

### 基于依赖类型剪枝的双特征自适应融合网络用于方面级情感分析

郑诚, 石景伟, 魏素华, 程嘉铭

引用本文

郑诚, 石景伟, 魏素华, 程嘉铭. 基于依赖类型剪枝的双特征自适应融合网络用于方面级情感分析[J]. 计算机科学, 2024, 51(3): 205-213.

ZHENG Cheng, SHI Jingwei, WEI Suhua, CHENG Jiaming. [Dual Feature Adaptive Fusion Network Based on Dependency Type Pruning for Aspect-based Sentiment Analysis](#) [J]. Computer Science, 2024, 51(3): 205-213.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

#### [CheatKD:基于毒性神经元同化的知识蒸馏后门攻击方法](#)

CheatKD: Knowledge Distillation Backdoor Attack Method Based on Poisoned Neuronal Assimilation  
计算机科学, 2024, 51(3): 351-359. <https://doi.org/10.11896/jsjcx.221200035>

#### [异质信息网络中基于解耦图神经网络的社区搜索](#)

Community Search Based on Disentangled Graph Neural Network in Heterogeneous Information Networks

计算机科学, 2024, 51(3): 90-101. <https://doi.org/10.11896/jsjcx.221200029>

#### [基于缺失数据的交通速度预测算法](#)

Traffic Speed Forecasting Algorithm Based on Missing Data

计算机科学, 2024, 51(3): 72-80. <https://doi.org/10.11896/jsjcx.230100045>

#### [基于启发式粗化算法的半监督图神经网络的训练加速框架及算法](#)

Framework and Algorithms for Accelerating Training of Semi-supervised Graph Neural Network Based on Heuristic Coarsening Algorithms

计算机科学, 2024, 51(3): 48-55. <https://doi.org/10.11896/jsjcx.221200158>

#### [一种抗屏摄攻击的DCT域深度水印方法](#)

Screen-shooting Resilient DCT Domain Watermarking Method Based on Deep Learning

计算机科学, 2024, 51(2): 343-351. <https://doi.org/10.11896/jsjcx.221200121>

# 基于依赖类型剪枝的双特征自适应融合网络用于方面级情感分析

郑 诚<sup>1,2</sup> 石景伟<sup>1,2</sup> 魏素华<sup>1,2</sup> 程嘉铭<sup>1</sup>

1 安徽大学计算机科学与技术学院 合肥 230601

2 计算智能与信号处理教育部重点实验室(安徽大学) 合肥 230601

**摘要** 现有的模型将基于依赖树的图神经网络用于方面级情感分析,一定程度上提升了模型的性能。然而,由于依赖解析技术的限制,语法解析结果的不精确导致依赖树存在大量噪声,使得模型的性能提升有限。此外,一些句子本身并不符合标准的句法结构。以往的研究以同样的置信度利用句法信息和语义信息,没有充分考虑它们对于确定方面词极性的贡献的不同,导致模型在相应的数据集上性能较差。为了克服这些困难,文中提出了一种基于依赖类型剪枝的双特征自适应融合网络。具体来说,该模型使用一种新型的混合方法,命名为依赖关系类型剪枝和邻接矩阵平滑,来缓解句法解析产生的噪声。此外,该模型通过双特征自适应融合模块充分考虑句子的句法信息的可用程度,以一种更灵活的方式将句法特征和语义特征结合起来用于方面级情感分析。在5个公开可用的数据集上进行广泛的实验,结果证明了该方法明显优于基线模型。

**关键词:** 方面级情感分析;图神经网络;依赖类型剪枝;双特征自适应融合;深度学习;自然语言处理

中图分类号 TP391

## Dual Feature Adaptive Fusion Network Based on Dependency Type Pruning for Aspect-based Sentiment Analysis

ZHENG Cheng<sup>1,2</sup>, SHI Jingwei<sup>1,2</sup>, WEI Suhua<sup>1,2</sup> and CHENG Jiaming<sup>1</sup>

1 School of Computer Science and Technology, Anhui University, Hefei 230601, China

2 Key Laboratory of Intelligent Computing & Signal Processing(Anhui University), Ministry of Education, Hefei 230601, China

**Abstract** Existing models use graph neural network based on dependency trees for aspect-based sentiment analysis, which improves the classification performance of the model to a certain extent. However, due to technical limitations of dependency parsing, the inaccuracy of the dependency parsing results leads to a large amount of noise in the dependency tree, which makes the performance improvement of the model is limited. In addition, some sentences themselves do not conform to the standard syntactic structure. Previous studies utilized syntactic and semantic information with the same confidence level without fully considering the difference in their contributions to determining the polarity of aspect words, resulting in poor model performance on the corresponding datasets. To overcome these challenges, a dual feature adaptive fusion network based on dependency type pruning is proposed in this paper. Specifically, the model uses a novel hybrid approach, named dependency type pruning and adjacency matrix smoothing, to mitigate the noise generated by dependency parsing. In addition, the model fully considers the availability of syntactic information of sentences through a dual feature adaptive fusion module to combine syntactic features and semantic features for aspect-level sentiment analysis in a more flexible way. Extensive experiments on five publicly available datasets demonstrate that the proposed method significantly outperforms baseline models.

**Keywords** Aspect-based sentiment analysis, Graph neural networks, Dependency type pruning, Dual feature adaptive fusion, Deep learning, Natural language processing

## 1 引言

方面级情感分析(Asspect-based Sentiment Analysis, AB-SA)是一项细粒度的情感分析任务,旨在识别出文本中的情感元素(方面词、观点词、方面类别及情感极性),其包含多个子任务,例如方面词提取、方面类别检测、观点词提取、方面词

情感分类、方面类别情感分析<sup>[1-2]</sup>、方面-观点对提取<sup>[3-5]</sup>、方面情感三元组提取<sup>[6-7]</sup>、方面情感四元组预测<sup>[8-9]</sup>等。本文的研究内容是方面词情感分类(Asspect Term Sentiment Classification),即识别给定句子中已经标注出的特定方面词的情感极性(积极、中性或消极)。例如,在“Great food but the service was dreadful”这句话中,包含两个方面,“food”和“service”。

到稿日期:2023-01-06 返修日期:2023-07-13

基金项目:安徽省重点研究与开发计划(202004d07020009)

This work was supported by the Key Research and Development Project of Anhui Province(202004d07020009).

