

# 基于动态提示学习和依存关系的生成式结构化情感分析模型

贾银涛, 崔佳佳, 穆玲玲\*, 咎红英

郑州大学计算机与人工智能学院, 河南省郑州市, 450001

ytjia2024@gs.zzu.edu.cn, jjcui@gs.zzu.edu.cn, iellmu@zzu.edu.cn, iehyzan@zzu.edu.cn

## 摘要

结构化情感分析旨在从文本中抽取所有由情感持有者、目标事物、观点表示和情感极性构成的情感元组, 是较为全面的细粒度情感分析任务。针对目前结构化情感分析方法错误传递, 提示模版适应性不足和情感要素构成复杂的问题, 本文提出了基于动态提示学习和依存关系的生成式结构化情感分析模型, 根据不同的情感元组构成情况分别设计提示模版, 并用模板增强生成式预训练模型的输入, 用依存关系增强生成效果。实验结果显示, 本文提出的模型在SemEval2022<sup>1</sup>数据集上的SF1值优于所对比的基线模型。

**关键词:** 结构化情感分析; 提示学习; 依存关系

## Dynamic Prompt Learning and Dependency Relation based Generative Structured Sentiment Analysis Model

Yintao Jia, Jiajia Cui, Lingling Mu\*, Hongying Zan

School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou Henan, 450001  
ytjia2024@gs.zzu.edu.cn, jjcui@gs.zzu.edu.cn, iellmu@zzu.edu.cn, iehyzan@zzu.edu.cn

## Abstract

Structured sentiment analysis aims to extract all sentiment tuples composed of holder, target, expression and polarity from the text, which is a more comprehensive fine-grained sentiment analysis task. Aiming at the problems of error propagation, insufficient adaptability of prompt templates and complex composition of sentiment tuples in current structured sentiment analysis methods, this paper proposes a dynamic prompt learning and dependency relations based generative structured sentiment analysis model. Prompt templates are designed according to different composition of sentiment tuples, and templates are used to enhance the input of the generative pre-trained model. Augment generation with dependency relationships. The experimental results show that the SF1 value of the proposed model on the SemEval2022 dataset is better than the compared baseline models.

**Keywords:** Structured sentiment analysis, Prompt learning, Dependency relations

<sup>1</sup><https://competitions.codalab.org/competitions/33556>

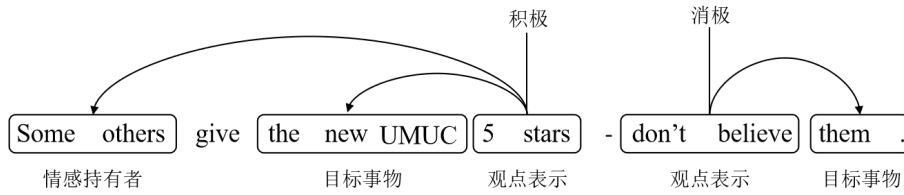


图 1. 结构化情感分析样例

## 1 引言

结构化情感分析(Structured Sentiment Analysis, SSA) (Barnes et al., 2021)旨在抽取文本中所有的情感元组, 情感元组由情感持有者(holder, h)、目标事物(target, t)、观点表示(expression, e)和情感极性(polarity, p)构成, 目的是分析情感持有者对目标事物表达的情感极性, 其中情感极性通过观点表示来表达。结构化情感分析样例如 1所示。相较于方面级情感分析, 结构化情感分析从句子中抽取的情感元素更多更复杂, 并且针对的是目标对象整体评价, 而方面级情感分析针对的是目标对象的方面或属性进行评价。

结构化情感分析在实际应用中具有非常重要的意义, 如在舆情分析方面, 结构化情感分析能够自动识别各类人群对于热点事件的情感表示和情感倾向, 对政府了解民意, 及时做出具体调整具有一定的意义; 在商品和服务评论上, 结构化情感分析能够自动识别各类用户对产品或者产品的某个方面属性所发表的观点及情感倾向, 对消费者挑选产品或服务和企业优化产品都有明确的指导和建议作用。

目前在结构化情感分析中主要的研究方法有图分析方法、序列标注方法和生成式方法。(Barnes et al., 2021)首次提出将结构化情感分析转化为关系图分析任务, 该方法将观点表示作为根节点, 通过与其他元素连接的弧表示它们之间的关系。(Zhai et al., 2023)通过改进情感元素标签和情感元素内部词与词的关系标签解决情感元素在句子中不连续和重叠的问题。序列标注方法通常用在管道方法的第一阶段对情感元素进行抽取, 然后通过第二阶段中的关系判断模块对第一阶段抽取到的情感元素之间的关系进行判断 (Poświata, 2022; Barikbin, 2022; Anantharaman et al., 2022)。生成式方法主要是通过构建提示模板直接将结构化情感分析任务转化为序列到序列的文本生成任务, 目前在结构化情感分析任务中的生成式方法设计的提示模板是构建不连续的文本序列, 模型直接生成情感元素位置索引或者带有特殊标签的情感元素文本(Morio et al., 2022; Raghav et al., 2022)。

上述方法虽然在结构化情感分析任务上取得了一定的成绩, 但是存在以下问题。(1) 图分析方法对关系图的设计有着很高的要求, 关系图设计的好坏对模型整体性能有着决定性的影响; (2) 序列标注方法存在错误传递的问题; (3) 生成式方法中设计的提示模板生成不连续的序列, 对于情感元素之间的关系不能充分表示, 而且对情感元组构成不同的情况使用静态的提示模板, 不利于模型区分不同情况的情感元组。

