特征提取为目标,设计了多通道图卷积网络模块。在图 1 中,初始化词性组合矩阵  $A^p \in R^{n \times n}$  和树的距离矩阵  $A^t \in R^{n \times n}$ , 然后分别输入 2 个通道进行图卷积操作。词性组合和基于树的距离组合的表示为  $P = [p_1, p_2, \dots, p_n]$  和  $T = [t_1, t_2, \dots, t_n]$ , 由此推得的公式可写为:

$$p_i^l = \sigma \left( \sum_{j=1}^n A_{ij}^p W^l h_j^{l-1} + b_p^l \right) \quad (4)$$

构建依存关系图,从而获得更长距离的依存关系。通过对句子中任意 2 个词之间的句法关系建模得到句法依存矩阵  $A^s$ , 并用于句法 GCN-S 模型的特征矩阵。通过描述不同词之间的词性得到词性组合矩阵  $A^p$ , 然后用于词性 GCN-P 模型的特征矩阵。句法依存图和词性组合图如图 4 所示。

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

通过设计 2 个不同的通道学习句子的结构以获得 2 种不同类型的语义表示,然后采用一种基于句法依存树的方法来表达自己的句法结构,并将这种方法作为语义信息的强化和补充。2 个 GCN 模块、即  $P$  和  $T$  的隐藏状态表示都被引入到句法依存 GCN 模块中进行信息融合。以词性组合和基于树的距离两个通道的输出作为句法学习模块的输入,句法依存树通过依存矩阵  $A^s$  表示。数学表达如式(6)、式(7)所示:

$$s_i^l = p_i^l + t_i^l \quad (6)$$

$$S_i^l = \sigma \left( \sum_{j=1}^n A_{ij}^s W^l s_j^{l-1} + b_s^l \right) \quad (7)$$

其中,  $\sigma$  表示激活函数;  $W$  表示权重矩阵;  $b$  表示偏差。

为了对单词之间的各种关系建模,本文研究采用双仿射注意力模块获取单词间的关系概率分布,由多关系邻接张量  $L^{ij} \in L^{n \times n \times k}$  表示。邻接张量表示对 1.1 节中定义的单词之间关系建模。最后,使用 GCN 为每个节点聚合每个关系的信息。这一过程可以描述为式(8)、式(9):

$$H_b^{ij} = \sigma(R_b^{ij} H W_b + b_b) \quad (8)$$

$$H^{ij} = f(H_1^{ij}, H_2^{ij}, \dots, H_k^{ij}) \quad (9)$$

基于句法依存图和关系概率分布得到邻接矩阵,然后通过图卷积来编码局部信息获得隐藏节点表示  $S_i^l$  和  $H^{ij}$ , 最后将池化函数和连接操作应用于

节点和所有边,可由式(10)、式(11)进行描述:

$$H = f(S_i^l, H^{ij}) \quad (10)$$

$$R = R_i^l \oplus R^{ij} \quad (11)$$

### 1.7 输出层

通过将节点表示  $h_i, h_j$  和边表示  $r_{ij}$  连接起来获得用于标签预测的词对  $(w_i, w_j)$  表示,推得的公式见式(12)、式(13):

$$\theta = h_i \oplus h_j \oplus r_{ij} \quad (12)$$

$$y = \text{softmax}(W\theta_{ij} + b) \quad (13)$$

### 1.8 损失函数

模型以最小化预测值和真实值之间的交叉熵损失为目的,数学定义公式具体如下:

$$\text{Loss} = - \sum_i^k \tilde{y}_i \log y_i + \partial \| \Gamma \|_2 \quad (14)$$

其中,  $\tilde{y}_i$  表示真实的概率分布;  $y_i$  表示预测的概率分布;  $\Gamma$  表示所有的训练参数;  $\partial$  表示  $L_2$  正则化的权重;  $k$  表示训练样本总数。

## 2 实验结果和分析

### 2.1 环境配置与超参数设置

实验基于 Ubuntu18.04 的操作系统,Python 版

本为 3.8,PyTorch 版本为 1.4。为了在实验过程中取得更优的分类结果,本文将最大句子长度设置为 102,图卷积层的向量维度设置为 300, *Batch Size* 设置为 8。详细超参数设置见表 2。