通信作者:郑诚(csahu@126.com)

方面词情感分类的目标就是分别将“food”和“service”的情感极性识别为积极和消极。因此,方面词情感分类中的主要挑战是如何将方面词与其对应的观点表达联系起来,特别是句子中有多个方面词时。

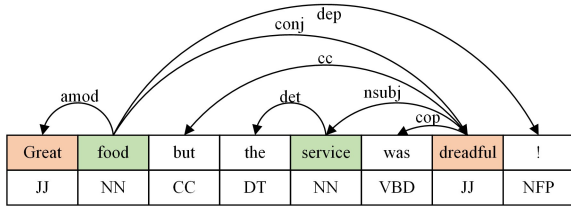


图1 句法解析存在噪声的示例(电子版为彩图)

Fig. 1 Example of syntactic parsing noise

最近的研究通过构造依赖树将句法信息应用到方面词情感分类任务中,成功捕获了句子中单词之间的长期依赖关系。Zhang等<sup>[10]</sup>提出了一种邻域加权卷积网络,通过句法解析器得到的依赖树来提供语法感知的上下文表示。ASGCN<sup>[11]</sup>使用Bi-LSTM学习句子上下文信息,并利用基于句子依赖树的图卷积网络成功地捕获了方面词相关的句法信息。然而,当前的依赖解析方法存在噪声。在这些噪声的干扰下,模型很难正确预测。图1给出了句子“Great food but the service was dreadful”的句法解析依赖树。方面词和观点词分别用绿色和橙色标记,单词下方是单词的词性。单词之间的曲线代表它们之间存在依赖关系,同时标注出了依赖关系的类型。方面词“food”和与其不相关的观点词“dreadful”被直接连接到一起。如果不考虑依赖类型的差异,那么“Great”和“dreadful”都会对方面词“food”的情感极性的判断做出相同的贡献。然而,根据句法和经验知识,对于预测方面词的情感,不同的依赖类型的重要性应该是不同的。之前的工作没有考虑到这点,从而导致预测错误。具体来说,“Great”和“food”之间的依赖类型是“amod”,“amod”为形容词修饰语,用来修饰名词,因此它更有可能将方面词和正确的观点词联系起来。而“dreadful”和“food”之间的依赖类型为“conj”,“conj”为连词,用于连接两个并列的词,它更可能将方面词和不相关的观点词联系起来。因此,当将依赖树应用于方面词情感分类时,句法解析生成的一些依赖类型不仅对方面词的情感极性判断没有帮助,还会干扰模型,致使模型预测错误。Wang等<sup>[12]</sup>提出使用注意力机制来区分依赖关系中重要的部分,虽然取得了一定成功,但是当句子中包含多个方面词时,注意力机制很容易关注与方面词不相关的观点词。

此外,一些句子不符合标准的句法结构,模型只使用句法信息很难预测正确。虽然Chen等<sup>[13]</sup>考虑同时使用句法信息和语义信息,但都是采用同样的置信度来结合句法特征和语义特征,没有充分考虑它们对于确定方面词极性的贡献的不同,导致模型性能不佳。

针对上述两个问题,本文提出了一种新的架构——双特征自适应融合网络(DFAFN),用于方面词情感分类。具体来说,模型使用一种新型的混合方法来缓解依赖解析中的噪声,命名为依赖类型剪枝和邻接矩阵平滑。依赖类型剪枝的目的是根据依赖关系类型对依赖树进行修剪,去除那些对于方面词情感分类没有贡献的依赖类型。邻接矩阵平滑的目的是对

剪枝后的依赖树的邻接矩阵进行平滑处理,避免模型过拟合。去噪后的依赖树有助于模型更加关注方面词真正的观点词。此外,本文设计了一个双特征自适应融合模块,并提出依赖关系置信度(Dependency Confidence Measure, DCM)来衡量句法信息的可用程度。对于不符合标准句法结构的句子,使用自注意力机制进一步学习它的语义特征,最后根据依赖关系置信度将句法特征和语义特征选择性结合起来,用于预测特定方面词的情感极性。

本文的贡献如下:

1)提出了一种新的缓解依赖解析噪声的组合方法,即依赖类型剪枝和邻接矩阵平滑,可以有效地降低依赖树中的噪声。

2)提出了一种新的架构,即DFAFN,用于方面词情感分类。考虑到不同句子的句法信息的可用程度,DFAFN利用依赖关系置信度将句法特征和语义特征自适应结合起来用于方面词情感分类。

3)在SemEval 2014和Twitter数据集上进行了大量的实验,实验结果证明了所提模型的有效性。

## 2 相关工作

随着深度学习技术的快速发展,深度神经网络目前被广泛应用于方面词情感分类中。与传统机器学习方法<sup>[14-16]</sup>手工选取特征不同,深度神经网络可以从大量的数据中自动学习特征表示。目前,方面词情感分类领域中现有的工作根据所使用技术的不同主要可以分为两类。第一类是利用注意力机制<sup>[17]</sup>捕获语义信息,它们使用结合注意力机制的神经网络模型来捕获方面词和上下文单词的语义关系,然后将提取到的特征输入分类器中进行情感分类。ATAE-LSTM<sup>[18]</sup>是一个基于注意力的LSTM模型,它使用目标嵌入,并为每个目标设置一个注意力向量,强制模型关注句子中与特定方面词相关的重要部分。Liu等<sup>[19]</sup>通过区分从给定方面词的左上下文和右上下文获得的注意力来拓展基于注意力的模型。MGAN<sup>[20]</sup>使用多粒度的注意力网络来获取方面词和上下文之间单词级别的交互信息,缓解了粗粒度的注意力在一定情况下会导致信息丢失的问题。Bao等<sup>[21]</sup>发现注意力机制会过度关注句子中的特定部分,而忽略了为判断方面词的情感极性提供关键信息的位置。为此,本文提出了一种利用词典信息增强模型鲁棒性的简单而有效的方法。虽然注意力机制可以关注方面词相关的重要上下文信息,但是当方面词和观点词的距离较远或是句子中包含多个方面词时,基于注意力的模型往往会由于难以关注到重要信息而出现预测错误。

第二类是通过依赖解析挖掘句法信息,句法信息可以辅助模型捕获方面词相关的单词,缩短方面词与观点词之间的距离,可以很好地弥补注意力机制的不足,对于识别方面词的情感极性有着重要的作用。He等<sup>[22]</sup>将句法信息引入注意力机制,显著改进了传统的基于注意力的LSTM。随着依赖解析和图神经网络的发展,基于依赖树的图神经网络在方面词情感分类领域取得了很多成果。CDT<sup>[23]</sup>利用Bi-LSTM来学习句子的特征表示,并使用直接作用于依赖树上的图卷积网络以进一步增强句子嵌入。TD-GAT<sup>[24]</sup>直接从方面词的句法上下文传播情感特征,使观点词更接近方面词,并且解决了

潜在的语法歧义。考虑到之前的研究没有利用依赖树的依赖类型信息,Tian等<sup>[25]</sup>提出了一种类型感知的图卷积网络(T-GCN),使用注意力机制区分依赖树中不同的依赖类型。Hou等<sup>[26]</sup>提出了一种图集成技术 GraphMerge,将多棵依赖树组合起来用于方面词情感分类。DualGCN<sup>[27]</sup>通过 SynGCN 模块和 SemGCN 模块分别整合了句法信息和语义信息,与单独考虑句法信息或语义信息的模型相比取得了更好的效果。

虽然基于依赖树的图神经网络及其变体模型在方面词情感分类领域中取得了较好的效果,但是依赖树中存在的噪声使得模型的性能不能进一步提高。为此,本文提出了一种基于依赖类型剪枝的双特征自适应融合网络(DFAFN),在使用依赖解析器得到依赖树的基础上,该模型根据依赖类型对依赖树进行剪枝,以缓解依赖树的噪声。此外,以往的研究中没有考虑到句子的语义信息和句法信息对于确定方面词极性的贡献的不同,DFAFN设计了一种双特征自适应融合模块,充分考虑了句子的句法信息的可用程度,以一种更灵活、合理的方式将句法特征和语义特征结合起来用于方面词情感分类。