结构化情感分析任务的数据集主要有 $OpeNER_{EN}$ 、MPQA和 $DS_{Unis}$ 等, 数据集以观点表示为基础标注情感元组, 其中情感持有者相对于观点表示数量很少, 而且目标事物的数量不大于观点表示的数量, 即目标事物可能为空。除此之外,  $OpeNER_{EN}$ 大量句子包含多个情感元组, 多个数据集中标注的情感元素不连续。因此, 数据集中存在的问题是 (1) 完整的情感元组很少; (2) 情感元组的构成情况复杂。

针对上述研究方法的局限性和数据集的问题, 本文提出了基于动态提示学习和依存关系的生成式结构化情感分析模型(dynamic prompt learning and dependency syntactic relations based generative structured sentiment analysis model,  $DPLG_{Dep}$ )。在模型中针对不同情感元组动态使用提示模板, 并且将依存关系融入到生成式预训练模型中提高模型性能。本文的研究工作如下:

- 提出了根据不同情感元组构成设计动态提示模版的思想, 针对完整的情感四元组和情感持

\*通讯作者

项目基金: 期货领域黑嘴检测研究, 横向项目, 郑州易盛信息技术有限公司。

©2024 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

有者为空的情感元组分别设计提示模板将其转化为连续的自然语句形式。

- 提出了用提示模板增强原始输入的方法对预训练模型微调，增强预训练模型在下游任务上的知识。
- 提出了用依存关系与预训练模型融合的方法提高模型对情感元素之间的关系建模，实验结果表明融入依存关系之后，模型性能有所提升，对于依存关系复杂的情感元素的识别能力也得到增强。

## 2 相关工作

**提示学习** 提示学习是将原始输入语句 $x$ 通过使用提示转化为一个提示字符串 $x'$ ，这个提示字符串中有一些空的位置，使用预训练语言模型利用概率去补充这些空置的位置，从而得到最终的输出 $y$  (Liu et al., 2023)。在提示学习中，提示模板的构建是基础的工作，也是本文研究方法中的重要研究内容。提示模板的形式主要包括完形填空式和前缀式，完形填空式提示模板会在提示模板中空出一部分位置，这些位置需要使用预训练模型预测的结果进行填充，这种形式的提示模板常用在信息抽取任务中，如对“Readers think the book is boring”进行情感分析，设计提示“The book is [MASK]”，其中[MASK]就是模型要预测的内容；而前缀式提示模板则是在输入句子和目标答案之间设计一些词或者特殊符号进行提示，在文本生成任务如机器翻译和摘要生成任务中常用，如机器翻译中设计提示模板为“English: [X] Chinese: [Y]”，其中[X]是输入的英文句子，[Y]是模型生成的中文的翻译文本。在本文针对不同的情感元组，设计了对应的完形填空式的提示模板，然后为增强提示效果，将提示模板与输入句子进行拼接作后缀。

**生成式情感分析方法** 生成式情感分析方法通过设计提示模板，将文本中包含的情感元素填充到提示模板对应位置形成目标序列，然后利用生成式预训练模型将情感分析任务转化为文本生成任务。模型输出的目标序列有两种形式：一种是不连续的文本序列形式，如直接生成情感元素位置索引或者带有特殊标签的情感元素文本，另一种是连续的自然语言句子形式的目标序列。不连续的文本序列形式在第1节中已经介绍，下面介绍自然语言句子为目标序列的生成式情感分析方法。(Liu et al., 2021)在方面类型情感分析任务(ACSA)任务中设计自然语言形式的提示模板，将任务转换为自然语言生成任务。(Zhang et al., 2021b)提出直接给生成式预训练模型输入句子中的情感元素添加注释，将经过注释的输入句子作为模型生成的目标句子。(Zhang et al., 2021a)在方面情感四元组预测任务中设计自然语言句子形式的提示模板，不仅可以直接识别四种情感元素，而且也能够充分利用自然语言形式的标签中的语义信息。然而这些生成式方法使用的提示模板为静态模板，本文在(Zhang et al., 2021a)的基础上提出动态提示模板。

**融入句法信息的情感分析方法** 在预训练模型的基础上添加句法或语义信息对于解决结构化情感分析任务是有效的，能够增强预训练模型获取文本句子中结构化信息的能力，提升模型的性能(Sarangi et al., 2022; Yu et al., 2023)。(Sarangi et al., 2022)在结构化情感分析任务中，使用BERT(Devlin et al., 2018)模型获取输入句子的上下文表示，同时使用分析器和神经网络获取句子的语义特征，最后将上下文表示和语义特征进行拼接，这种融合的特征比只使用BERT获取的上下文表示的效果更好。(Yu et al., 2023)在方面类型情感分析任务中首次提出将依存关系融入到生成式预训练模型中，使用图注意力网络(Veličković et al., 2018)获取依存特征表示，同时使用BART(Lewis et al., 2020)的编码器获取输入句子的上下文表示，然后将两种特征采用门控机制进行融合。在本文中同样使用门控机制将依存关系与预训练模型T5(Raffel et al., 2020)进行融合。

