

基于动态前缀提示及数据增强的情感 四元组提取方法

钟 将 刘雨轩 戴启祝 王佳祺 赖心怡 胡雯月

(重庆大学计算机学院 重庆 401331)

摘 要 在方面级情感分析 (Aspect-Based Sentiment Analysis, ABSA) 中,情感四元组提取是一个能全面分析情感且最具挑战性的任务。当前基于生成式的方法存在两方面局限性:(1)依赖于提示设计,无法针对任务动态优化,导致提示次优的问题;(2)未能充分解决隐含情感数据不平衡的问题,导致在处理这类数据时性能不佳。为解决这些问题,本文提出了一种动态前缀提示方法 (Dynamic Prefix Prompt),该方法利用可调整的前缀和注意力机制来动态优化提示。此外,本文设计了一种基于大语言模型的数据增强策略,该策略通过微调的方式来对齐数据扩充任务以平衡隐含情感数据。在两个真实应用的数据集上的实验表明,本文所提出的方法在 Restaurants-ACOS 和 Laptop-ACOS 数据集上 F1 分数分别提升 3.60 和 2.20,同时在隐含情感数据中 F1 分数平均提升了 4.23 和 4.67,达到目前最先进的水平,验证了本文方法的有效性和优越性。

关键词 方面级情感分析;情感四元组提取;动态前缀提示;隐含情感数据;大语言模型;数据增强

中图法分类号 TP18

DOI 号 10.11897/SP.J.1016.2025.01082

Sentiment Quadruple Extraction Based on Dynamic Prefix Prompts and Data Augmentation

ZHONG Jiang LIU Yu-Xuan DAI Qi-Zhu WANG Jia-Qi LAI Xin-Yi HU Wen-Yue

(College of Computer Science, Chongqing University, Chongqing 401331)

Abstract In Aspect-Based Sentiment Analysis (ABSA), sentiment quadruple extraction aims to extract quadruples consisting of Aspect, Category, Opinion, and Sentiment from review texts. This task, which enables comprehensive parsing of fine-grained sentiment information, is currently highly challenging. A series of prompt generation approaches have recently attracted much attention. These approaches reconfigure the traditional annotation task into a generative task, which can utilize the rich semantic labels in the text and avoid the challenge of over-focusing on multi-task interactions. Existing prompt-generation-based methods have two main limitations. Firstly, they depend on manually designed discrete prompt templates, which makes it hard to dynamically optimize task-related features, thus resulting in suboptimal prompts. Secondly, the high cost of annotating implicit sentiment data leads to its low proportion in datasets. Consequently, sentiment analysis data often exhibits long-tailed distribution, causing existing models to struggle with implicit sentiment in reviews and suffer from low accuracy. To address these issues, we propose a sentiment quadruple extraction method combining dynamic prefix prompts and data augmentation. For the suboptimal prompt problem, we design a generative-model-based

收稿日期:2024-06-27;在线发布日期:2025-03-14。本课题得到国家自然科学基金(No. 62176029)、重庆市科技创新与应用发展专项(CSTB2023TIAD-KPX0064、CSTB2022TIAD-KPX0206)的部分资助。钟 将(通信作者),博士,教授,中国计算机学会(CCF)会员,主要研究领域为大数据分析、自然语言处理、云网融合、网络安全等。E-mail:zhongjiang@cqu.edu.cn。刘雨轩,硕士,主要研究方向为情感分析和信息提取。戴启祝,博士,主要研究方向为关系提取。王佳祺,硕士,主要研究方向为机器阅读理解。赖心怡,硕士,主要研究方向为图像标注。胡雯月,硕士,主要研究方向为关系提取。

framework for the extraction task and introduce a Dynamic Prefix Prompt method. This method uses continuously adjustable prefix prompts to dynamically optimize the model's overall task representation. By incorporating category information into the prompt prefixes, it enables the model to dynamically generate context-relevant quadruples. Specifically, it initializes specific prefix vectors for predefined categories, combines them with context-encoded sentence representations, and applies attention weights to generate task-relevant prompt prefixes, enhancing the model's ability to focus on and generate context-relevant sentiment quadruples. To handle imbalanced implicit sentiment data, We designs a data augmentation strategy based on a large language model, which utilizes the rich knowledge of the large language model to synthesize new implicit sentiment data, in order to enhance the model's ability in processing implicit sentiment. Firstly the strategy aligns the big language model with the data generation task through instruction fine-tuning, and secondly generates diverse review texts inversely using quadruple labels as inputs. Specifically, we firstly constructs instruction fine-tuned data pairs to map real quadruples as the generation target, and then employs highly efficient fine-tuning of the big language model in order to control the diversity and quality of the generated texts, and finally retains high-quality enhanced samples that are consistent with the original data distributions by means of clustering and semantic filtering. high-quality augmented samples consistent with the distribution of the original data, realizing the mitigation of the data long-tail distribution problem. To verify the effectiveness of the method, extensive experiments are conducted on two real application datasets in this paper. The experiments show that our method improves $F1$ scores by 3.60 and 2.20 on the Restaurants-ACOS and Laptop-ACOS datasets, respectively, while the $F1$ scores in the implied sentiment data are improved by an average of 4.23 and 4.67, which is up to the current state-of-the-art level, verifying the validity and superiority of the present method.

Keywords aspect-based sentiment analysis; task of sentiment quadruple extraction; dynamic prefix prompt; implicit sentiment; large language models; data augmentation

1 引言

1.1 问题分析

方面—类别—观点—情绪四元组提取(ACOS)是方面级情感分析(ABSA)中的一项最具有挑战性的任务。其中包含四个子任务,分别是:(1)方面预测(Asspect Prediction, AP);(2)情绪预测(Sentiment Prediction, SP);(3)类别检测(Category Detection, CD);(4)观点提取(Opinion Extraction, OE)。这些任务都有助于理解评论中的方面级观点,并提供一个完整的情感分析视野。例如,评论“这家日料店的寿司令人垂涎欲滴”。该评论可以提取出一个由方面、类别、观点、情绪组成的情感四元组,即方面(Asepect)为“寿司”,类别(Category)为“食物质量”,观点(Opinion)为“垂涎欲滴”,情绪(Sentiment)为“积极”。传统的研究将此问题视为任务交互问题,如使用 BERT 编码、专家网络^[1]、任

务交互学习策略^[2]。但考虑到子任务交互难度和情感元素空间搜索,这些方法难以满足应用需要。近期,针对提取多情感元素的 ABSA 任务,一系列基于生成的方法^[3-7]引起了广泛关注。这些方法将传统的标注任务重构为生成任务,能利用文本中丰富的语义标签,同时避免过度关注多任务交互而带来的挑战。目前,相关基于生成的方法,其不足可以归纳为两点,即存在提示次优和隐含情感数据的不平衡。本文对这些问题进行深入的分析:

(1)提示次优

当前,相关方法^[3,5]通过给模型输入针对任务设计离散提示词,进行情感四元组的提取,这样的提示词需要人为设计,而且无法根据任务进行调优。Gou 等人^[6]采用了一种自适应的离散化提示模板方法,为情感四元组提取提供了多种候选的提示模板。Jun 等人^[7]采用了实例级熵来动态生成候选的离散提示模板,用于提取情感四元组,这些方案虽然进一步优化了提示模板的选择,但仍然属于离散提示的

范畴。

为了进一步提升离散提示的效果,我们采用了连续可调的前缀提示来替代传统的离散提示词。这种方法不需要针对特定任务进行人为设计,已经在实际应用中得到了验证,证明了其可行性和有效性^[8-10]。然而,尽管这种方式初步优化了提示的效果,但现有的提示方法仍存在次优问题,即无法动态地针对特定任务进行优化。这限制了模型在生成情感四元组时捕捉文本中语义信息的能力。

在前缀调优(Prefix-Tuning)中,前缀提示作为一小段连续的任务特定向量,虽然能表示整体任务的关键特征,但如果提示词只包含整体任务的特征,就会忽略任务间的内在联系,导致提示次优。具体来说,当直接使用前缀提示模型生成情感四元组时,由于当前的前缀提示只引导模型完成情感四元组任务,而没有考虑评论中的主题偏好,这使得前缀提示无法充分关注评论中的主题信息。为了解决这一问题,我们引入了类别信息作为主题信息的一部分,参与到提示前缀中。类别信息,如“餐厅—食物”,不仅包含了情感表达的上下文,还与可能的观点词有直接关联。通过将类别信息纳入提示前缀,模型能够根据当前上下文动态生成更符合主题(与类别相关)的四元组。例如,在评论“餐厅的寿司令人垂涎欲滴”中,如果前缀提示未能关注到上下文评论中的主题信息,模型将仅按照前缀的任务指示生成情感四元组。相反,如果前缀提示能基于当前评论上下文动态关注到主题(类别)信息,如“餐厅—美食”,这将帮助模型按照“餐厅—美食”的提示生成与“餐厅—美食”相关的词,如“寿司”、“垂涎欲滴”。因此,这种由各个类别动态集成组成的前缀提示,能帮助模型关注评论中的主题信息,即任务间的内在联系,实现提示的进一步优化,生成更准确的情感四元组。