### 3 基于依赖类型剪枝的双特征自适应融合网络

本章将详细介绍所提出的模型 DFAFN。给定一个长度为  $n$  的包含特定的方面词的句子  $s = \{s_1, s_2, s_3, \dots, s_n\}$ , 其中方面词  $a = \{a_1, a_2, a_3, \dots, a_m\}$  为句子中的长度为  $m$  的一段跨度。方面词情感分类的目标是预测方面词  $a$  的情感极性。

DFAFN 的整体架构如图 2 所示。DFAFN 由 4 个模块组成:词嵌入层、特征提取层、双特征自适应融合模块和分类

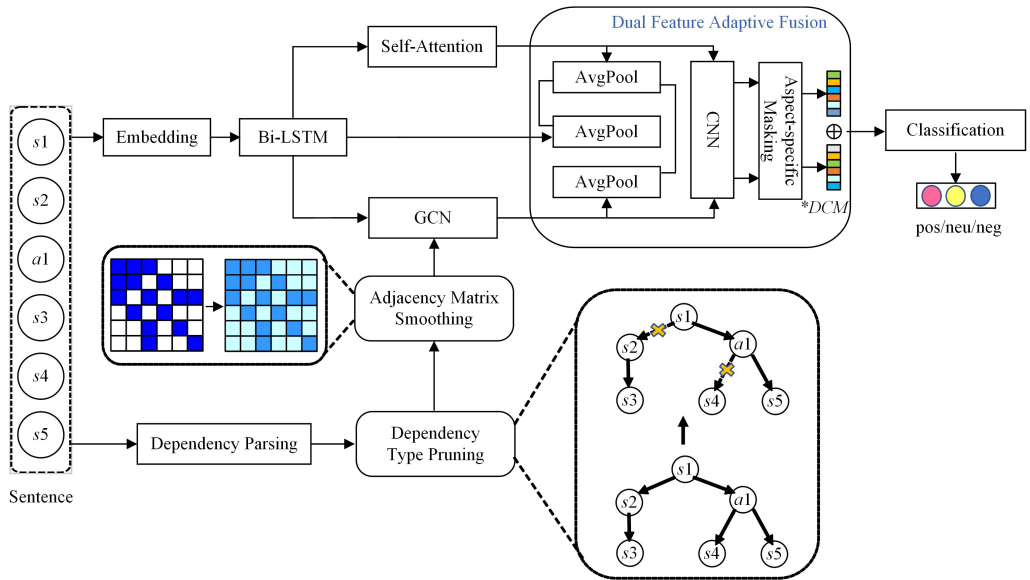


图 2 基于依赖类型修剪的双特征自适应融合网络(DFAFN)的主要架构

Fig. 2 Main structure of DFAFN based on dependency type pruning

#### 3.2 特征提取层

在词嵌入层中,得到了融合了 3 种信息的句子嵌入  $\mathbf{H}$ 。在特征提取层中,自注意力机制和基于依赖树的图卷积网络(Graph Convolutional Networks,GCN)分别被用来提取句子的深层语义特征和与方面词相关的句法特征。

##### 3.2.1 自注意力机制

对于过长的序列,Bi-LSTM 依然无法很好地传输序列

层。词嵌入层用于将输入的句子进行向量嵌入,以获得句子中每个单词的向量表示。特征提取层主要包括自注意力机制和基于依赖树的图卷积网络,用来提取整个句子的句法特征和语义特征。双特征自适应融合模块接收特征提取层输出的句法特征和语义特征作为输入,根据句法信息的可用程度自适应地整合句法特征和语义特征。最后,融合后的特征表示被输入分类层中进行分类,得到情感极性的概率分布。模型的各个模块的介绍如下。

#### 3.1 词嵌入层

为了更好地表示句子的语义信息,本研究使用预训练的词向量 Glove<sup>[28]</sup>,将文本序列  $s$  中的每个单词映射成低维向量。 $\mathbf{W}_{\text{tok}} \in \mathbf{R}^{V_{\text{tok}} \times d_{\text{tok}}}$  是 Glove 的词嵌入矩阵,其中  $d_{\text{tok}}$  是词向量的维度, $V_{\text{tok}}$  是词典中单词的个数。对于  $s$  中的每一个单词,都可以在  $\mathbf{W}_{\text{tok}}$  中按照建立好的索引表找到它们对应的向量嵌入。整个句子的向量表示记为  $\mathbf{e}_{\text{tok}} = \{e_{\text{tok}1}, e_{\text{tok}2}, \dots, e_{\text{tok}n}\}$ 。此外,单词的词性和单词与方面词之间的相对物理位置都对识别方面词的情感极性有着重要的作用。因此,除了词嵌入矩阵  $\mathbf{W}_{\text{tok}}$ ,还随机初始化了词性嵌入矩阵  $\mathbf{W}_{\text{pos}} \in \mathbf{R}^{V_{\text{pos}} \times d_{\text{pos}}}$  和位置嵌入矩阵  $\mathbf{W}_{\text{post}} \in \mathbf{R}^{V_{\text{post}} \times d_{\text{post}}}$ 。然后,根据每个单词的词性和位置信息得到整个句子对应的词性向量表示  $\mathbf{e}_{\text{pos}} = \{e_{\text{pos}1}, e_{\text{pos}2}, \dots, e_{\text{pos}n}\}$  和位置向量表示  $\mathbf{e}_{\text{post}} = \{e_{\text{post}1}, e_{\text{post}2}, \dots, e_{\text{post}n}\}$ 。最后,将这 3 种不同类型的向量拼接起来得到句子嵌入  $\mathbf{E} = \{e_1, e_2, \dots, e_n\}$ 。本文使用 Bi-LSTM<sup>[29]</sup> 来编码句子嵌入,得到结合上下文信息的向量表示  $\mathbf{H} = \{h_1, h_2, \dots, h_n\}$ ,也记作句子的初始特征向量。

起点的信息,但是自注意力机制<sup>[30]</sup>并没有依赖词与词之间的顺序,而是通过计算词与词之间的相似度来挖掘信息,可以更好地提取整个句子的语义信息。将句子的嵌入  $\mathbf{H}$  作为自注意力机制的输入,自注意力机制可以表述为:

$$\mathbf{H}_{\text{atten}} = \text{softmax}\left(\frac{\mathbf{W}_q \mathbf{H} \times (\mathbf{W}_k \mathbf{H})^T}{\sqrt{d}}\right) \mathbf{W}_v \mathbf{H} \quad (1)$$

其中, $\mathbf{W}_q, \mathbf{W}_k$  和  $\mathbf{W}_v$  是可训练的权重矩阵, $d$  是输入特征的

维度,  $\mathbf{H}_{\text{atten}}$  是进一步融合了深层语义信息的特征表示。

### 3.2.2 依赖类型剪枝

句法依赖解析产生噪音的原因是它不是专门用于方面词情感分类的。方面词情感分类重点关注与方面词相关的单词, 而依赖解析会引入与方面词的情感极性无关的依赖类型。当句子中包含多个方面词时, 这些依赖类型还会将方面词和不相关的观点词联系起来, 导致模型预测错误。因此, 需要对依赖解析的结果进行去噪处理。具体来说, 基于语言学和知识, 本研究从依赖树中去除了一些对方面词情感分类有害的依赖类型, 使依赖树更适合方面词情感分类。