## 3 模型

本文设计的 $DPLG_{Dep}$ 模型包括两个模块：基于动态提示学习的生成模块和句法分析模块。在生成模块中根据输入文本句子包含的情感元组使用对应的提示模板，在句法分析模块获取输入文本句子的依存特征，然后使用融合方法将依存特征融合到生成式预训练模型中。

### 3.1 基于动态提示学习的生成模块

生成模块的主体是一个编码器-解码器架构的生成式预训练模型T5，输入序列包括两部分，一部分是原始的输入文本句子，另一部分则是输入文本句子包含的情感元组对应的提示模板。目标序列是情感元素填充到提示模板中对应的位置得到的文本序列。如图2中输入文本句

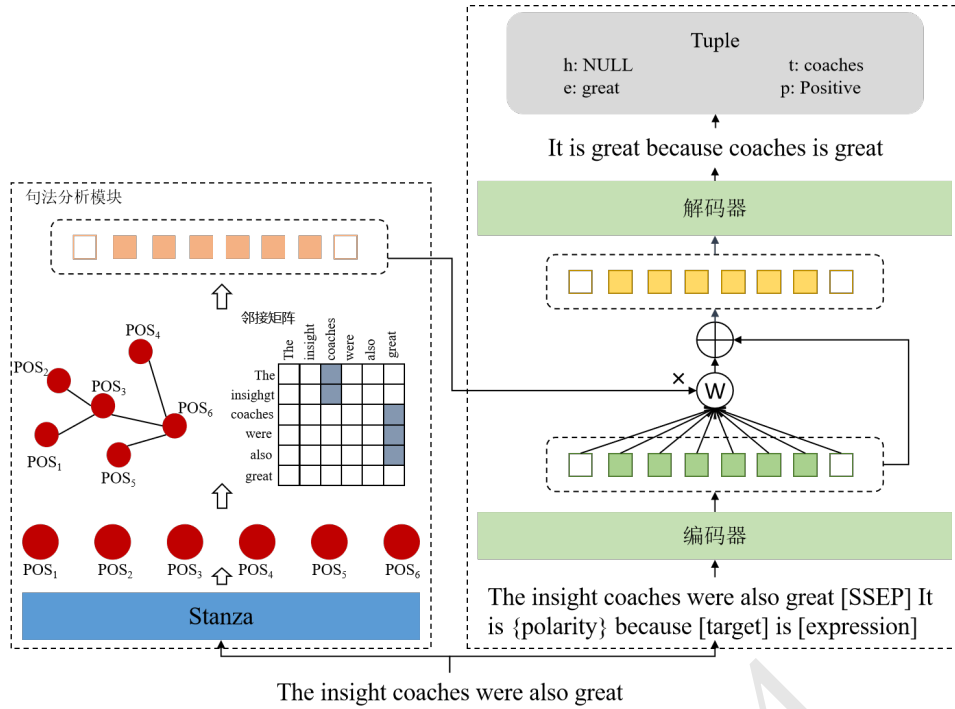


图 2.  $DPLG_{Dep}$  模型

子“The insight coaches were also great”包含情感元组(NULL, “coaches”, “great”, “Positive”), 根据这个情感元组得到对应的提示模板, 将输入句子和提示模板使用 “[SSEP]” 拼接, 得到输入序列。然后将情感元素填充到提示模板中对应的位置得到模型输出的目标序列 “It is great because coaches is great”。

本文提出根据不同情感元组构成设计动态提示模版的思想, 所以当情感元组中四个情感元素都存在时, 设计的提示模板为:

$$[h] \text{ said } [t] \text{ is } [e], [p]$$

其中[h]表示情感持有者, [t]表示目标事物, [e]表示观点表示, [p]表示情感极性对应的词。当情感元组中情感持有者为空时, 设计对应的提示模板为:

$$\text{It is } [p] \text{ because } [t] \text{ is } [e]$$

当情感元组为空时, 使用空字符串 $\emptyset$ 作为提示模板。

本文对目标事物进行判断, 如果目标事物为空则用 “it” 代替, 否则将目标事物填充到模板中:

$$[t] = \begin{cases} [t] & [t] \neq \emptyset \\ it & [t] = \emptyset \end{cases} \quad (1)$$

情感极性分为积极(Positive), 中性(Neutral)和消极(Negative), 直接将情感极性填充到提示模板中不符合自然语言的表达, 而且在不同的提示模板中, 情感极性在句子中充当的语义角色也不同, 所以本文需要将情感极性映射为可以表达对应情感的词, 并且针对不同的情况需要转化成不同的词。当情感元组中情感持有者存在时, 在对应的提示模板中情感极性的映射过程为:

$$[p] = \begin{cases} \text{satisfactorily}, & \text{Positive} \\ \text{objectively}, & \text{Neutral} \\ \text{dissaprovngly}, & \text{Negative} \end{cases} \quad (2)$$

当情感持有者为空时, 对应的提示模板中情感极性的映射过程为:

$$[p] = \begin{cases} \text{great}, & \text{Positive} \\ \text{ok}, & \text{Neutral} \\ \text{bad}, & \text{Negative} \end{cases} \quad (3)$$

如果输入语句中包括多个情感元组，那么给每个情感元组生成对应的目标语句 $y_i$ ，并使用“[SSEP]”将目标语句进行拼接作为预训练的生成模型的输出即 $Target = y_1[SSEP]y_2[SSEP] \dots y_k$ 。