值得说明的是,本文选择类别作为主题信息,是因为在评论文本中,类别往往决定了细粒度情感的相关内容^[11]。例如,类别“餐厅—食物”不仅能对应到“寿司”“牛排”等属于食物类别的方面词,也能联系到“美味”“难吃”等观点词。与其他情感元素(如方面、观点、情绪)相比,类别的数量更合理,指代的语义也更明确。因此,本文将类别视为评论的主题表示,作为主题信息动态融合到前缀提示中,有助于模型生成更精确的情感四元组。这种方法不仅提高了模型对评论中主题信息的关注度,还增强了生成的情感四元组的准确性和相关性。

(2) 隐含情感数据不平衡

在现实中文本存在大量未明确指代关系的实体^[12],而在评论文本里,也可以观察到许多评论并未直接明确给出其情感所指的具体对象^[13],这类评论也称为隐含情感数据^[3,14]。由于隐含情感的标注成本高昂,且需要专家知识对其标注,至于现有数据集中包含相当多的隐含情感数据,且存在不平衡的问题。

如表 1 所示,其中 EA&EO (Explicit Aspect and Explicit Opinion) 表示评论有明显所指代的方面词和观点词,它是属于显式情感。而 EA&IO (Explicit Aspect and Implicit Opinion)、IA&EO (Implicit Aspect and Explicit Opinion)、IA&IO (Implicit Aspect and Implicit Opinion) 分别表示评论中方面词或者观点词无明显提及或者均未提及,它们属于隐含情感。由表 1 可以看到,餐厅数据集 (Restaurant-ACOS) 中三类隐含情感分别只占 14.49%、9.57% 和 9.54%,这说明数据集存在长尾分布现象。目前,现有模型在处理评论中隐含情感面临精度不高的问题。这主要源于隐含情感数据的不平衡性而导致的数据长尾分布的限制。这一困难限制了当前模型的发展。

表 1 数据集统计

数据集		Restaurant-ACOS	Laptop-ACOS
句子总计		2286	4076
情感四元组总计		3658	5587
显式情感	EA&EO	2429 (66.40%)	3269 (56.77%)
	IA&EO	530 (14.49%)	910 (15.80%)
隐含情感	EA&IO	350 (9.57%)	1237 (21.48%)
	IA&IO	349 (9.54%)	342 (5.94%)

1.2 研究内容

为了解决上述问题,本文提出一种新的基于动态前缀提示及数据增强的情感四元组提取方法。首先,设计了一种基于生成模型的情感四元组提取任务的生成框架。并在此基础上,进一步引入了动态前缀提示机制,通过连续可调的前缀提示来动态优化模型对任务的整体表示。

具体来说,本文为每个类别训练了特定的前缀提示,并通过注意力机制将这些前缀提示与上下文信息动态集成,以实现模型生成过程的动态优化。该方法不仅能够提高模型充分捕捉到文本中的语义信息,还能够根据当前上下文的类别信息,引导模型关注评论句子中更相关的类别信息,从而生成更为准确的、贴切上下文主题的相关情感四元组。

此外,为了应对隐含情感数据不平衡的问题,本文采用了基于大模型的数据增强方法。通过指令微调方式,将大语言模型与数据扩充任务对齐,利用大语言模型的丰富知识来合成新的隐含情感数据,以增强模型在处理隐含情感方面的能力。最后通过在两个真实数据集上的实验,以验证本研究的优越性和数据增强的有效性。

本文主要贡献总结如下:

(1)为了缓解现有方法存在的提示次优的问题,我们提出了一种基于上下文动态前缀提示的改进方法。该方法提供了可优化的前缀提示,并且通过注意力机制将上下文类别信息与前缀提示动态地整合在一起,以实现模型的动态提示优化。

(2)为了解决隐含情感数据不平衡而导致模型性能下降的问题,我们设计了一种基于大语言模型的数据增强策略,该方法利用大语言模型的丰富知识来合成新的隐含情感数据,以增强模型在处理隐含情感方面的能力。

(3)在两个真实数据集上的实验表明,我们的方法在 Restaurants-ACOS 和 Laptop-ACOS 数据集上 F1 分数分别提升 3.60 和 2.20,达到目前最先进的水平,验证了本方法的有效性和优越性。

2 任务定义

对于情感四元组提取任务。在数据集 D 中给

定一个输入的句子 $x = \{w_1, w_2, \dots, w_n\}$, 每个情感四元组 q 由方面 a 、观点 o 、类别 c 和情绪 s 组成,即 $q = \{a, c, o, s\}$ 。对于 $a \subset x$, 表示句子 x 中的一个或多个词代表被评价的方面;对于 $o \subset x$, 表示句子 x 中的一个或多个词代表关于方面 a 的评价观点;对于 $c \in \mathcal{C}$, 表示 c 是一个预定义的类别标签,用于分类方面 a , \mathcal{C} 表示数据集 D 中所有类别的集合;对于 $s \in S$, 表示表达的情感极性, S 含有积极 (Positive)、消极 (Negative) 和中立 (Neutral)。情感四元组提取任务需要识别并提取句子 x 中的所有的情感四元组集合 Q , 即 $\{q_1, q_2, \dots, q_n\}$, 其中每个 $q_m = (a_m, c_m, o_m, s_m)$ 是句子 x 中的第 m 个情感四元组。

如图 1 所示,评论示例中包含四个四元组,这些四元组对应的类别分别是“Ambience General”、“Service General”、“Restaurant General”和“Restaurant miscellaneous”,与之对应的方面词是“ambience”、“restaurant”、“null”、“null”。在这些类别中,其描述的观点分别为“attentive”、“welcoming”、“null”、“null”。同时,所属的情绪极性也为积极、积极、消极和积极。当方面和观点被标记为“null”时,意味着没有特定的方面词或观点被明确提及,但仍然可以从文本中感知到明显的情绪倾向^[15]。例如,“Restaurant miscellaneous”类别所在的四元组虽无明显提及的方面和观点,但其情绪被标注为负面,这表示其中蕴含了一种隐含的情感表达。



图 1 方面预测、情绪预测、类别检测和观点提取的例子

3 模型描述

本文首先介绍基础的离散提示生成模型 Gen-ACOS,它通过生成机制来提取情感四元组。针对其局限性,我们在基础模型之上进一步引入了 Gen-CDPT 模型,其采用了上下文的动态前缀提示,以实

现提示的动态优化。

3.1 基于离散提示的生成模型

通常基于离散提示的情感四元组生成方法依赖于静态提示模板和对应的生成框架。对此,我们首先提出基础模型 Gen-ACOS (Generative model for Aspect-Category-Opinion-Sentiment extraction), 以生成式的方法获取情感四元组。整体的基础模型

Gen-ACOS 结构如图 2 所示。

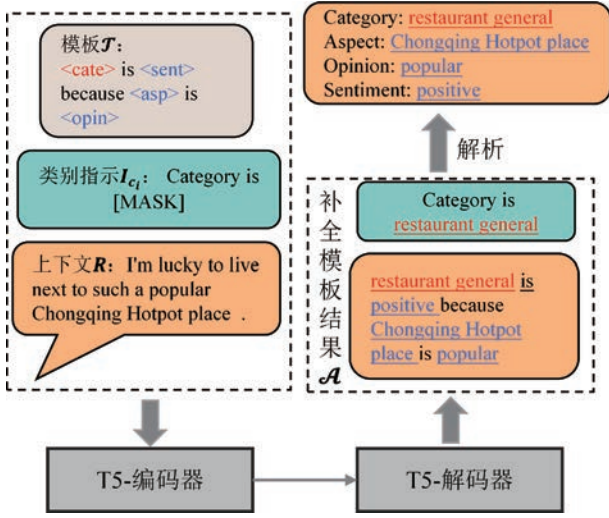


图 2 基础模型 Gen-ACOS 结构示意图

(1) 类别指示

类别指示由一句简短的自然语言序列构成,通常人工设计的形式是“类别是[MASK]”。例如,对于类别构造指令,类别“餐厅—饮品”的类别指令是“Category is restaurant drink.”。本文用表示离散提示方法的类别指示。

(2) 模板

模板通常包含几个占位符,用于表示期望生成的情感元素,通常“<cate>”表示类别,“<asp>”表示方面,“<opin>”和“<sent>”分别表示观点和情绪。由于本文的重点不是提示模板的设计,所以本文采用相关研究人员^[3,6]设计的模板,并用 \mathcal{T} 表示模板。

(3) 生成框架

Gen-ACOS 采用编码器—解码器架构的生成模型 T5 进行生成建模。在生成过程中, T5 利用自回归生成方式来建模输入序列 X 在先前生成的令牌(token)即 $y_{<i}$ 的条件下生成下一个令牌 y_i 的概率。该模型能够计算在给定输入序列 X 的情况下生成输出序列 Y 的整体概率 $p(Y|X)$, 因此给定输入序列 X , 生成输出序列的 Y 的整个概率 $p(Y|X)$ 计算为:

$$p(Y|X) = \prod_{i=1}^{|Y|} p(y_i | y_{<i}, X)$$

$$P_{c_i} = [I_{c_i}; \mathcal{T}]$$

$$X = [P_{c_i}; R]$$

$$Y = A \quad (1)$$

其中 $[\cdot; \cdot]$ 代表序列连接操作, R 代表当前输入的上下文, P_{c_i} 代表在类别 c_i 下的提示模板,它由上述

的 I_{c_i} 和 \mathcal{T} 组成。 \mathcal{A} 表示最终补全模板的结果,其中包含了对应的目标情感四元组。

(4) 训练策略

训练策略。用 ϕ 来表示 Gen-ACOS 模型中所有的可训练参数。因此,训练目标是最小化目标输出序列 $Y = \{y_1, y_2, \dots, y_m\}$ 和模型预测序列的差异,其中 X_j 表示 D 中的第 j 个样本与提示模板组成的输入。损失函数如下:

$$\mathcal{L}(D) = - \sum_{j=1}^{|D|} \sum_{i=1}^m \log p_{\phi}(y_i | y_1, \dots, y_{i-1}, X_j) \quad (2)$$