首先, 对于每个输入句子, 使用依赖解析器来获得其相应的句法依赖树。然后将依赖类型按照它在句子中承担的作用, 划分为 3 组。第一组  $A$  包含了句法解析的标识符, 例如根节点“root”和标点符号“punct”等依赖类型, 它们对于方面词的极性预测几乎没有影响。第二组  $B$  是具有修饰关系的依赖类型的集合, 这些依赖类型关系会连接名词和与其对应的修饰词, 例如名词主语“nsubj”和形容词修饰语“amod”。实际上, 方面词往往通过这些依赖类型与观点词连接起来。第三组  $C$  是句中用来衔接句子结构的依赖类型的集合, 例如协调连词“cc”和开放从句补语“xcomp”, 用来保持句子的连贯性。表 1 列出了分组后每个组中包含的依赖类型。之后, 构建一个空的冗余依赖类型列表  $L$ 。依赖类型剪枝的策略如下: 首先,  $A$  中的所有依赖类型都被剪枝, 这意味着将  $A$  中全部的依赖类型都添加到  $L$  中。然后, 依次检验  $B$  和  $C$  中的依赖类型。若去掉某一个依赖类型, 模型的性能得到了提升, 那么就将它添加到  $L$  中。此外, 对于句法结构相似的 Restaurant(14Rest) 和 Laptop(14Lap) 数据集, 为了得到对于它们普适且鲁棒的冗余依赖类型, 在对这两个数据集进行修剪时, 只有当模型在 14Rest 和 14Lap 数据集上的性能因删除某个依赖类型而同时得到改善时, 才会将它添加到  $L$  中。这种剪枝策略被记为联合剪枝策略。而对于句法结构混乱的 Twitter 数据集, 则单独对其剪枝, 这种剪枝策略称为独立剪枝策略。不同数据集的冗余依赖关系类型列表  $L$  如表 2 所列。需要注意的是, 由于 15Rest, 16Rest 和 14Rest 数据集来自相同的领域, 且句法结构严谨性相似, 我们直接根据 14Rest 数据集的最优依赖类型列表对 15Rest 和 16Rest 数据集进行剪枝, 以验证剪枝策略和冗余依赖关系类型列表的通用性。

表 1 依赖类型分组情况

Table 1 Groupings of dependency types

Grouping	Dependency type
A	<pad>, <unk>, ROOT, root, punct, dep, nummod, nmod, tmod
B	nsubj, amod, advmod, advcl, neg, acl, relcl, acl, appos, nmod, npmod
C	det, case, nmod, conj, cc, cop, dobj, compound, mark, aux, nmod, poss, xcomp, ccomp, auxpass, nsubjpass, compound, prt, parataxis, mwe, expl, csbj, det, predet, iobj, cc, preconj, discourse, csbjpass

表 2 数据集的冗余依赖项类型列表  $L$

Table 2 Redundant dependency type list  $L$  of datasets

Dataset	$L$
14Rest & 14Lap	<pad>, <unk>, ROOT, root, punct, dep, nummod, nmod, tmod, advcl, conj, parataxis, nmod, auxpass, aux, mark, expl, xcomp
Twitter	<pad>, <unk>, ROOT, root, nummod, nmod, tmod, punct, dep, advcl, acl, appos, nmod, compound, expl, mark, parataxis, ccomp, cc, cc, preconj, auxpass, det, predet

### 3.2.3 邻接矩阵平滑

Liu 等<sup>[31]</sup> 提出标签的不可靠性, 在损失函数中加入了一个标签平滑正则化项, 有效地提高了模型的泛化性能。受这项工作的启发, 本文提出对邻接矩阵进行平滑, 以预防模型因过度信任依赖解析和依赖类型剪枝的结果而过拟合。邻接矩阵平滑的表达式如式(2)所示:

$$A'_{i,j} = \begin{cases} 1 - eps, & \text{if } A_{i,j} = 1 \\ eps, & \text{if } A_{i,j} = 0 \end{cases} \quad (2)$$

其中,  $eps \in [0, 1]$  是邻接矩阵的平滑项,  $\mathbf{A}$  是建立在句子的经过剪枝的句法依赖树上的邻接矩阵。 $\mathbf{A}$  中的元素  $A_{ij}$  代表第  $i$  个节点是否连接到第  $j$  个节点(即第  $i$  个单词和第  $j$  个单词间存在句法关系)。具体来说, 如果第  $i$  个单词和第  $j$  个单词之间存在依赖关系, 则  $A_{ij} = 1$ , 否则  $A_{ij} = 0$ 。 $\mathbf{A}'$  为平滑后的邻接矩阵。值得注意的是, 我们尝试了其他的方法对邻接矩阵进行平滑, 例如考虑单词之间的句法距离信息或是单词节点在依赖图中的度信息。然而, 简单地使用平滑项  $eps$  反而能取得更好的效果。产生这个结果的原因主要在于, 对于句法距离等信息, 很难找到一个完美的表达式来契合邻接矩阵平滑。其次, 与具体的表达式相比, 本文方法通过引入超参数, 给予了模型更多的训练空间, 使得模型的泛化性能更优。

### 3.2.4 图卷积网络

GCN 被用来从去噪后的依赖树中学习节点特征的代表。在图卷积网络层中, 每一个节点的隐藏表示根据它的邻居节点计算而来。对于第  $l$  层的第  $i$  个节点, 其隐藏状态表示形式为  $\mathbf{h}'_i$ , 更新公式如下:

$$\mathbf{h}'_i = \sigma \left( \sum_{j=1}^n A'_{ij} \mathbf{W} \mathbf{h}'_j + \mathbf{b}' \right) \quad (3)$$

其中,  $A'_{ij}$  是平滑后的邻接矩阵  $\mathbf{A}'$  中节点  $i$  和节点  $j$  之间的依赖权重,  $\mathbf{W}$  是可以训练的权重矩阵,  $\mathbf{h}'_j$  是第  $j$  个节点在第  $l-1$  层的隐藏状态表示,  $\mathbf{b}'$  是偏置项,  $\sigma$  是 ReLU 激活函数。

不同于以往的大多数工作中直接将 GCN 第  $l$  层的输出作为第  $l-1$  层的输入。本文将 GCN 第  $l$  层的输出与第  $l$  层的输入的转置相乘, 并使用 softmax 激活函数对其进行处理, 得到特征的重要性分数。然后, 使用得到的分数矩阵对第  $l$  层的输出加权后作为 GCN 的第  $l+1$  层的输入, 如式(4)所示。通过这种处理, 可以有效地缓解 GCN 的过平滑<sup>[32]</sup> 问题。因此, DFAFN 可以通过堆叠 GCN 层来提取节点的深层特征。

$$\mathbf{H}'_{\text{input}}{}^{l+1} = \text{softmax}(\mathbf{H}'_{\text{output}}{}^l \mathbf{H}'_{\text{input}}{}^l) \mathbf{H}'_{\text{output}}{}^l \quad (4)$$

### 3.3 双特征自适应融合模块

对于在特征提取层捕获到的句法特征和语义特征, 以往的工作大多简单地将它们拼接起来作为最后的特征表示, 用于方面词情感分类任务。然而, 不同句子的句法信息和语义

信息对于方面词情感极性的预测的贡献是不同的。因此,在双特征自适应融合模块,考虑到依存关系的可用程度,本文提出了依赖关系置信度(DCM),让模型根据 DCM 自适应地调整句法特征和语义特征融合时所占的比重。