针对数据集中存在的情感元素的短语不连续的问题，本文直接将不连续的短语使用特殊符号“[SEP]”进行拼接，以区分在标准答案中情感元素连续和不连续的情况，优点是更符合标准答案形式，减少在模型训练过程添加额外的噪音。由情感元组线性化为目标句子以及对于不连续情感元素处理的样例如表 1所示。其中输入句子1中包含两种情感元组，第一个情感元组为完整情感元素四元组，第二个为情感持有者为空的情感元组。输入句子2中情感持有者和目标事物都为空。在输入句子3中展示了标注不连续的观点表示[“definitely going”, “again”]，在转化为目标句子时，转化为“definitely going [SEP] again”。

文本类型	文本序列
输入句子1	<b>Clean sizeable room with good facilities</b> in the part of the hotel <b>we were in</b> .
情感元组	h: we, t: facilities, e: good, p: Positive
目标句子	we said facilities is good, satisfactorily
情感元组	h:NULL, t: room, e: Clean sizeable, p: Positive
目标句子	It is great because room is Clean sizeable
输入句子2	<b>Bit pricey</b> and but away from center
情感元组	h: NULL, t: NULL , e: Bit pricey, p: Negative
目标句子	It is bad because it is Bit pricey
输入句子3	<b>I'm definitely going there again</b> whenever I get a chance .
情感元组	h: I, t: there, e: [definitely going, again], p: Positive
目标句子	I said there is <b>definitely going [SEP] again</b> , satisfactorily

表 1. 目标句子构建和非连续情感元素处理

本文考虑只将提示模板经过情感元素填充之后形成目标句子，在这个过程中预训练模型学习到的结构化情感分析的信息不够充分，并且比较难理解情感元素在目标句子中的语义角色，因此，将提示模板添加到输入句子的后面作为后缀，给预训练模型提示更多的信息。

### 3.2 句法分析模块

**依存关系计算：**本文使用邻接矩阵表示依存关系，对输入句子使用句法分析工具Stanza (Qi et al., 2020)进行分析，获取到句子的句法信息包括词性，(依存关系，头结点，尾结点)的依存关系三元组，然后构建依存关系的邻接矩阵 $A \in R^{n \times n}$ ，构造过程如公式4所示，需要注意的是，为符合结构化情感分析任务特征，本文的依存关系保留词在句子中的“root”关系，如图 3，“hotel”就是句子的依存关系的根节点，使用邻接矩阵中的(0, index)表示，其中index是“hotel”在句子中的位置索引，在图 3中index的值为7:

$$A_{i,j} = \begin{cases} 1, & \text{词}_{i,j}\text{之间有依存关系} \\ 0, & \text{词}_{i,j}\text{之间无依存关系} \end{cases} \quad (4)$$

**节点特征表示：**边节点特征使用分析得到的词性表示，根据分析器得到的句子的词性 $P = [pos_1, pos_2, pos_3, \dots, pos_n]$ ，其中 $pos_1$ 前添加一个代表“root”节点对应的词性，如图 3，本文使用“< pad >”作为“root”的词性。然后初始化一个 $Embedding_{pos}$ ，获取词性的特征表示。

$$E^{pos} = Embedding_{pos}(P) \quad (5)$$

其中 $E^{pos} \in R^{n \times d}$ ， $n$ 是预训练模型T5设置的最大长度， $d$ 表示模型T5的维度。

**依存特征的获取：**本文使用图注意网络GAT获取依存特征表示。使用词性嵌入作为GAT模型的节点特征，使用依存关系的邻接矩阵 $A$ 实现掩码注意力。GAT模型获取依存特征表示的计算：

$$h_{l+1} = ELU(a_l h_l W_l) \quad (6)$$

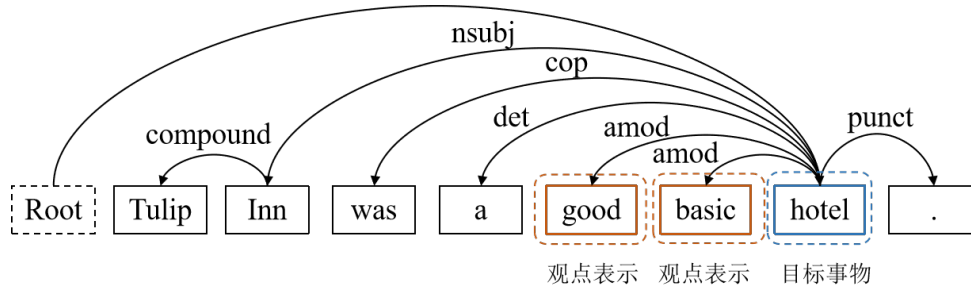


图 3. 依存分析样例

$$a_l = \text{Dropout}(\text{soft max}(e_l)) \quad (7)$$

$$e_l = \text{LeakyRelu}(h_l W_l a[:d, :] + (h_l W_l a[d, :])^T) \quad (8)$$

其中ELU为激活函数， $h_l \in R^{n \times d}$ 为第 $l$ 层句子的词性特征表示， $W \in R^{d \times d}$ ， $a \in R^{2d \times 1}$ ， $h_0 = E^{pos}$ ， $l$ 设置为2。将GAT模型最后一层隐状态表示作为依存特征表示 $E^{dep}$ 。