(5) 推理和解析策略

在推理阶段,基础生成模型通过束搜索(Beam search)的方式生成序列,并选择生成概率最大的序列作为结果。其生成的最大序列长度根据数据集统计量设置,略大于最长的句子。在结果解析时,按照模板顺序,依次从补全的模板 \mathcal{A} 中顺序解析出“<cate>”、“<sent>”、“<asp>”和“<opin>”这些情感元素占位标记对应的词,如果不能匹配当前尝试解析的情感元素,则直接跳过对该情感元素的解析。

3.2 上下文动态前缀提示

为了解决基于离散提示的基础生成模型中提示次优的不足,本文在 Gen-ACOS 的基础生成模型上,提出 Gen-CDPT(Generative model for Contextual Dynamic Prefix-Tuning)模型。其整体框架结构如图 3 所示,其中 Gen-CDPT 主要由类别前缀提示、上下文动态前缀提示、重参数化(Re-parametrization)层构成。

(1) 类别前缀

在基础生成模型中需要手动设计的一串自然语言的离散提示。受到前缀调优(Prefix-Tuning)的启发,本文将原有的类别指示 I_{c_i} 设计为一串可以调优的前缀向量。具体来说,对于每一种类别 c_i ,先初始化一个长度为 L ,维度为 D 的随机向量,用来表示对应类别(主题)的前缀提示。然后训练一个嵌入矩阵 $\mathbf{P} \in \mathcal{R}^{|C| \times L \times D}$ 建模前缀向量的映射关系,从而得到所有的类别前缀提示 $sp_c \in \mathcal{R}^{|C| \times L \times D}$,其中类别 c_i 对应的前缀提示是:

$$sp_c^i = \mathbf{P}[c_i, i, :] \quad (3)$$

其中 i 表示所有的类别前缀 sp_c 的第 i 个位置索引,即第 i 个类别前缀表示 $sp_c^i \in \mathcal{R}^{L \times D}$ 。通过只训练该类别的数据,以得到对应的类别前缀 sp_c^i 。这样本文实现原有的特定类别指示到特定类别前缀的

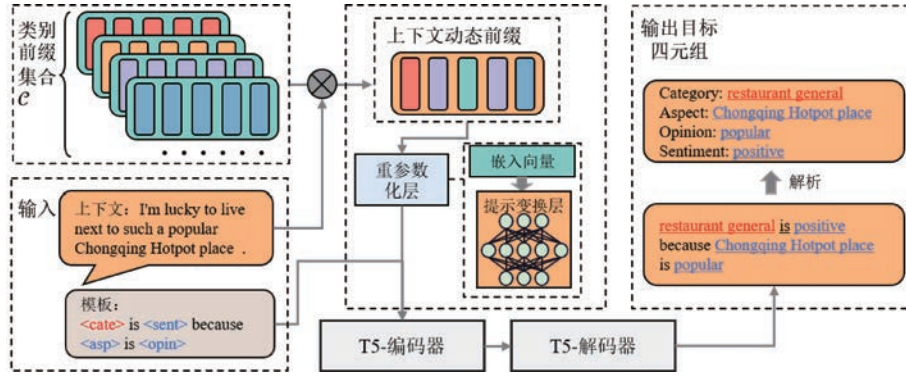


图3 Gen-CDPT 模型结构示意图

转化,即 $I_{c_i} \rightarrow sp_c^i$ 。注意,由于这里的所有的类别前缀 sp_c 的映射矩阵嵌入矩阵 P ,与 Transformer 中历史激活序列进行拼接,其主要优势在于它们为编码器和解码器提供了可训练的上下文以此实现前缀提示调优,所以这在计算上也是可行的。接下来,将继续阐述如何对 sp_c 进行动态结合,使得模型能关注不同的主题信息。

(2) 上下文动态前缀

本文提出基于上下文、包含各类别信息的动态前缀,以动态集成类别(主题)信息到前缀提示中,其用 dp_c 表示。如图4所示,上下文动态前缀 dp_c 和单一类别前缀 sp_c^i 序列长度相同,长度为 L 。为了让模型获取到在上下文信息下不同类别前缀的细粒度表示,本文采用注意力机制将上下文信息和类别前缀 sp_{c_i} 动态组合,具体步骤如下所示:

$$\begin{aligned} r_{cls} &= \text{BERT}(R)_{cls} \\ e_i &= \mathbf{v}^T \tanh(W_{sp} sp_c^i + W_r r_{cls}) \\ a_i &= \text{softmax}(e_i) \\ dp_c &= \sum_{i=1}^{|C|} a_i sp_c^i \end{aligned} \quad (4)$$

其中 r_{cls} 由 BERT 模型进行编码并取第一个索引位置的向量,是当前上下文 R 的平均句子表示。公式(4)是整体注意力机制的公式,其中它将上下文表示 r_{cls} 和类别前缀向量输入到多层感知机里,从而得到一个注意力分数 e_i 。把注意力分数 e_i 进行 Softmax 归一化得到权重系数 a_i ,将 a_i 和各个类别前缀进行加权平均,就得到了在上下文信息下的动态类别表示 dp_c 。注意这里基于上下文特定前缀 dp_c 是动态加权的,因为它在当前上下文信息下考虑了 C 中所有类别的前缀信息。这样本文实现了将各个类别信息动态集成到前缀提示中,从而让模型更加关注对于四元组提取更有用的类别前缀提示。

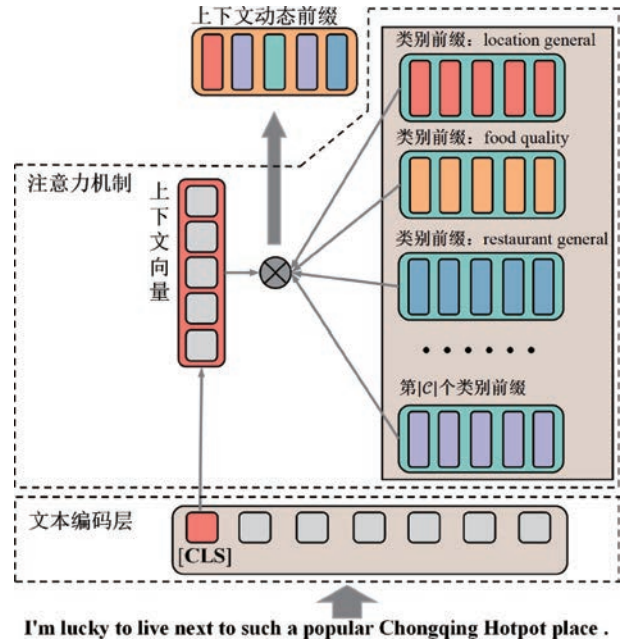


图4 采用注意力机制获得上下文动态前缀提示

(3) 重参数化(Re-parametrization)层

本文遵循前缀调优的训练指导,通过建模一个词嵌入向量 P' 和由多层感知机组成的提示变换层对前缀进行重新参数化,嵌入向量 P 计算如下:

$$P[c_i, i, :] = \text{MLP}(P'[c_i, i, :]) \quad (5)$$

其中 $P' \in \mathcal{R}^{|C| \times L \times D'}$, 通常 $D' < D$, 这是因为直接更新原始的嵌入向量 $P \in \mathcal{R}^{|C| \times L \times D}$ 的参数会导致模型优化不稳定且性能下降。因此采用重参数化的方式能降低整体模型的容量,以限制其模型自由度,实现其训练稳定性^[8-9]。

3.3 整体训练流程

本方法中训练参数除了基础模型 Gen-ACOS 的可训练的参数 ϕ 外,还包括前缀提示嵌入向量和上下文信息编码器 BERT 组成动态上下文前缀的参数 θ_{prf} 。

在具体实验中发现采用联合训练的方式优化上

述参数,效果欠佳,这是因为 ϕ 和 θ_{prf} 之间需要不同尺度的训练超参数以达到最佳优化效果。所以,本文分三个步骤进行训练:

(1)首先,使用基础模型 Gen-ACOS 训练参数 ϕ ,以此学习基础生成任务的能力,得到训练好的参数 ϕ' 。

(2)其次,冻结住已经训练好的 Gen-ACOS 模型的参数 ϕ' 。为了获得主题信息,对每一种类别 c_i ,都训练出对应的类别前缀提示 sp_c^i ,进而组合成完整的类别前缀表示 sp_c 。