依赖关系置信度是依赖解析结果的可用程度,用来衡量句子中的句法信息对于方面词的情感极性的预测的重要性。

将 Bi-LSTM 得到的原始上下文特征表示记为  $\mathbf{H}_{ori}$ ,自注意力机制提取到的上下文语义特征表示为  $\mathbf{H}_{sem}$ ,图卷积网络提取到的上下文句法特征表示为  $\mathbf{H}_{syn}$ 。为了得到 DCM,对上述 3 种特征向量进行平均池化,分别得到特征表示  $\mathbf{H}_{ori}^{squeeze}$ ,  $\mathbf{H}_{sem}^{squeeze}$  和  $\mathbf{H}_{syn}^{squeeze}$ 。然后,定义特征相似度函数来计算两种特征向量之间的相似度,如式(5)所示:

$$CalFeaSim(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^d (x_i y_i)}{2 \sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} + 0.5 \quad (5)$$

其中,  $x_i$  和  $y_i$  分别是输入的特征向量  $\mathbf{X}$  和  $\mathbf{Y}$  的第  $i$  个元素,  $d$  是特征向量的维度。将  $\mathbf{H}_{ori}$  和  $\mathbf{H}_{sem}$  之间的相似度记作  $Sim4HoHa$ ,  $\mathbf{H}_{sem}$  和  $\mathbf{H}_{syn}$  之间的相似度记作  $Sim4HsHa$ 。最后, DCM 由两种相似度结合得到,作为整合句法特征和语义特征的权重系数。这一过程的表述如下:

$$Sim4HoHa = CalFeaSim(\mathbf{H}_{ori}^{squeeze}, \mathbf{H}_{sem}^{squeeze}) \quad (6)$$

$$Sim4HsHa = CalFeaSim(\mathbf{H}_{syn}^{squeeze}, \mathbf{H}_{sem}^{squeeze}) \quad (7)$$

$$DCM = \sqrt{Sim4HoHa \times Sim4HsHa} \quad (8)$$

得到 DCM 后,依次将  $\mathbf{H}_{sem}$  和  $\mathbf{H}_{syn}$  输入卷积神经网络(Convolutional Neural Networks, CNN)来提取方面词相关的局部信息。利用特定方面词来屏蔽由 CNN 层学习到的输出向量中的非方面词表示,并保持方面词表示不变,分别用  $\tilde{\mathbf{H}}_{sem}$  和  $\tilde{\mathbf{H}}_{syn}$  表示。最后,使用 DCM 将  $\tilde{\mathbf{H}}_{sem}$  和  $\tilde{\mathbf{H}}_{syn}$  拼接起来,得到最终的方面词表示  $\mathbf{H}_{final}$ ,如式(9)所示:

$$\mathbf{H}_{final} = \tilde{\mathbf{H}}_{sem} \oplus DCM * \tilde{\mathbf{H}}_{syn} \quad (9)$$

其中,  $\oplus$  表示向量拼接操作。

### 3.4 分类层

模型使用多层感知机作为情感极性分类器,将最终的特征表示  $\mathbf{H}_{final}$  作为输入,然后经过 softmax 函数处理得到情感极性概率分布  $p$ 。

$$p = \text{softmax}(\mathbf{W}\mathbf{H}_{final} + \mathbf{b}) \quad (10)$$

其中,  $\mathbf{W}$  和  $\mathbf{b}$  分别表示可训练的权重矩阵和偏置项。

### 3.5 模型训练

考虑到样本中存在少量错误标签,本文采用标签平滑来预测模型过拟合。标签平滑即按照式(11)处理原始情感标签编码  $(y_1, y_2, y_3)$ ,平滑后的情感标签编码为  $(\hat{y}_1, \hat{y}_2, \hat{y}_3)$ ,其中  $\mu \in [0, 1]$  为超参数,用来控制标签平滑的力度。

$$\hat{y}_i = (1 - \mu)y_i + \mu(1 - y_i) \quad (11)$$

训练分类器的目标被定义为最小化预测分布和真实分布之间的交叉熵损失。此外,将 1-DCM 作为句法依赖正则化项添加到损失函数中,并引入 L2 正则化项防止模型过拟合。因此,模型的训练目标是 minimized 以下总目标函数:

$$L = - \sum_{j=1}^S \sum_{i=1}^K \hat{y}_i^j \log(p_i^j) + \lambda \| \theta \|_2 + \beta(1 - DCM) \quad (12)$$

其中,  $S$  是训练样本的个数,  $K$  是情感极性类别数,  $\hat{y}$  是样本的真实标签,  $p$  是模型预测的情感概率分布,  $\theta$  代表所有可训练的参数,  $\lambda$  和  $\beta$  是正则化系数。

## 4 实验

### 4.1 数据集

为了评估模型的性能,本文在以下 5 个公开的基准数据集上进行了大量的实验:来自 SemEval 2014 Task 4<sup>[33]</sup> 的 14Rest 和 14Lap 数据集、来自 SemEval 2015 Task 12<sup>[34]</sup> 的 15Rest 数据集、来自 SemEval 2016 Task 5<sup>[35]</sup> 的 16Rest 数据集以及来自 ACL 14 Twitter dataset<sup>[36]</sup> 的 Twitter 数据集,并进行了官方的训练集和测试集分割。这些数据集集中的每个句子都被标注出了方面词和对应的情感极性。去除了数据集中带有“conflict”情感极性标签的句子。这些数据集的统计数据如表 3 所列。

表 3 5 个基准数据集的统计信息

Table 3 Statistics of five benchmark datasets

Dataset		Negative	Positive	Neutral
Twitter	train	1560	1561	3127
	test	173	173	346
14Rest	train	807	2164	637
	test	196	728	196
14Lap	train	870	994	464
	test	128	341	169
15Rest	train	382	1178	50
	test	328	439	35
16Rest	train	709	1620	88
	test	190	597	38

### 4.2 对比实验

为了验证模型的有效性,在上述 5 个数据集上进行了实验,并使用以下主流模型作为基准模型。这些模型的相关介绍如下:

1)ATAE-LSTM 将注意力机制引入 LSTM 中,关注特定方面的上下文信息。

2)IAN<sup>[37]</sup> 通过交互注意力网络学习方面和上下文表示之间的交互关系。

3)AOA-LSTM<sup>[38]</sup> 通过 Attention over Attention(AOA) 模块交互式地学习 Bi-LSTM 生成的隐藏层序列和原始文本序列。