表 2 超参数设置

Table 2 Hyperparameter setting

参数名	参数值
<i>Batch Size</i>	8
<i>BERT</i> <sub>Large</sub> 层向量维度	1 024
GCN 层向量维度	300
<i>Epoch</i>	100
学习率	0.001/0.000 02
<i>Dropout</i>	0.5
$\lambda$	0.000 1

### 2.2 数据集

本文实验是在 4 个公共数据集上进行,本文模型实验数据集来自 Wu 等学者<sup>[22]</sup> 标注的 Lap14、Rest14、Rest15 和 Rest16。Lap14 来自笔记本领域的评论,Res14、Res15 和 Res16 来自餐厅领域的评论。每个数据集的详细信息见表 3。

表 3 数据集

Table 3 Experimental datasets

Datasets	Res14		Lap14		Res15		Res14	
	#S	#T	#S	#T	#S	#T	#S	#T
Train	1 259	2 356	899	1 452	603	1 038	863	1 421
Dev	315	580	225	383	151	239	216	348
Test	493	1 008	332	547	325	436	328	525

### 2.3 评价指标

使用精确率( $P$ )、召回率( $R$ )和  $F1$  值为评价指标。各指标的数学定义公式分别如下:

$$P = \frac{T_{TP}}{T_{TP} + F_{FP}} \quad (15)$$

$$R = \frac{T_{TP}}{T_{TP} + F_{FN}} \quad (16)$$

$$F = \frac{2 \times P \times R}{P + R} \quad (17)$$

其中,  $T_{TP}$  表示实体为真,预测为真;  $F_{FP}$  表示实体为真,预测为假;  $F_{FN}$  表示实体为假,预测为真。

### 2.4 实验结果分析

为验证迭代次数对 BMGCN(BERT-MC-GCN)模型性能的影响,本文模型测试了在同一数据集下不同

迭代次数的实验结果,并测试在不同维度下对  $F1$  值的影响。最后,与 Peng-two-stage<sup>[10]</sup>、BMRC<sup>[23]</sup>、Dual-MRC<sup>[24]</sup>、GTS-BERT、GTS-BiLSTM、 $S^3 E^2$ <sup>[25]</sup> 进行对比,以验证提出模型的有效性。

#### 2.4.1 迭代次数对比

为了验证本文提出的 BMGCN 模型的有效性,本文对该模型的迭代次数做了进一步研究。在前述研究基础上,本文进行了多次迭代实验,这样一来就能反映模型在不同迭代次数下精确率、召回率和  $F1$  值的动态变化。实验结果如图 5~图 7 所示。

为了验证 *Epoch* 参数取值的有效性,本文对不同 *Epoch* 取值下的  $P$ 、 $R$  和  $F1$  值进行分析。从图 5~图 7 可以看出,当取值为 100 时,  $P$ 、 $R$  和  $F1$  值均趋于收敛。