预训练模型T5模型编码器的输出包括对输入句子 $x$ 和提示模板 $Template$ 拼接之后的序列的输出，所以编码器输出的语义特征 $O_{encoder}$ 为：

$$O_{encoder} = \text{Encoder}_{T5}(x + [SSEP] + Template) \quad (9)$$

本文为了能够使得获取的依存特征与T5模型的编码器的输出 $O_{encoder}$ 在形式上尽可能的一致，在构建依存关系的邻接矩阵时对其边长进行填充和裁剪，使得邻接矩阵的形状固定，填充采用的T5中tokenizer的“< pad >”特殊字符，并且在 $E^{dep}$ 开始和结尾添加T5模型中表示序列开始和结束的特殊符号的嵌入，在T5-base模型中，开始符号默认为“< pad >”，结束符号默认为“< /s >”。所以 $E^{dep}$ 最终还需进行转换：

$$E^{dep} = \text{Embedding}_{T5}(\langle \text{pad} \rangle) || E^{dep}[: -2, :] || \text{Embedding}_{T5}(\langle \text{/s} \rangle) \quad (10)$$

### 3.3 特征融合方式

本文在将依存关系特征表示 $E^{dep}$ 与T5的输出 $O_{encoder}$ 进行融合时设计了四种融合方式，包括直接相加，拼接，门控机制 (Yu et al., 2023)和交叉注意力。

**直接相加：**由于在 $E^{dep}$ 已经将开始和结束位置替换为T5的开始和结束位置，所以直接相加在理论上可行的，计算方式为：

$$H = O_{encoder} + E^{dep} \quad (11)$$

**拼接：**直接对 $O_{encoder}$ 和 $E^{dep}$ 进行拼接，然后经过全连接层，对拼接之后的表示进行维度的转换以及获取拼接之后更重要的信息。

$$H = [O_{encoder} || E^{dep}] W^{trans} + b^{trans} \quad (12)$$

**门控机制：**这种方法能够根据编码器输出的上下文表示动态的调整依存特征的权重，具体就是使用一个线性网络对 $O_{encoder}$ 进行维度转换，然后使用激活函数获取到权重值，对依存特征加权，然后与 $O_{encoder}$ 做残差连接。

$$H = O_{encoder} + W^O E^{dep} \quad (13)$$

其中 $W \in R^{d \times 1}$ 。

**交叉注意力：**本文考虑在将依存特征 $E^{dep}$ 与编码器输出的上下文表示 $O_{encoder}$ 进行融合时， $E^{dep}$ 和 $O_{encoder}$ 的权重都应该是动态调整的，但是经过交叉注意力之后输出的特征表示与 $O_{encoder}$ 存在较大差距，所以需要额外使用一个上下文表示进行求和，做残差连接，所以在交叉注意力中，将 $O_{encoder}$ 作为查询， $E^{dep}$ 作为键和值。

$$H = O_{encoder} + \text{Att} * E^{dep} \quad (14)$$

$$\text{Att} = \text{softmax}\left(\frac{O_{encoder} * (E^{dep})^T}{\sqrt{d}}\right) \quad (15)$$

### 3.4 训练目标

将融合之后的特征用 $e$ 表示，而解码器之前生成的结果用 $y_{<i}$ 表示，解码器解码当前输出的概率分布 $y_i$ 需要基于 $e$ 和 $y_{<i}$ 进行计算：

$$y_i = \text{Decoder}_{T5}(e, y_{<i}) \quad (16)$$

模型训练时采用最大似然估计，对T5模型的输入序列和输出序列进行映射训练更新两个模块的模型参数：

$$\max_{\theta} \log f_{\theta}(y|e) = \sum_{i=1}^n \log f_{\theta}(y_i|e, y_{<i}) \quad (17)$$

## 4 实验设置

### 4.1 数据集

本文在 $OpeNER_{EN}$  (Agerri et al., 2013)、MPQA (Wiebe et al., 2005)和 $DSUnis$  (Toprak et al., 2010)数据集上进行实验。三个数据集的统计信息如表 2所示。

数据集	句子数	情感持有者	目标事物	观点表示	情感极性	
$DSUnis$	训练集	2253	63	806	806	806
	验证集	232	9	98	98	98
	测试集	318	12	130	130	130
MPQA	训练集	5873	1425	1481	1706	1706
	验证集	2063	406	494	570	570
	测试集	2112	434	462	518	518
$OpeNER_{EN}$	训练集	1744	266	2679	2884	2884
	验证集	249	49	371	400	400
	测试集	499	98	793	865	865

表 2. 数据集统计信息

### 4.2 参数设置

$DPLG_{Dep}$ 使用的生成式预训练模型为T5-base，本文训练时batch size设置为16，epoch设置为20，学习率设置为 $3e-4$ ，设置的输入句子的最大长度为128，模型维度为768，预测时采用贪心解码策略。在句法分析模块实现中，包括邻接矩阵的构建、节点特征的设置、依存特征的捕获以及特征融合方式的设置，模型在NVIDIA RTX3090显卡上加速训练，参数信息如表 3所示。

参数类型	参数值
预训练模型	T5-base
输入最大长度	128
预训练模型维度	768
邻接矩阵大小	$128 \times 128$
节点特征维度	768
GAT层数	2
GAT的dropout	0.5
交叉注意力头数	8