4)PWCN 探索了一种近似加权卷积网络,并使用句法解析器得到句法信息,同时考虑了句法距离和物理距离。实验结果证明了句法信息对于方面级情感分类任务的有效性。

5)ASGCN 利用基于句子依赖树的图卷积网络成功地捕获了句法信息和长距离的单词依赖。

6)BiGCN<sup>[39]</sup> 建立了一种结合层次句法和词汇图的新型网络结构,有效地融合了词对共现信息和句法依赖信息。

7)R-GAT 对依赖解析得到的依赖树进行重塑,并使用图注意力网络(Graph Attention Networks, GAT)编码不同的依赖类型信息。

8)BSSCN 提出了一种类脑语义和句法认知网络,同时考虑了句法信息和语义信息。

9)BERT-SPC<sup>[40]</sup> 是一种基于 Transformer 的双向编码

器,通过获取方面词上下文的深层语义特征进行情感分类。

10)RGAT-BERT与R-GAT的模型架构相同,但词嵌入部分使用预训练语言模型BERT。

#### 4.3 超参数设置

对于所有的实验,使用预训练的300维Glove词向量来初始化单词嵌入。单词词性嵌入和位置嵌入的维度分别为50和100,Bi-LSTM和GCN的隐藏状态向量维度分别为100和200,GCN层数设置为2。为了缓解过拟合,对于输入Bi-LSTM,GCN和注意力模块的单词向量,分别使用0.6,0.2和0.2的dropout。所有的句子都由LAL-Parser句法解析器解析。协调因子 $\beta$ 被设定为0.2,L2正则化项的系数 $\lambda$ 被设定为 $10^{-4}$ ,邻接矩阵平滑项在14Rest,14Lap和Twitter中的取值分别为0.002,0.05,0.1,标签平滑系数 $\mu$ 为0.05,使用学习率为0.003的Adam优化器训练模型。DFAFN模型在30个epochs中训练,批次大小为32。

#### 4.4 主要实验结果

本文使用正确率(Acc)和宏观F1分数(F1)作为模型性能的评价指标,主要实验结果如表4所列。BSSCN在15Rest和16Rest两个数据集上缺少结果的原因是,该模型的作者在论文中没有提供可以用来复现的代码,且该模型只在Twitter,

14Lap和14Rest这3个数据集上进行了实验,由于无法复现,我们将其在论文中的实验结果作为对比结果。与其他模型相比,DFAFN在绝大多数数据集和评价指标上都取得了最好的效果。此外,除了DFAFN,R-GAT也没有使用句法解析得到的完整的依赖树。与其他使用整个依赖树的模型相比,R-GAT在Twitter和14Rest数据集上取得了更好的结果,只有同时考虑了句法信息和语义信息的BSSCN在14Lap数据集上性能更优。因此,对依赖树剪枝有助于模型关注方面词相关的信息。然而,对依赖树进行剪枝,在一定程度上也会损失有用信息。而DFAFN同时考虑了句子的语义信息,通过双特征自适应融合模块更灵活地结合了句法特征和语义特征。与R-GAT或BSSCN相比,DFAFN在14Lap和14Rest数据集上取得了更好的效果,说明在特征融合时考虑将句法信息和语义信息选择性融合起来更有利于ABSA。此外,当使用BERT进行词嵌入时,模型的性能可以得到进一步的提升。分析模型在15Rest和16Rest数据集上的实验结果可以看出,即使是直接根据14Rest的冗余依赖类型列表进行剪枝,模型仍然可以得到很好的泛化性能,证明了设计的剪枝策略及实验得到的冗余依赖类型列表在方面词情感分类领域具有一定的通用性。

表4 模型在5个公开数据集上的实验结果比较

Table 4 Experimental results comparison on five publicly datasets

Models	Twitter		14Lap		14Rest		15Rest		16Rest	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
Models with Glove										
ATAE-LSTM	67.36	66.45	68.57	64.52	76.78	67.42	78.32	60.47	83.78	61.71
IAN	68.56	67.55	71.10	65.93	76.84	68.35	78.67	52.79	84.32	55.02
AOA-LSTM	71.68	69.25	72.62	67.52	79.96	70.40	78.27	57.09	87.37	66.62
PWCN	72.03	69.74	75.78	72.02	80.87	72.19	78.78	61.57	88.32	67.21
ASGCN	72.66	70.53	75.47	71.03	81.27	72.96	78.86	61.89	88.79	67.35
BiGCN	74.16	73.35	74.59	71.84	81.97	73.48	81.86	64.79	88.96	70.84
R-GAT	75.77	73.87	77.26	73.64	83.30	76.12	81.36	64.55	88.98	70.92
BSSCN	75.14	73.27	77.74	74.21	82.59	74.22	—	—	—	—
DFAFN(ours)	76.66	75.31	78.96	75.60	83.56	76.60	82.03	65.89	89.42	71.36
Models with Pre-trained Language Models										
BERT-SPC	75.46	74.21	77.59	73.28	84.11	76.68	83.48	66.18	90.10	74.16
RGAT+BERT	76.35	74.87	78.17	74.27	86.62	81.36	83.27	69.82	89.82	76.65
DFAFN+BERT(ours)	77.21	75.39	79.42	76.82	86.38	81.79	83.77	70.21	90.53	77.56

注:“—”表示实验结果缺少数据。

#### 4.5 消融实验

为了验证DFAFN中不同组件的有效性,本研究进一步进行了消融实验。分别从原始模型中去除依赖类型剪枝(即w/o DTP)、邻接矩阵平滑(即w/o AMS)和依赖关系置信度(即w/o DCM),实验结果如表5所列。一些观察结果如下:首先,从完整的模型中去除任何组件,模型在所有数据集上的分类性能都会下降。这一结果表明,3个组成部分都发挥着重要作用。其次,当模型不使用DCM时,模型在大多数指标上的表现都是最差的,说明了根据句法信息的重要性自适应地融合句法信息和语义信息是对方面词情感分类最重要的改进。此外,去除依赖类型剪枝和邻接矩阵平滑,该模型在Twitter数据集上遭受了最严重的性能下降。该实验结果符合预期,因为Twitter数据集不符合标准的句法结构,经过句法解析得到的依赖树中存在大量的噪声,证明了依赖类型修剪和

邻接矩阵平滑在缓解依赖解析的噪音方面的有效性。

表5 消融实验结果

Table 5 Experimental results of ablation study

Models	Twitter		14Lap		14Rest	
	Acc	F1	Acc	F1	Acc	F1
DFAFN(ori)	76.66	75.31	78.96	75.60	83.56	76.60
w/o DTP	75.48	74.75	78.01	74.72	83.20	76.38
w/o AMS	76.22	74.46	78.32	74.55	83.30	76.08
w/o DCM	75.78	74.68	77.74	74.21	82.75	74.57

注:“w/o”表示去除某一特定的模块。

#### 4.6 不同剪枝策略对14Rest和14Lap数据集的影响

在对依赖树中的依赖类型进行剪枝时,模型根据数据集的句法结构对14Rest和14Lap数据集采取了联合剪枝策略,从而得到了在这两个数据集上普适的冗余依赖关系类型。为了验证联合剪枝策略的正确性,对14Rest和14Lap数据集

分别使用独立剪枝策略。表 6 列出了模型采用两种不同的剪枝策略的对比实验结果。与独立剪枝策略相比,模型使用联合剪枝策略可以得到更好的分类性能。产生这种预期的结果的原因是,在数据集句法结构相似的情况下,使用联合剪枝策略可以综合考虑模型在这两个数据集上的泛化情况,更有利于得到可靠的冗余依赖关系类型列表。

表 6 不同的剪枝策略在 14Rest 和 14Lap 数据集上的实验结果

Table 6 Experimental results of different pruning strategies on

减枝策略	14Rest and 14Lap datasets			
	14Lap		14Rest	
	Acc	F1	Acc	F1
独立剪枝	78.80	75.36	83.38	75.82
联合剪枝	78.96	75.60	83.56	76.62