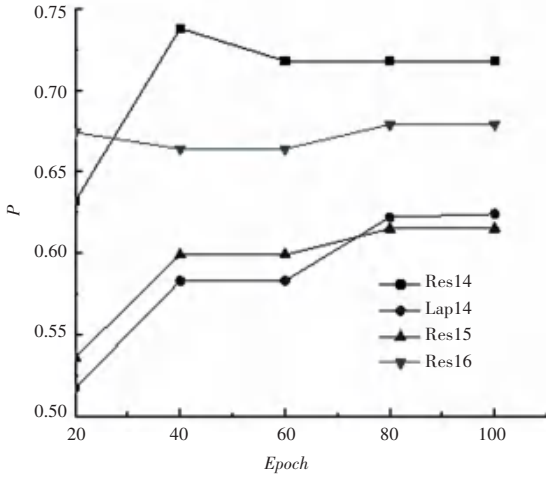


图 5 不同数据集的精确率 P

Fig. 5 Accuracy P of different datasets

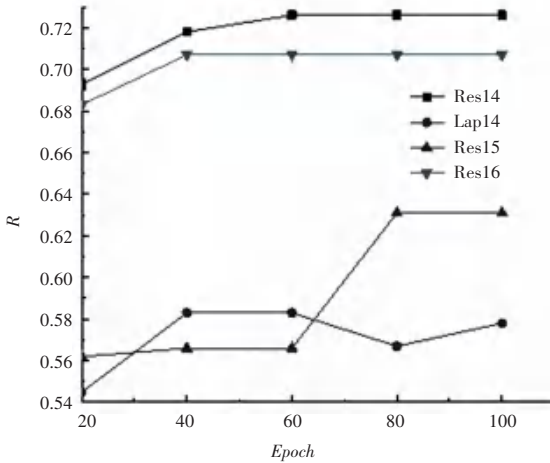


图 6 不同数据集的召回率 R

Fig. 6 Recall rate R of different datasets

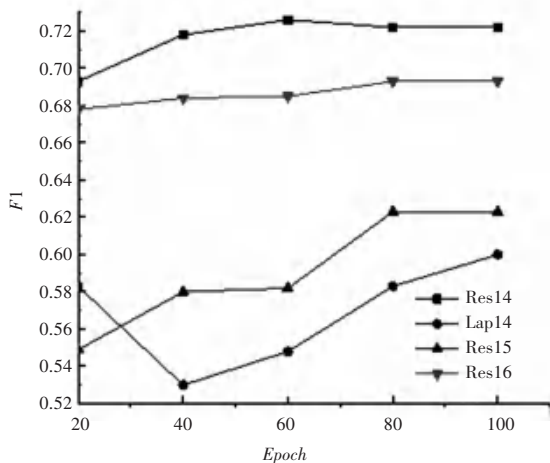


图 7 不同数据集的 F1 值

Fig. 7 F1 values of different datasets

### 2.4.2 词向量维度对比

词向量维度过大或者过小都会产生无法有效表示句子语义的问题,因此选择合适的词向量维度至关重要。由于本文采用  $Bert_{Large}$  版本的编码器,维度为 1 024,因此与使用维度为 768 的编码器 ( $Bert_{base}$ ) 进行对比实验。实验结果如图 8 所示。

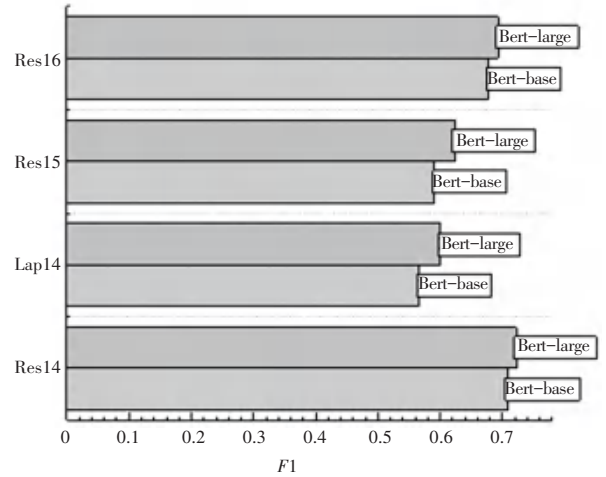


图 8 不同编码器的 F1 值

Fig. 8 F1 values of different encoders

由图 8 可知,使用  $Bert_{Large}$  作为编码器在精确率、召回率和 F1 值上取得更好的实验结果,原因可能是更大的 BERT 可以提供更好的语言特征表达能力和语义特征提取能力。