表 3. 参数设置

### 4.3 基线模型

与本文的模型 $DPLG_{Dep}$ 进行对比的模型有图分析模型，序列标注模型和生成式模型。

**图分析模型：**Graph-parser (Barnes et al., 2021)是结构化情感分析任务的基线模型，USSA (Zhai et al., 2023)在Graph-parser的基础上改进以解决数据集中情感元素不连续问题。

**序列标注模型：**SeqL (Anantharaman et al., 2022)仅通过双层BiLSTM (Graves and Graves, 2012)模型实现序列标注和情感元素关系判断的模型，TBST (Poświata, 2022)首先使用基于Transformer (Vaswani et al., 2017)的模型实现序列标注，然后使用R-BERT<sup>1</sup>模型判断情感元素之间的关系。

**生成式模型：**GenA (Raghav et al., 2022)直接使用BART模型从输入句子生成情感元素索引，SeqGen (Morio et al., 2022)使用T5-large模型从输入句子生成带有标签的情感元素，以及直接使用大语言模型Llama3<sup>2</sup>预测的方法。

### 4.4 评价指标

在结构化情感分析任务中，为了评估模型预测的情感元组的性能，提出使用情感图F1(Sentiment Graph F1, **SF1**) (Samuel et al., 2022)作为评估指标，在SF1的计算中，一个正确的样例就是情感元素完全匹配，否则计算预测的情感元素与标准答案的情感元素之间重叠的词数权重，然后取情感持有者、目标事物和观点表示这三个情感元素权重的平均值作为当前情感元组的权重值。

## 5 结果及分析

### 5.1 模型结果

模型	类型	数据集		
		OpENER <sub>EN</sub>	MPQA	DS <sub>Unis</sub>
Graph-parser	图分析方法	52.1	12.5	20.4
USSA		-	30.5	33.2
SeqL	序列标注方法	32.9	0.02	0.06
TBST		67.0	32.6	39.5
GenA	生成式方法	62.6	32.7	33.0
SeqGen		<u>67.1</u>	<u>40.9</u>	<u>40.5</u>
Llama3		18.3	0.1	4.9
$DPLG_{Dep}$		<b>70.4</b>	<b>43.3</b>	<b>51.6</b>

表 4.  $DPLG_{Dep}$ 模型结果(SF1, %)

表 4中展示了 $DPLG_{Dep}$ 模型与其他方法的对比结果， $DPLG_{Dep}$ 模型的结果是在5个随机种子的平均值。从表中可以看到，本文提出的模型在所有数据集上的效果达到了最佳，并且生成式方法相较于图分析方法和序列标注方法都有更好的性能。在图分析方法中可以看到不同的关系图的设计会导致模型的性能会有很大的差距，其中USSA方法在解决数据集中情感元素不连续和重叠的问题，相较于Graph-parser方法提升效果明显。在序列标注方法中，可以看到仅仅使用BiLSTM作为节点抽取和关系判断在结构化情感分析任务中的效果不理想，而通过使用预训练模型可以极大地提高模型的性能。而在生成式方法中，相较于GenA和SeqGen模型中使用静态的提示模板生成不连续的目标序列，本文提出的 $DPLG_{Dep}$ 模型效果更好，证明使用动态提示模板，并且用提示模板对原始句增强可以使得T5模型学习到更多的结构化情感分析的知识，而SeqGen方法中使用的模型为T5-large，相较于本文方法使用的T5-base，耗费的资源更多。而没有微调直接使用Llama3进行预测的效果比较差。



模型	特征融合方式	数据集			平均值
		OpeNER <sub>EN</sub>	MPQA	DS <sub>Unis</sub>	
DPLG	-	69.6	42.1	<b>52.2</b>	<u>54.6</u>
<i>DPLG<sub>Dep</sub></i>	直接相加	0.2	0.0	0.0	0.1
	拼接	2.9	0.4	1.3	1.5
	交叉注意力	<u>69.8</u>	<u>43.0</u>	49.3	54.0
	门控机制	<b>70.4</b>	<b>43.3</b>	<u>51.6</u>	<b>55.1</b>

表 5. 特征融合方式结果(SF1, %)

## 5.2 特征融合方式结果

本文针对将依存关系特征与编码器输出的上下文表示进行融合的方式也进行了对比验证, 实验结果如表 5 所示。其中DPLG模型表示*DPLG<sub>Dep</sub>*消去句法分析模块的模型。由表可知, 当使用门控机制时模型的性能最好, 说明使用门控机制将依存特征融合到T5-base模型时对模型整体框架产生的影响最小。使用直接相加和拼接方式时模型的性能剧烈下降, 在T5模型中已经包含丰富的语义信息, 依存信息中可能包含与其重叠或者冲突的信息, 所以在直接进行拼接或者相加时, 模型需要处理这些冗余或者冲突, 可能导致信息混淆。

## 5.3 消融实验结果

模型	数据集			平均值
	OpeNER <sub>EN</sub>	MPQA	DS <sub>Unis</sub>	
DPLG-suffix	65.0	33.0	39.3	45.8
DPLG	69.6	42.1	<b>52.2</b>	<u>54.6</u>
<i>DPLG<sub>Dep</sub></i>	<b>70.4</b>	<b>43.3</b>	<u>51.6</u>	<b>55.1</b>

表 6. 消融实验(SF1, %)