(3)最后,加载已经训练好的所有的类别前缀表示 sp_c ,同时冻结住参数 ϕ' ,训练动态前缀和上下文编码器 BERT 的参数 θ_{prf} 。

4 数据增强策略

为了解决隐含情感数据不平衡而导致数据长尾分布以至于模型在处理隐含情感时性能欠佳。本文提出的基于大语言模型的数据增强策略。该策略本质上是一种大模型知识蒸馏方式^[16-17]。近期研究表明^[18-21],通过将大型语言(教师)模型(LLMs)的通用知识蒸馏至轻量级学生模型中,能够有效提升下游任务的模型性能。这种知识迁移方法已被证明是提升模型效率的实用技术路径。

基于此,本研究提出一种基于知识蒸馏的数据增强策略,重点解决隐式情感样本稀疏性问题,所提出的数据增强策略整体流程如图5所示,其详细步骤如下。

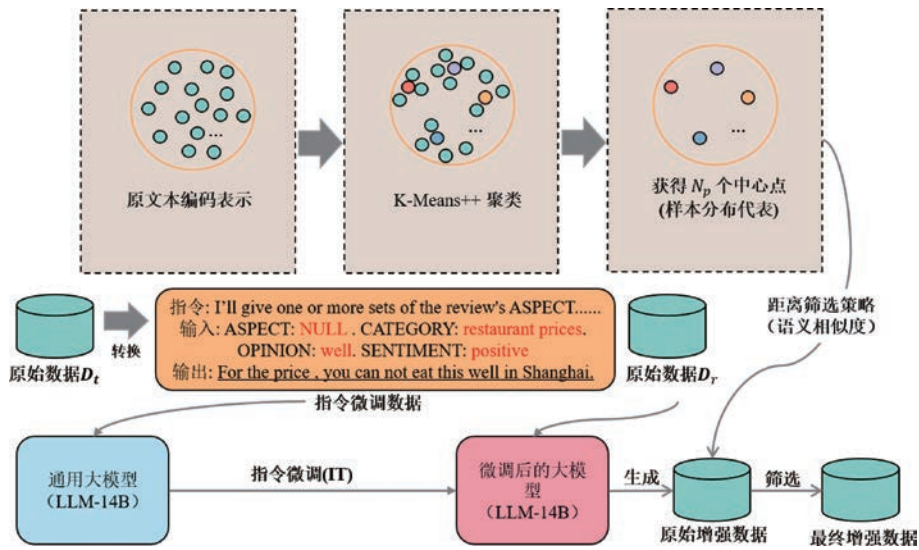


图5 基于大语言模型的数据增强整体流程

4.1 指令数据构造

为了让通用大语言模型适应数据的分布,同时学习特定任务相关的模式和规则,本文采用指令微调(Instruction Tuning, IT)方式对大语言模型进行数据扩展意图任务的对齐,而指令数据是指令微调必不可少的指令微调数据,其中每条数据实例的形式都如: (“指令”: [prompt], “输入”: [input], “输出”: [output])。其中[prompt]一般代表由自然语言描述的指令,[input]代表输入的信息,而[output]则代表期望模型输出的回复。

本文指令微调数据构造流程如图6所示。将整体的数据集 D 平均随机划分为 D_t 和 D_r 。对于 D_t 中的每一条数据对 $(x_i, \{q_i\})$,即文本评论和其对应的情感四元组,为其构造指令微调数据对 (x_i^I, y_i^I) 。具体地讲,本文定义一个映射函数 f ,该函数接受一个真实的情感四元组集合 $\{q_i\}$ 以及指

令 I ,输出指令微调数据对 (x_i^I, y_i^I) ,其中 x_i^I 由原始的情感四元组集合 $\{q_i\}$ 和指令 I 得到,而 y_i^I 就是原始的评论句子 x_i 。这种构造方法的优点是,不需要引入额外专家知识,采用少部分的真实数据即可构造出指令微调数据,使大模型后续对齐数据扩展的任务。

4.2 指令微调

为了让大语言模型对齐数据扩充任务,本文选择解码器架构的通义千问 Qwen 大语言模型进行有监督指令微调。Qwen 是一款基于 Transformer 架构的大语言模型,在中英文领域上的理解、生成中具有行业领先优势^[22]。同时,Qwen 模型基于高质量预训练语料和人类反馈进行训练,展现出卓越的理解能力,所以通过对其指令微调,Qwen 能够更好地实现任务对齐^[22]。

该大语言模型利用自回归生成方式来建模输入

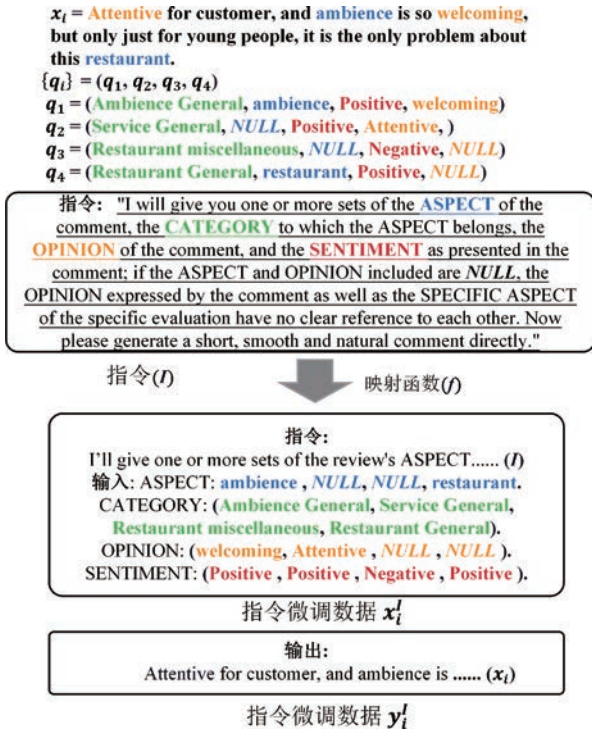


图6 指令微调数据的构造流程

序列在先前生成的令牌(token)即 $y_{<k}$ 的条件下生成下一个令牌 y_k 的概率。该模型能够计算在给定输入序列 X 的情况下生成输出序列 Y 的整体概率 $p(Y|X)$, 对于第个微调数据对 (x_l^I, y_l^I) , 其生成概率表示如下:

$$p(YX) = \prod_{k=1}^{|Y|} p(y_k | y_{<k}, X) \quad (6)$$

这里输入序列 X 和输出序列 Y 分别为 y_l^I 和 x_l^I 。

在训练阶段,其模型参数使用预训练的权重进行初始化,采用教师强制策略(Teacher forcing)进行微调训练,即同时输入指令微调数据对 (x_l^I, y_l^I) 。其有监督指令微调的损失函数如下:

$$\mathcal{L}(D_l) = - \sum_{j=1}^{|D_l|} \sum_{i=1}^m \log p_{\theta}(y_i | y_1, \dots, y_{i-1}, X_j) \quad (7)$$

其中 θ 表示所有可训练参数, $|D_l|$ 表示用于指令微调的训练集样本数(本文采用部分原始样本 D_l 来进行扩充), X_j 表示第 j 个指令微调数据的输入序列, m 表目标序列长度。然后使用优化器将可调的模型参数 θ 更新为 θ_l 。

4.3 数据生成策略

在对大语言模型进行指令微调后,本文选择高温、高 top-p 的生成参数来指导模型生成结果。

这是因为温度能调整单词的概率分布,在较低的温度下,自回归模型每次迭代生成的数据更具确定性,而在较高的温度下模型每次迭代生成的单词不确定性更高^[23]。同时较高的 top-p 也可以保留一些有趣或有创意的单词^[24]。为了生成足够多样性原始增强数据,本方法多次重复迭代生成足够多的原始数据。基于上述设置,本文根据微调更新后的参数 θ_l 进行原始增强数据的生成:

$$D_a = \sum_{t=1}^T G_{\theta_l}(f(D_r)) \quad (8)$$

其中 G_{θ_l} 代表指令微调后的大语言模型, $f(D_r)$ 代表经过映射函数 f 代表经过映射函数 D_r , T 代表对于每个样本重复生成的次数, D_a 代表迭代 T 次后生成的原始增强数据。注意到本文采用剩余数据集 D_r 用于生成推理,而不是先前用于构造指令微调的数据 D_l ,这是为了防止大模型生成过于相似于用于微调的数据,目的是增加数据集的多样性,以提高模型的泛化能力和性能。

4.4 数据后筛选策略

考虑到原始的增强数据 D_a 是在高温、高 top-p 参数指导下生成的,其中数据分布的不确定性更高,对此本文采用距离筛选策略过滤掉部分数据,使其整体分布趋近于原始数据集 D 。具体的步骤如下:

(1)首先采用 K-Means++^[25] 聚类算法以获取原始数据集 D 的中心表示,即中心点集合 P_n 。

(2)计算生成数据集 D_a 中每个样本与每个中心点 $\{p | p \in P_n\}$ 的语义相似度之和 ϵ 。其中具体的计算方式如下:

$$\epsilon = \sum_{j \in D_a} \sum_{c \in P_n} \cos(j, c) = \sum_{j \in D_a} \sum_{c \in P_n} \sigma$$

$$\sigma = \cos(j \cdot c) = \frac{j \cdot c}{\|j\| \|c\|} \quad (9)$$

(3)取语义相似度和 ϵ 最高的样本,加入筛选后的增强数据集 D_f 。

(4)不断迭代步骤(2)、(3)直到 D_f 的数量达到需要保留增强数据集个数 N_d 。算法1描述了数据后筛选策略的整体过程。

算法1. 数据后筛选策略的整体过程。

输入:

D_a : 增强数据集;

D : 完整数据集;

N_p : 聚类中心簇数量;

N_d : 最终保留的增强数据个数

输出:最终增强数据集 D_f


```

1. 初始化  $D_f = \Phi$ 
2. 采用 K-Means++ 聚类算法, 从  $D$  中分别获取  $N_p$  个中心点集合  $P_n$ 
3. for  $i$  in range( $N_d$ ) do
4.   初始化  $min\_sim\_sum = +\infty$ 
5.   初始化  $selected\_element = 0$ 
6.   for  $j$  in  $D_a$  do
7.     初始化  $sim\_sum = 0$ 
8.     for  $c$  in  $P_n$  do
9.       计算  $c$  与  $j$  的文本余弦相似度, 得到相似度  $\sigma$ 
10.       $sim\_sum += \sigma$ 
11.    end for
12.    if  $sim\_sum < min\_sim\_sum$  then
13.       $min\_sim\_sum = sim\_sum$ 
14.       $selected\_element = j$  # 保留与中心点最小的距离和的元素
15.    end if
16.  end for
17.  将  $selected\_element$  加入到集合  $D_f$ 
18. end for
19. return  $D_f$ 

```

5 实验设计与分析

我们在两个真实的数据集进行了广泛的实验, 以进一步验证所提出的方法的有效性。在这里, 我们主要对以下 5 个方面进行实验分析:

方面 1: Gen-CDPT 方法与基线工作相比的性能表现。

方面 2: 促进情感四元组有效提取的 Gen-CDPT 关键组成部分。

方面 3: 上下文动态前缀提示对优化提示的作用。

方面 4: 前缀提示对 Gen-CDPT 性能的影响。

方面 5: 采用数据增强策略生成的数据可靠性及其对模型性能的贡献。

5.1 实验设置

(1) 数据集

所选用的数据集分别为 Restaurant-ACOS 和 Laptop-ACOS, 它们来源于 SemEval 测评任务所采用的 Restaurant 和 Laptop 数据集, 由 Cai 等人^[14]在原有基础上进行了重新标注, 以适应更为复杂的情感四元组提取任务。这两个数据集涵盖了餐厅和电子产品领域, 其整体数据统计详见表 1。

(2) 实现细节

在实验中, 对于 Laptop-ACOS 和 Restaurant-

ACOS 中的训练集和测试集, 本文保持与官方相同的划分。该方法在 2 张 NVIDIA RTX A6000 GPU 上使用 PyTorch、Transformers、Peft 实现所有模型的构建和训练的程序。采用 T5-base^[26]作为主干生成网络, 采用 BERT-base^[27]作为上下文编码器。在实验中使用学习率为 $1e-4$ 的 AdamW (Adam weight decay) 优化器^[28]优化 T5-base 模型, 使用带权重衰减的学习率为 $3e-5$ 的 AdamW 优化器优化 BERT-base 编码器和各个类别前缀以及动态前缀。对于前缀向量, 其维度大小为 768, 长度为 120。

关于数据增强策略, 本文决定对含有隐含情感的句子进行扩充, 将其数量增加至原有的一倍, 并使用学习率为 $1e-5$ 的 AdamW 优化器微调 Qwen 基础大语言模型, 采用 LoRA 高效微调^[29]的方式, 仅对 Qwen 注意力层和前馈神经网络模块进行优化, LoRA 秩为 128。更多的实验参数如表 2 所示。

表 2 实验参数

类别	Gen-ACOS	Gen-CDPT	数据增强
最大长度	256	256	256
前缀长度	—	120	—
前缀维度	—	512	—
优化器	AdamW	AdamW	AdamW
权重衰减	$1e-5$	$1e-5$	$1e-5$
Batch size	64	16	128
学习率	$1e-4$	$3e-5$	$1e-5$
Epochs	20	30	10
LoRA 秩	—	—	128
梯度累加	1	1	8

(3) 评价指标

遵循现有的研究^[30-31], 我们采用精确度 (Precision)、召回率 (Recall) 和 F1 分数作为主要的评估指标。值得注意的是, 只有当一个样本中情感四元组中的所有情感元素及其组合与真实标签 (ground truth) 的四元组完全一致时, 才被视为正确的匹配。

5.2 基线对比方法

本文将提出的 Gen-CDPT 和多种基线工作进行了对比。相关基线方法如下:

(1) DP-ACOS^[14]: 一种基于双传播的方法, 其也是具有代表性的基于规则方法的情感四元组提取方法。

(2) JET-ACOS^[14]: 一种位置感知端到端框架, 结合了方面、对应的观点和情绪极性的标记方案。

(3) TAS-BERT-ACOS^[14]: 由 Cai 等人的提出的一种两步流水线方法, 利用 TAS-BERT 模型提取情感四元组。

(4)EC-ACOS^[14]:采用了一个具有代表性的方面—观点联合提取系统,以适应情感四元组提取。

(5)BARTCRN^[32]:提出了一种基于 BART 模型的方法,该方法通过生成并解码过程来提取情感四元组。

(6)GAS^[4]:一种采用注释与提取范式生成情感元素的模型,针对提取情感四元组任务进行了调整。

(7)PARAPHRASE^[5]:是一种有效的方面—类别—情绪—观点生成模型,它是一种基于文本序列到文本描述的生成方法。

(8)USI^[33]:通过多任务指令微调的方式统一生成情感四元组。

(9)SI-T2S^[3]:一种通过离散提示生成结构信息的方式提取情感四元组。

(10)Seq2Tree^[34]:采用语法树作为序列生成的目标结构,进行情感四元组提取任务。

(11)iACOS^[35]:通过多头注意力和多标签序列标注的方式实现方面—类别—情绪—观点的联合提取。

(12)MVP^[6]:一种基于多视角离散提示的情感四元组提取方法,通过不同顺序生成和聚合情感元素。

(13)GenDA^[36]:提出了一种四元组到文本模型的生成式数据增强方法用于解决情感四元组提取,针对这一模型,本文在 ACOS 数据集上进行了重新适配和复现。

(14)ADA^[37]:提出了一种自适应数据增强框架进行情感四元组提取,针对这一模型,本文在 ACOS 数据集上进行了重新适配和复现。

(15)ATOSS^[38]:介绍了一种基于大型语言模型的情感四元组提取方法,该方法通过将句子拆分成更简单的部分,并重新组合成提示信息。

(16)LCD/&CD^[39]:通过引入潜在分布来捕捉文本与情感元素之间的关系强度,并利用约束解码策略来生成情感四元组。

(17)Chat-GPT^[40]:最近流行的一种大型语言模型,本实验设计了少样本的提示指令来激发 ChatGPT 的理解能力以此实现方面—类别—情绪—观点的四元组提取。

(18)Gen-ACOS:是本文提出的采用离散提示的基础生成模型,以实现情感四元组提取。

5.3 主要实验结果

为了比较 Gen-CDPT 方法与基线工作的性能

表现(即方面 1),本文将 Gen-CDPT 模型与基线结果进行比较,其中标有星号(*)的条目表示该实验在 ACOS 数据集上进行了重新复现。主要实验结果如表 3 所示,其中 P 代表精确度, R 代表召回率, $F1$ 代表 $F1$ 分数。总体而言,提出的 Gen-CDPT 模型在基线上取得了良好的结果。这些实验结果表明 Gen-CDPT 能够实现优秀的性能。

表 3 主要实验结果

模型/数据集 评价指标	Restaurant-ACOS			Laptop-ACOS		
	P	R	$F1$	P	R	$F1$
DP-ACOS	34.67	15.08	21.04	13.04	5.71	8.01
JET-ACOS	59.81	28.94	39.01	44.52	16.25	23.81
TAS-BERT-ACOS	26.29	46.29	33.53	47.15	19.22	27.31
EC-ACOS	38.54	52.96	44.61	45.56	29.48	35.8
BARTCRN	56.62	55.35	55.98	41.63	40.46	41.05
GAS	56.01	56.01	56.01	42.04	40.91	41.47
PARAPHRASE	58.76	59.30	59.08	45.06	41.88	43.47
USI	60.07	61.14	60.61	44.57	43.91	44.24
SI-T2S	62.36	61.41	61.89	46.71	43.58	45.29
iACOS	57.24	53.21	55.15	49.59	34.65	40.80
MVP	62.43	60.77	61.59	45.59	42.32	43.92
Seq2Tree	63.96	61.74	62.83	46.11	44.79	45.44
GenDA*	64.26	62.42	63.33	44.71	43.49	44.09
ADA*	62.63	62.46	62.54	45.01	44.09	44.55
ATOSS	63.12	58.30	60.61	45.32	43.17	44.22
LCD&CD	60.30	62.30	61.33	44.70	46.30	45.50
Chat-GPT	41.01	41.01	41.01	34.89	33.65	34.26
Gen-ACOS	61.70	58.96	60.30	45.22	42.86	44.01
Gen-CDPT	63.57	61.51	62.34	46.32	45.96	46.20
Gen-CDPT (数据增强)	66.32	65.56	65.94	51.83	45.39	48.40