### 2.4.3 情感分析模型对比

为验证 BMGCN 模型在情感分析任务中的有效性,本文与以下几种情感分析模型进行对比。这里给出阐释分述如下。

(1) Peng-two-stage: 通过 2 个阶段抽取三元组。一阶段,提取潜在的方面词、情感极性和观点词;二阶段,对方面词和观点词进行配对。

(2) BMRC: 将 ASTE 任务转化为多轮机器阅读理解任务,设计 3 种类型的查询和双向 MRC 结构。

(3) Dual-MRC: 通过联合训练 2 个具有共享参数的 BERT-MRC 框架来解决三元组抽取任务。

(4) GTS-BERT 和 GTS-BiLSTM: 提出的一种网格标记方案,以端到端的方式完成三元组抽取。

(5)  $S^3E^2$ : 为 ASTE 任务设计了图序列对偶表示和建模范式来学习句子中词对之间的语义和句法关系,并使用图卷积网络进行编码以提取三元组。

为了评估模型的综合性能,该模型分别与上述模型进行对比,测试了精确率 P、召回率 R 和 F1 值,其中 F1 值用于评估框架的综合性能,实验结果见表 4。

表4 实验结果对比

Table 4 Comparison of experimental results

模型	Res14			Lap14			Res15			Res16		
	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>
Peng-two-stage	41.48	62.99	51.68	40.40	47.24	43.50	40.97	54.68	46.79	46.76	62.97	53.62
BMRC	-	-	70.01	-	-	57.83	-	-	58.74	-	-	67.49
Dual-MRC	71.55	69.14	70.32	57.39	53.88	55.58	63.79	51.87	57.21	68.60	66.24	67.40
GTS-BERT	70.92	69.49	70.20	57.52	51.92	54.58	59.29	58.07	58.67	68.58	66.60	67.58
GTS-BiLSTM	67.28	61.91	64.49	59.42	45.13	51.30	63.26	50.71	56.29	66.07	65.05	65.56
$S^3E^2$	69.08	64.55	66.74	59.43	46.23	52.01	61.06	56.44	58.66	71.08	63.13	66.87
<b>BMGCN</b>	71.80	72.59	<b>72.19</b>	62.38	57.80	<b>60.00</b>	61.46	63.08	<b>62.26</b>	67.89	70.75	<b>69.29</b>

表4给出在4个不同数据集上进行的实验结果对比。与其他几种方法相比,本文模型在综合评价指标 *F1* 值上取得最好的实验结果。相比于 Peng-two-stage、BMRC 和 Dual-MRC 这几种管道方法, BMGCN 模型在精确率和 *F1* 值上都取得更好的结果,这表明在捕获方面词和上下文之间的关系方面,句法结构与 GCN 的方法优于上述方法。与 GTS-BERT、GTS-BiLSTM 和  $S^3E^2$  这3种端到端的方法相比,本文模型在召回率和 *F1* 值上都取得更好的结果,这说明通过语义和句法的集成有助于利用单词和语言知识之间的关系来进行单词表示学习,这进一步提高模型的特征提取能力。同时,本文提出的方法不仅连接几种类型的信息,而是旨在学习每个句子单词的位置信息和句子间的词性组合信息,然后将2种集成到句法依存结构中。

相比于使用 BiLSTM 编码器的模型,使用各类型 BERT 编码器的模型(GTS-BERT, BMGCN)在上述4个数据集上取得更好的结果,原因可能是 BERT 已经在大规模数据上进行了预训练,可以提供更好的语言特征表达能力。

### 3 结束语

本文提出一种解决 ASTE 的评论文本情感分析模型。在输入层,采用人工处理的方式,建立多通道句法 GCN,从词性组合信息和位置语义信息中获取语义特征。利用双仿射注意力机制为文本信息分配权重,获取单词间的关系概率分布,并将其输入到图卷积神经网络中,获取文本信息的局部与全局特征。最后,构建了句法依存 GCN 模型学习句子的句法结构。实验结果表明,在4个数据集上与现有的几种常用的情感分析模型相比,本文提出的 BMGCN 模型在整体性能上得到了有效提升。由于基于句法依

存和结构语义的模型性能与句子的依存关系树存在很大的关系,因此,下一步将考虑通过减少由于句法性能不完善和评论句子的随意性引起的噪声信息来细化单词的依存表示,以解析结构更复杂的评论文本。

### 参考文献

- [1] ZHANG Lei, WANG Shuai, LIU Bing. Deep learning for sentiment analysis: A survey [J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018, 8(4): e1253.
- [2] PANG Bo, LEE L. Opinion mining and sentiment analysis [J]. Foundations and Trends® in Information Retrieval, 2008, 2(1-2): 1-135.
- [3] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. Semeval-2016 task 5: Aspect based sentiment analysis [C]//International Workshop on Semantic Evaluation. San Diego, USA: ACL, 2016: 19-30.
- [4] MA Dehong, LI Sujian, WU Fangzhao, et al. Exploring sequence-to-sequence learning in aspect term extraction [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: ACL, 2019: 3538-3547.
- [5] VEYSEH A P B, NOURI N, DERNONCOURT F, et al. Introducing syntactic structures into target opinion word extraction with deep learning [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Punta Cana: ACL, 2020: 8947-8956.
- [6] CHEN Shaowei, LIU Jie, WANG Yu, et al. Synchronous double-channel recurrent network for aspect-opinion pair extraction [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. ACL, 2020: 6515-6524.
- [7] WANG Yequan, HUANG Minlie, ZHU Xiaoyan, et al. Attention-based LSTM for aspect-level sentiment classification [C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA: ACL, 2016: 606-615.
- [8] ZHANG Y, DING Q, ZHU Z, et al. Enhancing aspect and opinion terms semantic relation for aspect sentiment triplet extraction [J]. Journal of Intelligent Information Systems, 2022, 59: 523-542.
- [9] CHEN F, YANG Z, HUANG Y. A multi-task learning framework