#### 4.7 数据集的冗余依赖类型分析

表 2 中的实验结果说明,14Rest&14Lap 数据集和 Twitter 数据集都有它们自己对应的最优冗余依赖类型列表。为了进一步研究冗余依赖类型和数据集之间的关系,通过替换 14Rest&14Lap 数据集和 Twitter 数据集的冗余依赖类型列表来进行额外的实验。具体的实验结果如表 7 所列。可以发现,DFAFN 只有根据数据集对应的冗余依赖类型列表进行剪枝才能获得最佳性能。此外,可以看出, Twitter 比 14Rest&14Lap 数据集具有更多的冗余依赖类型,且 Twitter 比 14Rest&14Lap 数据集删除了更多分组 C 中的依赖类型。其次,14Rest&14Lap 和 Twitter 数据集都尽可能避免将分组 B 中的依赖类型添加到冗余依赖类型列表中。产生前一结果主要有两方面的原因。首先,联合剪枝策略比独立剪枝策略

要求严格。其次, Twitter 数据集句法结构混乱,通过删除方面词无关的依赖类型更有利于模型捕获方面词相关的信息,而在 14Rest&14Lap 数据集过度剪枝可能会破坏句子的句法连贯性,从而导致方面词无法和观点词联系起来。后一个结果说明了与其他的依赖类型相比,具有修饰关系的依赖类型对帮助模型确定方面词的情感极性起着更重要的作用。

表 7 数据集使用不同的冗余依赖类型列表 L 的实验结果

Table 7 Experimental results of different redundant dependency

L	type list L on each dataset					
	Twitter		14Lap		14Rest	
	Acc	F1	Acc	F1	Acc	F1
L(14Rest&14Lap)	75.92	74.70	78.96	75.60	83.56	76.60
L(Twitter)	76.66	75.31	78.48	74.70	83.38	76.46

#### 4.8 邻接矩阵可视化

为了更直观地展示依赖类型剪枝和邻接矩阵平滑对邻接矩阵的影响,本文可视化了未去噪和去噪后的依赖树的邻接矩阵。“Great food but the service was dreadful”这句话的邻接矩阵如图 3(a)所示。原始的邻接矩阵由于句法解析结果的不精确而存在噪声,例如“food”和“dreadful”之间的依赖。依赖类型剪枝不仅将这种噪声去除,还去除了那些对于方面词的情感极性预测没有贡献的冗余信息,如图 3(b)所示。此外,为了进一步增加模型的鲁棒性,防止模型过分相信句法解析的结果,我们对邻接矩阵进行了平滑处理。平滑后的邻接矩阵如图 3(c)所示。在消融实验中进行的大量实验证明了依赖类型剪枝和邻接矩阵平滑对缓解句法解析噪声的有效性。

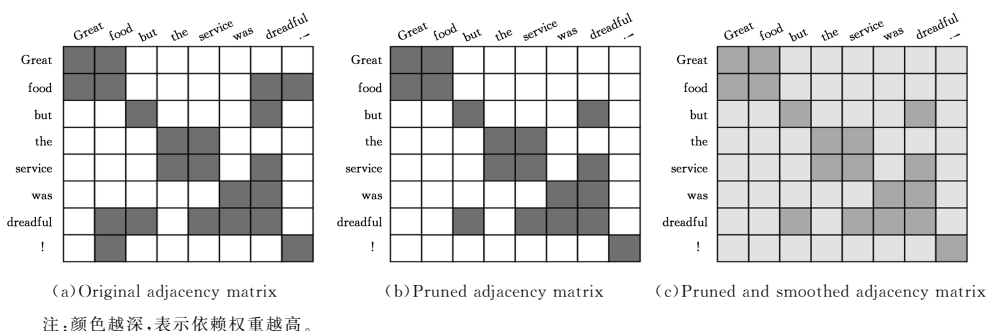


图 3 依赖类型修剪和邻接矩阵平滑对邻接矩阵做出贡献的说明

Fig. 3 Illustration of how dependency type pruning and adjacency matrix smoothing contribute to the adjacency matrix

#### 4.9 个案研究

表 8 列出了使用不同模型分析的几个案例,其中 IAN 是基于注意力的模型,ASGCN 是基于句法依赖的模型。这两个模型分别对应了当前研究工作中的两个研究方向,即利用注意力机制捕获语义信息和通过依赖解析挖掘句法信息,具有代表性。第一个句子中包含了两个情感极性相反的方面词。对于方面词“food”,基于注意力的模型 IAN 错误地关注到“dreadful”这个观点词,因此 IAN 失败。基于依赖的模型 ASGCN 虽然很好地捕捉到了“food”和“Great”之间的依赖关系,但是也错误地引入了“food”和“dreadful”之间的依赖关系,最终,ASGCN 将“food”错误地预测为消极的。然而,DFAFN 通过依赖类型剪枝成功地消除了“food”和“dreadful”之间的依赖类型“conj”,因此,DFAFN 成功预测出“food”的

情感极性。此外,第三个句子的依赖树也存在严重的噪声,方面词“speakers”和不相关的观点词“great”通过依赖类型“parataxis”联系起来。而 DFAFN 仍能很好地应对这种情况。虽然基于句法依赖的模型可以捕获长距离单词依赖关系,但是如果句子中不包含情感极性明显的观点词,它们也很难预测正确。正如第二个句子和第四个句子,ASGCN 失败。而 IAN 却精准地捕获了句子的语义信息。DFAFN 通过双特征自适应融合模块整合了句法特征和语义特征,因此 DFAFN 得到了正确的预测。此外,由于 Twitter 数据集中句子的非正式表达,往往会给方面词情感分类任务带来额外的困难。例如,对最后一句话进行解析时,“hell”会和方面词“britney spears”通过“follow”产生二阶依赖关系,从而导致 ASGCN 失败。由于“hell”和

“follow”之间的依赖关系类型为“dep”，DFAFN 可以通过 依赖类型剪枝去掉这个噪声，从而成功预测。

表 8 个例研究的实验结果  
Table 8 Experimental results of case study

Dataset	Review	Label	Prediction		
			IAN	ASGCN	DFAFN
14Rest	Greatfood but the service was dreadful	(P,N)	(N,N)	(N,N)	(P,N)
	Kenny is always there and he treats my family like we are part of his family.	P	P	O	P
14Lap	Tehfeatures are great,the only thing it needs is better speakers.	(P,N)	(P,P)	(P,P)	(P,N)
	I can barely use anydevices because they will not stay connected properly.	N	N	O	N
Twitter	Why the hell do i followbritney spears on twitter ?	O	O	N	O

注：方面词用下划线标记；表中的 P、N 和 O 分别表示积极、消极和中性；“Prediction”是模型预测的结果；“Label”是方面词的真实标签。

**结束语** 本文提出了一种基于依赖类型剪枝的双特征自适应融合网络架构，即 DFAFN，用于弥补基于句法依赖的方面词情感分类方法的不足。针对以往工作中存在的句法解析噪声问题，DFAFN 结合语言学知识，通过依赖类型剪枝和邻接矩阵平滑来缓解这种噪声。此外，考虑到不同句子中的句法和语义信息对预测方面词的情感极性有着不同的贡献，DFAFN 通过双特征自适应融合模块以更灵活的方式动态融合捕获到的句法特征表示和语义特征表示，并将其用于情感分析。基准数据集上的大量实验证明了所提模型的有效性。虽然本文基于语言学和经验知识制定了适合于方面词情感分类的剪枝策略，但是搜寻冗余依赖类型仍然需要花费大量的时间。尽管如此，我们已经获得方面词情感分类领域中 3 个基准数据集各自最优的冗余依赖类型列表，并通过实验验证了它们的通用性。因此，未来的工作可以直接使用这些冗余依赖类型列表对依赖树进行剪枝。此外，仅仅使用自注意力机制可能无法充分提取方面词相关的语义关系，在未来的工作中，我们将结合领域特定的外部知识来辅助模型提取语义特征。