具体来讲,在无数据增强策略情况下,Gen-CDPT 在 Restaurant-ACOS 和 Laptop-ACOS 数据集上比基于离散模板生成的 SI-T2S 模型的 $F1$ 分数提升了 0.45 和 0.91;相比于本文提出的基础生成模型, $F1$ 分数提升了 2.04 和 2.19;与当前最先进的 Seq2Tree 模型相比 $F1$ 分数在电子产品数据集上提升了 0.76。在本研究中,我们利用了 OpenAI 推出的 ChatGPT3.5 模型来探索情感四元组的提取效果。

在实际的实验中,我们设计了一系列的提示词,以期引导模型准确识别文本中的关键要素。具体来说,我们设定了以下提示词结构:{"role": "system", "content": "You are a helpful assistant. Now I will give you a review such as, 'The hamburgers at this restaurant are delicious.' where aspect is hamburgers, opinion is delicious, category is restaurant food, and sentiment is positive."},

{"role": "user", "content": [输入]}。这里[输入]代表真实输入的评论句子。

通过实验发现,直接利用语言模型来生成精确的情感四元组存在显著的挑战。这一现象主要归因于大语言模型“幻觉效应”,即模型倾向于产生看似合理但并不准确的输出,特别是在含有隐含情感的情感四元组的生成上。此外,由于缺乏针对性的任务微调,模型在识别和提取那些隐含情感的四元组时显得力不从心,无法有效地与情感四元组提取任务对齐。这些因素共同作用,导致了模型在情感四元组提取任务上的表现不尽如人意。

可以看出,本文提出的数据增强策略使得 Gen-CDPT 模型在原始的基础上 F1 分数提升显著,分别提升和 3.60 和 2.20。这些结果反映出了本方法的几个优势:

(1)该方法明显优于采用离散提示的方法,这是因为 Gen-CDPT 将离散的提示转为了可训练的提示前缀,能在训练中不断优化提示,可以提升模型生成的效果。

(2)为每一种类别训练其对应的前缀,在学习到了特定类别的相关信息后,再通过当前上下文动态的加权组合组成。这使模型能更加关注对于四元组提取更有用的类别前缀提示,以此进一步优化提示,缓解了提示次优问题。

(3)所提出的基于大语言模型的数据增强策略,通过知识蒸馏,从大语言模型中提取知识,以生成高质量的增强数据。这一过程验证了通过微调大语言模型来生成不平衡数据的有效性,其直观地带来了对于模型的效果提升。

5.4 消融实验结果

为研究促进情感四元组有效提取的 Gen-CDPT 关键组成部分(即方面 2),本文继续深入研究 Gen-CDPT 并进行了消融实验分析。

其中,-w/o CDPT 表示去除掉上下文注意力动态前缀而只保留连续提示前缀的模型,-w/o PFT 表示去除连续提示前缀的模型,其等同于基础模型 Gen-ACOS。其消融结果如表 4 所示。

表 4 消融实验结果(F1 分数)

模型	Restaurant-ACOS	Laptop-ACOS
Gen-CDPT	62.34	46.20
-w/o CDPT	61.65	45.84
-w/o PFT	60.30	44.01

可以发现在-w/o PFT 与含有连续前缀提示的模型-w/o CDPT 的对比,可以发现前者在 Restau-

rant-ACOS 和 Laptop-ACOS 数据集上的 F1 分数分别下降了 1.35 和 1.83。这一明显的变化表明,可调节连续前缀提示的有效性,其对于直接优化提示具有积极的影响。这实验结果与本文的预期相符。

在对比去除掉上下文注意力动态前缀的-w/o CDPT 与本文提出的上下文动态前缀提示模型 Gen-CDPT 时,可以发现,在两个不同的数据集上,前者相比于 Gen-CDPT 的性能分别下降了 0.69 和 0.36。这表明,通过动态地将上下文信息融入各类别前缀,形成的连续前缀提示,有助于更有效地提取情感四元组。此外,这一实验结果也在一定程度上证实了上下文动态前缀提示能帮助模型关注到评论中的主题信息,使模型聚焦于对四元组提取有用的类别前缀提示,以提升模型生成效果。

5.5 案例分析

针对方面 3,本文探讨上下文动态前缀提示是否有效提升模型对四元组提取中有效类别前缀提示的关注,从而优化提示,解决提示次优的不足。

本文设计的案例分析实验结果如图 7 所示,其中 Gen-pfx 表示在 Gen-CDPT 的基础上去除了上下文动态前缀提示,只含有连续前缀的模型,“√”和“×”表示正确和不正确的预测。这里的案例选自 Laptop-ACOS 数据集。在句子 1 中,可以看到 Gen-pfx 和 Gen-CDPT 都正确预测了情感四元组,然而在句子 2 中,Gen-CDPT 准确地预测了类别,这证明了动态前缀可以有效地引导模型关注对四元组提取有用的类别前缀提示,并进而生成与当前上下文主题相关的特定情感四元组。对于含有情感四元组较多的长文本句子 3,Gen-pfx 表现不佳,而 Gen-CDPT 继续保持对类别元素和部分观点元素的正确预测,这表明相比于单一的前缀提示,动态的前缀提示能引导模型根据上下文的具体信息调整生成过程,从而更准确地识别和提取情感元素。

5.6 前缀实验分析

为了探究前缀提示对 Gen-CDPT 性能的影响(即方面 4),本文设计了相关前缀实验分析如下。

(1)前缀在类别标签上的对比分析

为系统评估所提出的动态前缀提示机制对类别标签的优化效果,我们在 Restaurant-ACOS 和 Laptop-ACOS 数据集上设计并开展了针对类别细粒度的前缀对比实验。通过对比融合了上下文的主题信息的动态前缀模型(Gen-CDPT)与单一动态前缀模型(Gen-pfx)在不同类别上的性能差异,选定方面词提取准确率(AE-acc)、类别检测准确率(CD-acc)以

#句子1: The laptop is easy to use and just a bit more than a tablet .

真实结果: (laptop, laptop usability, positive, easy) ✓

Gen-pfx提取预测结果: (laptop, laptop usability, positive, easy) ✓

Gen-CDPT的预测结果: (laptop, laptop usability, positive, easy) ✓

#句子2: This does not really matter to me because the mx150 is

enough for my needs.

真实结果: (mx150, graphics design_features, neural, NULL)

Gen-pfx提取预测结果: (mx150, hardware quality, neural, NULL) ×

Gen-CDPT的预测结果: (mx150, graphics design_features, neural, NULL) ✓

#句子3: Consequently , i am no longer covered and my otherwise great chromebook becomes unusable anytime the phantom touches appear , and this leads to poor battery life.

真实结果: (chromebook, laptop general, positive, great),

(phantom touches, keyboard general, negative, NULL),

(battery, battery performance, negative, poor)

Gen-pfx提取预测结果:

(chromebook, laptop general, positive, great), ✓

(chromebook, laptop general, negative, poor), ×

(battery, battery general, negative, unusable) ×

Gen-CDPT的预测结果:

(chromebook, laptop general, positive, great), ✓

(phantom touches, keyboard general, negative, poor), ×

(battery, battery, performance, negative, poor) ✓

图7 案例分析结果

及情感四元组提取 $F1$ 分数作为核心评估指标。

如表5所示,实验结果表明,在案例类别标签中,所提出的动态前缀提示不仅显著提升了模型在整体情感四元组提取任务上的 $F1$ 分数(+1.04%),在类别检测准确度(+1.38%)和方面词提取准确度(+1.46%)上也实现了同步提升。这直接表明,Gen-CDPT模型通过动态调整类别的关注权重,能够更精准地捕捉细粒度情感元素,以提示模型生成更准确的情感元素。

具体而言,在 Laptop-ACOS 数据集上,融合了上下文的动态前缀使类别提取准确性显著提高,在4个基准案例中平均提升1.70个百分点。值得注意的是,该机制对低频类别的提升效果尤为突出。以占比仅1.15%的“支持-性能”类别为例,其 $F1$ 分数较单一动态前缀模型(Gen-pfx)提升4.42个百分点。这同时表明,通过注意力机制动态融合上下文信息的提示前缀,能有效引导模型捕捉低频类别特征,缓解数据长尾分布问题。

值得注意的是,该方法在复杂类别体系中的优势随任务复杂度提升而更加显著。当对比 Restaurant-ACOS(含有13种类别标签)与 Laptop-ACOS(含有121种类别标签)时,所提出的动态前缀的 CD-acc

表5 类别标签对比分析

数据集	类别	指标	前缀模型		
			Gen-CDPT	Gen-pfx	Δ
Restaurant-ACOS	食物-质量	AE-acc	80.23	79.87	0.36
		CD-acc	86.39	85.74	0.65
		$F1$	69.94	69.88	0.06
	类别占比		35.07%		
	餐厅-气氛	AE-acc	84.19	83.29	0.90
		CD-acc	85.73	84.19	1.54
		$F1$	70.21	69.95	0.26
	类别占比		9.19%		
	服务-质量	AE-acc	82.50	81.98	0.52
		CD-acc	88.89	88.14	0.75
		$F1$	62.40	62.26	0.14
	类别占比		5.37%		
Laptop-ACOS	食物-价格	AE-acc	82.29	82.02	0.27
		CD-acc	91.36	90.05	1.31
		$F1$	66.32	65.44	0.88
	类别占比		2.80%		
	笔记本-质量	AE-acc	53.50	52.67	0.83
		CD-acc	70.78	69.54	1.24
		$F1$	48.80	47.85	0.95
	类别占比		20.93%		
	电池-续航	AE-acc	54.13	53.21	0.92
		CD-acc	67.89	66.97	0.92
		$F1$	46.50	46.11	0.39
	类别占比		9.39%		
Laptop-ACOS	显示-性能	AE-acc	54.55	52.27	2.28
		CD-acc	65.54	63.63	1.82
		$F1$	43.27	42.03	1.24
	类别占比		4.73%		
	支持-性能	AE-acc	52.78	47.22	5.56
		CD-acc	63.89	61.11	2.78
	性能	$F1$	45.12	40.70	4.42
	类别占比		1.15%		