- for end - to - end aspect sentiment triplet extraction [ J ]. *Neurocomputing*, 2022, 479: 12-21.
- [ 10 ] PENG Haiyun, XU Lu, BING Lidong, et al. Knowing what, how and why: A near complete solution for aspect - based sentiment analysis [ C ] // Proceedings of the AAAI Conference on Artificial Intelligence. New York: ACM, 2020, 34 ( 5 ): 8600 - 8607.
- [ 11 ] XU Lu, LI Hao, LU Wei, et al. Position - aware tagging for aspect sentiment triplet extraction [ C ] // Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Punta Cana: ACL, 2020: 2339-2349.
- [ 12 ] ZHANG Chen, LI Qiuchi, SONG Dawei, et al. A multi - task learning framework for opinion triplet extraction [ C ] // Findings of the Association for Computational Linguistics: EMNLP 2020. Punta Cana: ACL, 2020: 819-828.
- [ 13 ] CHEN Hao, ZHAI Zepeng, FENG Fangxiang, et al. Enhanced multi - channel graph convolutional network for aspect sentiment triplet extraction [ C ] // Proceedings of the 60<sup>th</sup> Annual Meeting of the Association for Computational Linguistics ( Volume 1: Long Papers ). Dublin, Ireland: ACL, 2022: 2974-2985.
- [ 14 ] SHI Lingling, HAN Donghong, HAN Jiayi, et al. Dependency graph enhanced interactive attention network for aspect sentiment triplet extraction [ J ]. *Neurocomputing*, 2022, 507: 315-324.
- [ 15 ] CHEN Yuqi, ZHANG Zequn, ZHOU Guangyao, et al. Span - based dual - decoder framework for aspect sentiment triplet extraction [ J ]. *Neurocomputing*, 2022, 492: 211-221.
- [ 16 ] 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: ACL, 2019: 4171 - 4186.
- [ 17 ] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [ C ] // Advances in Neural Information Processing Systems. Long Beach, USA: NIPS Foundation, 2017, 30: 5998-6008.
- [ 18 ] DOZAT T, MANNING C D. Deep biaffine attention for neural dependency parsing [ J ]. arXiv preprint arXiv:1611.01734, 2016.
- [ 19 ] ZHANG Chen, LI Qiuchi, SONG Dawei. 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 9<sup>th</sup> International Joint Conference on Natural Language Processing (EMNLP - IJCNLP). Hong Kong: ACL, 2019: 4568-4578.
- [ 20 ] YAO Liang, MAO Chengsheng, LUO Yuan. Graph convolutional networks for text classification [ C ] // Proceedings of the AAAI Conference on Artificial Intelligence. Hawaii: AAAI, 2019, 33 ( 1 ): 7370-7377.
- [ 21 ] DAI An'an, HU Xiaohui, NIE Jianyun, et al. Learning from word semantics to sentence syntax by graph convolutional networks for aspect - based sentiment analysis [ J ]. *International Journal of Data Science and Analytics*, 2022, 14 ( 1 ): 17-26.
- [ 22 ] WU Zhen, YING Chengcan, ZHAO Fei, et al. Grid tagging scheme for aspect - oriented fine - grained opinion extraction [ C ] // Findings of the Association for Computational Linguistics: EMNLP 2020. Punta Cana: ACL, 2020: 2576-2585.
- [ 23 ] CHEN Shaowei, WANG Yu, LIU Jie, et al. Bidirectional machine reading comprehension for aspect sentiment triplet extraction [ C ] // Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2021, 35 ( 14 ): 12666-12674.
- [ 24 ] LI You, LIN Yongdong, LIN Yuming, et al. A span - sharing joint extraction framework for harvesting aspect sentiment triplets [ J ]. *Knowledge - Based Systems*, 2022, 242: 108366.
- [ 25 ] CHEN Zhixue, HUANG Hong, LIU Bang, et al. Semantic and syntactic enhanced aspect sentiment triplet extraction [ C ] // Findings of the Association for Computational Linguistics: ACL - IJCNLP 2021. ACL, 2021: 1474-1483.