本文对*DPLG<sub>Dep</sub>*模型进行了消融实验, 验证各个模块的作用, 其中DPLG-suffix模型表示在*DPLG<sub>Dep</sub>*的基础上, 不加依存特征, 同时在生成式模型的输入序列中不加提示模板作为后缀。从表中可以看到, 在将提示模板添加到输入句子中时, 在三个数据集上的性能都有很高的提升, 说明将提示模板拼接到生成式预训练模型的输入序列中可以提示模型捕获结构化情感分析任务的信息, 帮助模型理解各个情感元素在目标句子中的语义角色。而在DPLG的基础上融入依存关系特征对模型性能有所提升, 证明仅仅通过T5模型的注意力识别情感元素的性能有待提高, 并且这也是生成式预训练模型的预训练目标与结构化情感分析任务的差异, 而融入句法依存关系能够减小这种差异, 提高模型对情感元素之间关系的建模能力。

## 5.4 案例分析

在本节对*DPLG<sub>Dep</sub>*预测的结果与DPLG预测的结果进行对比, 如表 7 所示, *DPLG<sub>Dep</sub>*模型相较于DPLG, 可以对情感元素根据依存关系进行纠正以及帮助预训练模型捕获模型注意力忽略的句法关系。在样例1中, 使用DPLG没有捕获目标事物“it”, 而添加依存信息之后, 能够捕获到“really great”所修饰的“it”元素。在样例2中, DPLG模型已经准确的预测到第一个情感元组, 但是预测第二个情感元组时模型的注意力更多的集中在“everything”上, 而将“had a great time!”作为修饰语, 根据依存信息, 其中关系更强的是“great”对于“time”的修饰, 所以*DPLG<sub>Dep</sub>*可以准确的预测到两个情感元组。在样例3中, DPLG模型本身强大的性能已经可以准确的预测到其中的三个情感元组, *DPLG<sub>Dep</sub>*中的句法分析模块能够帮助预训练模型捕获“Clean”与“Hotel”的关系。综上可知, 在DPLG的基础上添加句法分析模块, 能够利用依存信息对预训练模型的注意力捕获到一些存在偏差的情感元素进行纠正, 还可以使用依存关系对模型忽略的情感元组关系进行补充, 帮助模型对情感元素之间的关系进行建模。

<sup>1</sup><https://github.com/monologg/R-BERT>

<sup>2</sup><https://github.com/meta-llama/llama3>

样例来源	样例文本
标准样例1	For people who want some rest and relaxatio <b>it is really great</b> . [(NULL, it, really great, Positive)]
DPLG	For people who want some rest and relaxatio it is <b>really great</b> . [(NULL, NULL, really great, Positive)]
$DPLG_{Dep}$	For people who want some rest and relaxatio <b>it is really great</b> . [(NULL, it, really great, Positive)]
标准样例2	<b>everything was quite good</b> , had a <b>great time</b> ! [(NULL, everything, quite good, Positive), (NULL, time, great, Positive)]
DPLG	<b>everything was quite good</b> ,had a <b>great time</b> ! [(NULL, everything, quite good, Positive), (NULL, everything, had a great time, Positive)]
$DPLG_{Dep}$	<b>everything was quite good</b> , had a <b>great time</b> ! [(NULL, everything, quite good, Positive), (NULL, time, great, Positive)]
标准样例3	<b>Wonderful Hotel</b> , <b>Clean Friendly very spacious</b> [(NULL, Hotel, Wonderful, Positive), (NULL, Hotel, Clean, Positive), (NULL, Hotel, Friendly, Positive) , (NULL, Hotel, very spacious, Positive)]
DPLG	<b>Wonderful Hotel</b> , Clean <b>Friendly very spacious</b> [(NULL, Hotel, Wonderful, Positive), (NULL, Hotel, Clean, Positive), (NULL, Hotel, Friendly, Positive) , (NULL, Hotel, very spacious, Positive)]
$DPLG_{Dep}$	<b>Wonderful Hotel</b> , <b>Clean Friendly very spacious</b> [(NULL, Hotel, Wonderful, Positive), (NULL, Hotel, Clean, Positive), (NULL, Hotel, Friendly, Positive) , (NULL, Hotel, very spacious, Positive)]

表 7.  $DPLG_{Dep}$ 与DPLG预测案例对比

## 5.5 模型泛化性验证

为验证本文设计的模型的泛化性, 额外增加了数据集进行实验, 但是包含情感持有者、目标事物、观点表示和情感极性的英文数据集只有 $OpenNER_{EN}$ 、MPQA和 $DS_{Unis}$ , 而这些数据集有一个共同特征, 就是大多数情感持有者为空, 所以本文在包含目标事物、观点表示和情感极性的ASTE-Data-V2数据集 (Xu et al., 2020)上进行额外的实验, 对模板泛化性进行验证。

**基线模型:** 与 $DPLG_{Dep}$ 对比的基线模型有Paraphrase(Zhang et al., 2021a)、BART-ABSA(Yan et al., 2021)、Span-BiDir(Chen et al., 2022)和CONTRASTE(Mukherjee et al., 2023)。

模型	14Lap	14Res	15Res	16Res
BART-ABSA	58.7	65.3	59.3	67.6
Paraphrase	60.4	71.5	62.1	71.9
Span-BiDir	62.7	<b>74.3</b>	<u>64.8</u>	72.1
CONTRASTE	<u>62.9</u>	<u>74.0</u>	<b>66.1</b>	<b>74.2</b>
$DPLG_{Dep}$	<b>66.2</b>	73.1	<u>64.8</u>	<u>73.9</u>