平均提升幅度从1.0%提升至1.70%,相对增幅达60.38%。这验证了 Gen-CDPT 的双重优势:其一,能通过上下文感知精准定位相关类别特征,克服传统方法在长尾分布下的性能衰减;其二,Gen-CDPT借助动态注意力机制自适应调整多类别权重,在高度细分的复杂类别标签识别场景中展现出更强的语义适配能力。

(2)前缀参数实验

连续前缀提示主要受到前缀长度和维度的影响。为了深入探究 Gen-CDPT 的内在特性,本研究通过分析前缀的超参数对模型性能的具体影响进行了分析,其结果展示在图8中,其中红色代表模型在 Laptop-ACOS 数据集上的表现,对应右边的纵坐标,而蓝色代表模型在 Restaurant-ACOS 数据集上的表现,对应左边的纵坐标。

首先,本研究对前缀长度进行了系统性的探索,

设定前缀长度 L 遵循集合 $\{L \mid L=10 * k, k \in \mathbb{N} \wedge k \leq 15\}$, 并在此范围内采用网格搜索方法, 以评估前缀长度 L 对模型生成性能的具体影响。研究结果在数据集 Restaurant-ACOS 和 Laptop-ACOS 上的 $F1$ 分数中得到了体现。从图 8(a) 中可以明显观察到, 随着前缀长度 L 的逐步增加, 直至 120, 尽管存在一定的波动, $F1$ 分数整体上呈现出上升趋势。这一现象可以归因于较长的前缀能够引入更多的可训练参数, 从而赋予模型更强的能力来捕捉和动态建模上下文中的信息。

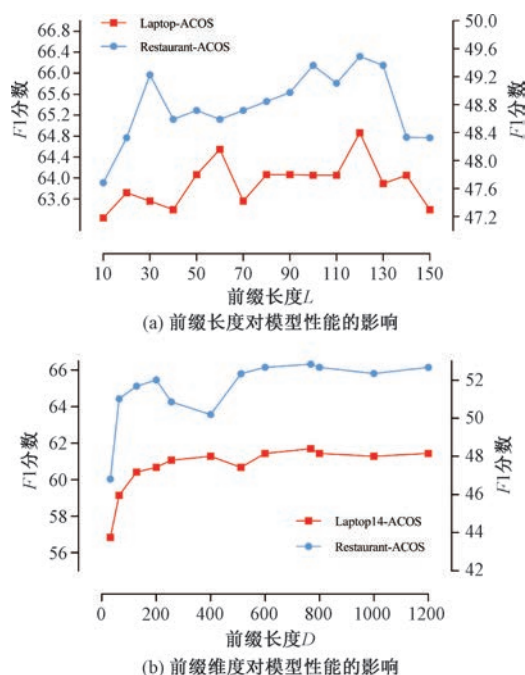


图 8 前缀超参数对模型性能的影响

然而, 当前缀长度超过 120 时, $F1$ 分数开始逐渐下降。这一趋势揭示了在本文所涉及的前缀提示中, 连续前缀的加入虽然能够增强模型的表达能力, 但过长的前缀却可能引入不必要的复杂性, 特别是在将连续前缀拼接到 Transformer 模型的历史激活

层中时, 可能导致基于 Transformer 的生成模型在生成目标序列时稳定性受损。因此, 基于上述分析, 本文推荐 Gen-CDPT 模型将前缀长度设定为 120, 以达到模型生成性能的最优化。这样的设定能够充分利用前缀带来的参数增加所带来的性能提升, 又能够避免因前缀过长而导致的模型稳定性下降的问题。

同样, 本文探讨了前缀嵌入向量维度 D 从 32 逐步增加至 1200 对模型性能的影响, 并以 Restaurant-ACOS 和 Laptop-ACOS 数据集的 $F1$ 分数作为评估指标, 结果展示在图 8(b)。研究发现, 尽管理论上更高维度的嵌入向量能够提供更丰富的信息以表达特定类别, 但实际观察到的 $F1$ 分数表现并非始终与维度成正比。具体来说, 当嵌入维度增至 768 时, $F1$ 分数显示出明显的稳定增长趋势。然而, 当维度超过 800 后, 这种增长趋势不再持续, 甚至出现了中断。

基于这些发现, 本文最终选择将 Gen-CDPT 模型中的前缀嵌入向量的维度设定为 768, 以确保模型在生成任务中达到最优性能。

5.7 数据增强策略实验分析

为了评估所提出的数据增强策略的性能提升及有效性(即方面 5), 本文进行了一系列分析实验进行深入探讨。

(1) 数据增强前后的性能对比

所提出的数据增强策略旨在解决数据中隐含情感不平衡的问题, 为此, 本文针对不同的数据类型进行了深入研究, 如表 6 所示, 其中, -llm 表示本文所提出的数据增强方式, -replace 和 -translate 分别表示同义词替换和回译的数据增强方式。实验结果表明, 本文提出的数据增强策略取得了显著成效。具体来说, 采用该策略, 在 3 种隐含情感数据(IA&EO、EA&IO 和 IA&IO)上表现尤为突出, $F1$ 分数值均

表 6 数据增强性能对比 ($F1$ 分数)

数据集	方法	情感类型			
		EA&EO (显式情感)	IA&EO (隐含情感)	EA&IO (隐含情感)	IA&IO (隐含情感)
Restaurant-ACOS	Gen-CDPT(-replace)	68.39	52.3	25.75	55.55
	Gen-CDPT(-translate)	66.68	51.06	23.68	54.13
	GenDA	67.21	55.85	27.63	55.55
	Gen-CDPT(-w/o DA)	67.21	50.45	23.68	53.33
	Gen-CDPT(-LLM)	69.94	55.85	27.63	56.66
	Gen-CDPT(-replace)	54.82	53.91	28.89	42.19
Laptop-ACOS	Gen-CDPT(-translate)	52.21	54.82	33.77	37.5
	GenDA	53.96	53.91	34.22	42.19
	Gen-CDPT(-w/o DA)	54.82	53.91	28.89	40.62
	Gen-CDPT(-LLM)	55.25	58.59	35.11	43.75

有所提升。可以发现在两个数据集上,在隐含情感数据中 F1 分数平均提升了 4.23 和 4.67。注意到模型在显式情感上的性能也获得了提升,这是因为部分评论中既包含显式情感也包含隐含情感。通过扩充这类评论,模型在处理显式情感数据方面的性能也将得到进一步提升。本文针对其他的相关的数据增强工作同样进行了对比实验,可以看到本文所提出的数据增强的方法在显式和隐含情感上的表现不仅超越了同义词替换和回译的数据增强方法,同时也超越了直接基于大模型生成的数据增强方法^[36],这一结果说明微调后的大语言模型能对齐数据扩充任务,进一步表明这种的基于大语言模型数据增强手段是一种高效且实用的方法。

(2) 增强数据分布的可视化

为了验证增强数据是否维持了原始数据的关键特征,本文对通过构造语义相似度热力图对原始数据与增强数据的分布差异进行直观分析,热力图的横轴和竖轴分别表示原始数据与和原始数据标签一致的增强数据。随机选择 20 个样本后对其构造热力图,如图 9 所示。

经过观察发现,在数据空间中,句子语义最相似的区域主要集中在对角线附近,而远离对角线的区域则呈现出明显的语义相似度降低趋势。这一结果表明,本文所采用的数据增强方法成功地最大限度地保留了原始数据的关键分布特性,证实了本文数

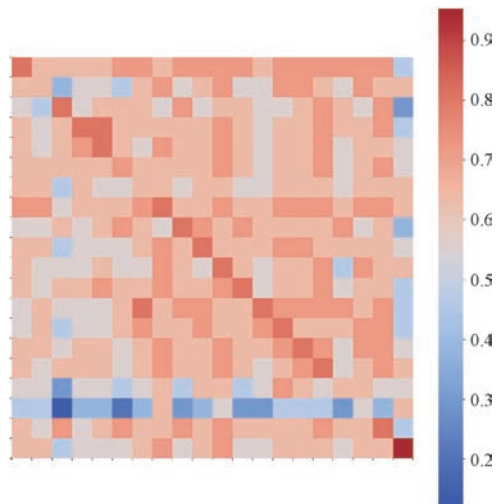


图 9 原始数据与增强数据句子之间的语义相关性热力图

据增强方法的可靠性。

(3) 增强数据样例展示

本文展示了部分增强数据样例,以直观展示原始数据与增强数据的形式。随机选取了几组与原始数据标签一致的增强数据样本进行展示,如图 10 所示。可以直观地观察到,相较于传统的数据增强方法(例如采用回译、同义词替换),基于大模型生成的增强数据展现出丰富的语义多样性。同时,这些增强数据在语义上与原始数据保持了一定程度的相似性。这进一步说明所采用数据增强策略能够有效提升模型的性能和鲁棒性。

标签 (情感四元组)	原始数据	大模型增强数据
(restaurant general, NULL, negative, doubt)	And I hate to say this but I doubt I'll ever go back .	I doubt it's what they call authentic .
(food quality, dinner, negative, NULL) (food prices, dinner, negative, NULL)	If it seemed possible to do so while there I would have fought my bill since my dinner portion of my meal was inedible !	Can't even justify the \$ 250 I spent on this dinner !!!
(restaurant general, NULL, positive, NULL)	You will go back again and again .	No doubt you'll be back !