## 参考文献

- [1] LIU J, TENG Z, CUI L, et al. Solving Aspect Category Sentiment Analysis as a Text Generation Task[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2021: 4406-4416.
- [2] LIN Y, FU Y, LI Y, et al. Aspect-based sentiment analysis for online reviews with hybrid attention networks[J]. World Wide Web, 2021, 24(4): 1215-1233.
- [3] LI Y, WANG C, LIN Y, et al. Span-based relational graph transformer network for aspect-opinion pair extraction[J]. Knowledge and Information Systems, 2022, 64(5): 1305-1322.
- [4] GAO L, WANG Y, LIU T, et al. Question-driven span labeling model for aspect-opinion pair extraction[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021: 12875-12883.
- [5] WU S, FEI H, REN Y, et al. Learn from Syntax: Improving Pair-wise Aspect and Opinion Terms Extraction with Rich Syntactic Knowledge[C]// Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2021: 3957-3963.
- [6] LI Y, LIN Y, LIN Y, et al. A span-sharing joint extraction framework for harvesting aspect sentiment triples[J]. Knowledge-Based Systems, 2022, 242(2): 108366.
- [7] LIU Y, LIU Q, DAI D, et al. Unified Structure Generation for Universal Information Extraction[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2022: 5755-5772.
- [8] MAO Y, SHEN Y, YANG J, et al. Seq2path: Generating sentiment tuples as paths of a tree[C]// Findings of the Association for Computational Linguistics. Stroudsburg: ACL, 2022: 2215-2225.
- [9] BAO X, WANG Z, JIANG X, et al. Aspect-based Sentiment Analysis with Opinion Tree Generation[C]// Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2022: 4044-4050.
- [10] ZHANG C, LI Q, SONG D. Syntax-aware aspect-level sentiment classification with proximity-weighted convolution network [C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and De-velopment in Information Retrieval. New York: ACM, 2019: 1145-1148.
- [11] 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. Stroudsburg: ACL, 2019: 4567-4577.
- [12] 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. Stroudsburg: ACL, 2020: 3229-3238.
- [13] CHEN J, HUANG Z, XUE Y. Bilateral-brain-like Semantic and Syntactic Cognitive Network for Aspect-level Sentiment Analysis[C]// 2021 International Joint Conference on Neural Networks. Piscataway: IEEE, 2021: 1-8.
- [14] PANG B, LEE L. Opinion Mining and Sentiment Analysis[J]. Applied and Environmental Microbiology, 2008, 2(1/2): 1-135.
- [15] KIRITCHENKO S, ZHU X, CHERRY C, et al. NRC-Canada 2014: Detecting Aspects and Sentiment in Customer Reviews [C]// Proceedings of the 8th International Workshop on Semantic Evaluation. Stroudsburg: ACL, 2014: 437-442.
- [16] MARCHEGGIANI D, TÄCKSTRÖM O, ESULI A, et al. Hierarchical multi-label conditional random fields for aspect-oriented opinion mining[C]// European Conference on Information Retrieval. Cham: Springer, 2014: 273-285.
- [17] MNIH V, HEES N, GRAVES A. Recurrent Models of Visual Attention[J]. Advances in Neural Information Processing Systems, 2014, 27(2): 2204-2212.
- [18] 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, Stroudsburg: ACL, 2016: 606-615.
- [19] LIU J, ZHANG Y. Attention modeling for targeted sentiment [C] // Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Stroudsburg: ACL, 2017: 572-577.
- [20] 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, Stroudsburg: ACL, 2018: 3433-3442.
- [21] BAO L, LAMBERT P, BADIA T. Attention and Lexicon Regularized LSTM for Aspect-based Sentiment Analysis [C] // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Stroudsburg: ACL, 2019: 253-259.
- [22] HE R, LEE W, NG H, et al. Effective Attention Modeling for Aspect-Level Sentiment Classification [C] // Proceedings of the 27th International Conference on Computational Linguistics, Stroudsburg: ACL, 2018: 1121-1131.
- [23] 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, Stroudsburg: ACL, 2019: 5678-5687.
- [24] HUANG B, CARLEY K M. Syntax-Aware Aspect Level Sentiment Classification with Graph 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, Stroudsburg: ACL, 2019: 5468-5476.
- [25] 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, Stroudsburg: ACL, 2021: 2910-2922.
- [26] HOU X, QI P, WANG G, et al. Graph Ensemble Learning over Multiple Dependency Trees for Aspect-level Sentiment Classification [C] // Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics, Washington: NAACL, 2021: 2884-2894.
- [27] 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, Stroudsburg: ACL, 2021: 6319-6329.
- [28] 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, Stroudsburg: ACL, 2014: 1532-1543.
- [29] KOTA V R, MUNISAMY S D. High accuracy offering attention mechanisms based deep learning approach using CNN/bi-LSTM for sentiment analysis [J]. International Journal of Intelligent Computing and Cybernetics, 2022, 15(1): 61-74.
- [30] LIANG S, WEI W, MAO X L, et al. BiSyn-GAT+: Bi-Syntax Aware Graph Attention Network for Aspect-based Sentiment Analysis [C] // Findings of the Association for Computational Linguistics: ACL 2022, Stroudsburg: ACL, 2022: 1835-1848.
- [31] LIU B, BEN A, GALDRAN A, et al. The Devil is in the Margin: Margin-based Label Smoothing for Network Calibration [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2022: 80-88.
- [32] LI Q, HAN Z, WU X, et al. Deeper Insights Into Graph Convolutional Networks for Semi-Supervised Learning [C] // Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, Palo Alto: AAAI Press, 2018: 3538-3545.
- [33] 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, Stroudsburg: ACL, 2014: 27-35.
- [34] PONTIKI M, GALANIS D, PAPAGEORGOU H, et al. SemEval-2015 task 12: Aspect based sentiment analysis [C] // Proceedings of the 9th International Workshop on Semantic Evaluation, Stroudsburg: ACL, 2015: 486-495.
- [35] PONTIKI M, GALANIS D, PAPAGEORGOU H, et al. SemEval-2016 task 5: Aspect based sentiment analysis [C] // Proceedings of the 10th International Workshop on Semantic Evaluation, Stroudsburg: ACL, 2016: 19-30.
- [36] 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, Stroudsburg: ACL, 2014: 49-54.
- [37] 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, San Francisco: Morgan Kaufmann, 2017: 4068-4074.
- [38] HUANG B, OU Y, CARLEY K M. Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks [C] // International Conference on Social Computing, Behavioral-cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, Cham: Springer, 2018: 197-206.
- [39] 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, Association for Computational Linguistics, Stroudsburg: ACL, 2020: 3540-3549.
- [40] BURSTEIN J, DORAN C, SOLORIO T, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [C] // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics, Stroudsburg: ACL, 2019: 4171-4186.



**ZHENG Cheng**, born in 1964, Ph.D., associate professor. His main research interests include data mining and text analysis, and natural language processing.