表 8. 模型在ASTE-Data-V2数据集的结果对比 (F1, %)

从表8中可以看到, 相较于当前在ASTE-Data-V2数据集上的先进模型,  $DPLG_{Dep}$ 也表现出具有竞争性的性能, 所以本文的模型具有一定的泛化能力。

## 6 总结与展望

针对当前结构化情感分析任务中研究方法的局限性和数据集的问题, 本文提出了基于动态提示学习和依存关系的生成式结构化情感分析模型 $DPLG_{Dep}$ , 在生成式模块中根据情感元组的构成情况动态地设计提示模板, 并将提示模板添加到输入句子的后面作为后缀给预训练模型T5提示更多的信息。为了减小生成式预训练模型T5的预训练目标与结构化情感分析

任务之间的差异和增强模型建模情感元素之间关系的能力，在生成式模块外添加句法分析模块，将句法分析模块获取的依存特征融合到预训练模型中，模型性能得到提升。实验结果显示， $DPLG_{Dep}$ 模型的性能优于所有对比的基线模型，并且通过案例分析发现，添加句法分析模块，能够利用依存信息对预训练模型的注意力捕获到一些存在偏差的情感元素进行纠正，还可以使用依存关系对模型忽略的情感元组关系进行补充。

## 参考文献

- Rodrigo Agerri, Montse Cuadros, Sean Gaines, and German Rigau. 2013. Opener: Open polarity enhanced named entity recognition. *Procesamiento del Lenguaje Natural*, (51):215–218.
- Karun Anantharaman, K Divyasri, Jayanthan Pt, S Angel, Rajalakshmi Sivanaiah, Sakaya Milton Rajendram, and TT Mirnalinee. 2022. Ssn\_mrlg1 at semeval-2022 task 10: Structured sentiment analysis using 2-layer bilstm. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1324–1328.
- Sadrodin Barikbin. 2022. Slpl-sentiment at semeval-2022 task 10: Making use of pre-trained model’s attention values in structured sentiment analysis. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1382–1388.
- Jeremy Barnes, Robin Kurtz, Stephan Oepen, Lilja Øvrelid, and Erik Velldal. 2021. Structured sentiment analysis as dependency graph parsing. In *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)*, pages 3387–3402.
- Yuqi Chen, Chen Keming, Xian Sun, and Zequn Zhang. 2022. A span-level bidirectional network for aspect sentiment triplet extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 4300–4309.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Alex Graves and Alex Graves. 2012. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, pages 37–45.
- Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.
- Jian Liu, Zhiyang Teng, Leyang Cui, Hanmeng Liu, and Yue Zhang. 2021. Solving aspect category sentiment analysis as a text generation task. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4406–4416.
- Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35.
- Gaku Morio, Hiroaki Ozaki, Atsuki Yamaguchi, and Yasuhiro Sogawa. 2022. Hitachi at semeval-2022 task 10: Comparing graph-and seq2seq-based models highlights difficulty in structured sentiment analysis. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1349–1359.
- Rajdeep Mukherjee, Nithish Kannan, Saurabh Pandey, and Pawan Goyal. 2023. Contrast: Supervised contrastive pre-training with aspect-based prompts for aspect sentiment triplet extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12065–12080.
- Rafał Poświata. 2022. Opi at semeval-2022 task 10: Transformer-based sequence tagging with relation classification for structured sentiment analysis. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1366–1372.
- Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D Manning. 2020. Stanza: A python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 101–108.

- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140):1–67.
- R Raghav, Adarsh Vemali, and Rajdeep Mukherjee. 2022. Etms@ iitkgp at semeval-2022 task 10: Structured sentiment analysis using a generative approach. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1373–1381.
- David Samuel, Jeremy Barnes, Robin Kurtz, Stephan Oepen, Lilja Øvrelid, and Erik Velldal. 2022. Direct parsing to sentiment graphs. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 470–478.
- Pratyush Sarangi, Shamika Ganesan, Piyush Arora, and Salil Joshi. 2022. Amex ai labs at semeval-2022 task 10: contextualized fine-tuning of bert for structured sentiment analysis. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1296–1304.
- Cigdem Toprak, Niklas Jakob, and Iryna Gurevych. 2010. Sentence and expression level annotation of opinions in user-generated discourse. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 575–584.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph attention networks. In *International Conference on Learning Representations*.
- Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39:165–210.
- Lu Xu, Hao Li, Wei Lu, and Lidong Bing. 2020. Position-aware tagging for aspect sentiment triplet extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2339–2349.
- Hang Yan, Junqi Dai, Tuo Ji, Xipeng Qiu, and Zheng Zhang. 2021. A unified generative framework for aspect-based sentiment analysis. In *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)*, pages 2416–2429.
- Chengze Yu, Taiqiang Wu, Jiayi Li, Xingyu Bai, and Yujiu Yang. 2023. Syngen: A syntactic plug-and-play module for generative aspect-based sentiment analysis. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE.
- Zepeng Zhai, Hao Chen, Ruifan Li, and Xiaojie Wang. 2023. Ussa: A unified table filling scheme for structured sentiment analysis. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14340–14353.
- Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing, and Wai Lam. 2021a. Aspect sentiment quad prediction as paraphrase generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9209–9219.
- Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2021b. Towards generative aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 504–510.