图 10 原始数据与增强数据样例展示

综上所述,该实验结果直观展示了本文提出的数据增强策略的可靠性,并有效提升模型在处理隐含情感数据时的性能。

6 相关工作

最近的研究^[41-46]表明,情感元素的多重提取在分析基于方面的情感原因是有效的。为了全面分析评论中的情感,部分研究者^[30,47]提出了方面—类别—观点—情感四元组(ACOS)任务。而先前的工作

采用基于任务交互的方法建模标注任务。如 Cai 等人^[14]采用一个位置感知的端到端模型联合四元组提取的联合进行,Zhang 等人^[48]则采用两阶段的多任务学习方法对情感元素进行标记,但因为多任务交互、情感元素空间搜索的困难,这些模型效果始终欠佳。

近期,生成式方法在多情感元素提取领域引起了广泛关注。这些方法通过将标注任务转化为生成任务,采用如 BART^[49]或 T5^[26]等预训练生成模型来直接生成情感四元组,从而简化了任务的处理和

交互流程。同时,这些方法还有效地利用了文本中蕴含的丰富语义信息,以增强生成结果的准确性和丰富性。Yan 等人^[50]设计了统一的生成框架来提取多个情感元素;Zhang 等人^[5]尝试使用提示短语构建情感元素生成目标,以直接提取情感四元组。Wang 等人^[33]提出了基于多任务指令调优的生成框架进行提取。同时,最近的相关研究者^[3,32,38,51]采用设计离散提示的方式提取相应的四种情感元素,Zhou 等人^[39]通过引入潜在类别分布和约束解码策略来解决类别语义包含和重叠的问题,进而采用生成模型进行情感四元组的直接生成,而 Bao 等人^[34]则采用语法树作为生成的目标结构进行情感四元组提取任务。Guo 等人^[6]和 Jun 等人^[7]改进了提示模板,采用自适应算法生成离散模板进行情感四元组提取,彭等人^[52]将四元组预测建模为完形填空任务,设计离散型和连续型的混合提示,提出 C-ASQP 框架集成判别式模型和提示增强的生成式模型,以提升元素间语义关系捕捉能力。另一方面,部分研究工作着眼于从数据层面提升情感四元组提取任务的性能。Zhang 等人^[37]提出了一种采用连接操作的自适应数据增强框架,旨在增强情感四元组任务的数据。Wang 等人^[36]则提出了一个基于四元组到文本的模型,用于扩充不平衡的数据集。在获得增强数据后,进一步利用生成模型来提取四元组,Zhong 等人^[53]提出了一种基于迭代数据生成框架的方法,通过从无监督的句子语料库中提取情感元素,然后利用大型语言模型生成伪标签数据,以增强情感四元组提取任务的数据集。

将情感四元组提取任务转化为生成问题已成为一种流行趋势。然而,目前这类方法也存在一些局限性,比如它们可能依赖于特定的任务提示设计,这导致提示需要人为设计,同时这些设计的提示无法进一步针对任务优化,这会导致出现提示次优。另外也未充分考虑到隐含情感数据的不平衡问题。为了克服这些局限性,提升情感四元组提取任务的性能,本文进一步探索和改进了生成式方法,以更好地适应复杂的情感分析任务的需求。

7 总 结

本文采用生成式的架构来支持情感四元组提取任务。首先提出了基于离散提示的基础生成模型用于方面—类别—情绪—观点四元组提取。针对现有方法的提示次优问题,我们继续在基础生成模型之

上提出了一种基于上下文动态前缀提示方法 Gen-CDPT。此外,该方法构建了连续可调的前缀来建模特定类别的提示信息,以优化离散提示;另一方面,Gen-CDPT 在模型动态计算出特定上下文的提示前缀,并利用注意力机制整合类别信息和上下文信息,使模型能够动态地关注和生成与当前上下文主题相关的情感四元组,实现对提示的进一步优化。

为了解决数据中隐含情感不平衡的问题,本文提出了一种基于大语言模型的数据增强方法,通过微调、生成和后筛选机制实现对数据长尾分布问题的缓解,提升了模型在隐含情感数据上的性能表现。

最后通过一系列实验,验证了本文方法在情感四元组生成能力上的提升,以及基于大语言模型的数据增强策略的有效性。这些结果充分证明了本文所提出方法的卓越性和实用性。

参 考 文 献

- [1] CHEN W, LIU Y, ZHANG Z, et al. Modeling adaptive inter-task feature interactions via sentiment-aware contrastive learning for joint aspect-sentiment prediction; 16//Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver, Canada, 2024, 38: 17781-17789
- [2] Zhang Huan-Xiang, Peng Jun-Jie. A deep semantic mining model based on aspect-level sentiment analysis. Acta Electronica Sinica, 2024, 52(7): 2307-2319. (in Chinese)
(张焕香, 彭俊杰. 基于方面级情感分析的深度语义挖掘模型. 电子学报, 2024, 52(7): 2307-2319)
- [3] LIU Y, ZHONG J. Implicit sentiment extraction using structure generation with sentiment instructor prompt template//Advanced Data Mining and Applications; 19th International Conference, ADMA 2023 (ADMA)). Shenyang, China, 2023: 94-108
- [4] ZHANG W, LI X, DENG Y, et al. Towards generative aspect-based sentiment analysis//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP) (Volume 2: Short Papers). Online, 2021: 504-510
- [5] ZHANG W, DENG Y, LI X, et al. Aspect sentiment quad prediction as paraphrase generation//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Online, 2021: 9209-9219
- [6] GOU Z, GUO Q, YANG Y. MvP: Multi-view prompting improves aspect sentiment tuple prediction//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 1: Long Papers). Toronto, Canada, 2023: 4380-4397
- [7] JUN Y, LEE H. Dynamic order template prediction for gener-

- ative aspect-based sentiment analysis. <http://arxiv.org/abs/2406.11130>
- [8] LI X L, LIANG P. Prefix-Tuning: Optimizing continuous prompts for generation//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP) (Volume 1: Long Papers). Online, 2021: 4582-4597
- [9] LIU X, HUANG H, SHI G, et al. Dynamic prefix-tuning for generative template-based event extraction//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 1: Long Papers). Dublin, Ireland, 2022: 5216-5228
- [10] ZHANG Z R, TAN C, XU H, et al. Towards adaptive prefix tuning for parameter-efficient language model fine-tuning//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 2: Short Papers). Toronto, Canada, 2023: 1239-1248
- [11] ZHANG W, LI X, DENG Y, et al. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. <http://arxiv.org/abs/2203.01054>
- [12] WEI K, SUN X, ZHANG Z, et al. Trigger is not sufficient: Exploiting frame-aware knowledge for implicit event argument extraction//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). Online, 2021: 4672-4682
- [13] VERMA K, DAVIS B. Implicit aspect-based opinion mining and analysis of airline industry based on user-generated reviews. *SN Computer Science*, 2021, 2(4): 286
- [14] CAI H, XIA R, YU J. Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions//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). Online, 2021: 340-350
- [15] Chen Zhuang, Qian Tie-Yun, Li Wan-Li, et al. A survey on low-resource aspect-based sentiment analysis. *Chinese Journal of Computers*, 2023, 46(7): 1445-1472. (in Chinese)
(陈壮, 钱铁云, 李万理等. 低资源方面级情感分析研究综述. *计算机学报*, 2023, 46(7): 1445-1472)
- [16] HO N, SCHMID L, YUN S Y. Large language models are reasoning teachers//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 1: Long Papers). Toronto, Canada, 2023: 14852-14882
- [17] MAGISTER L C, MALLINSON J, ADAMEK J, et al. Teaching small language models to reason//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 2: Short Papers). Toronto, Canada, 2023: 1773-1781
- [18] WEI K, SUN X, ZHANG Z, et al. Implicit event argument extraction with argument-argument relational knowledge. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(9): 8865-8879
- [19] YU P, XU J, WESTON J, et al. Distilling System 2 into System. <http://arxiv.org/abs/2407.06023>
- [20] DEEPSEEK-AI, GUO D, YANG D, et al. DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. <http://arxiv.org/abs/2501.12948>
- [21] MUENNIGHOFF N, YANG Z, SHI W, et al. s1: Simple test-time scaling. <http://arxiv.org/abs/2501.19393>
- [22] BAI J, BAI S, CHU Y, et al. Qwen technical report. <http://arxiv.org/abs/2309.16609>
- [23] DONG C, LI Y, GONG H, et al. A survey of natural language generation. *ACM Computing Surveys*, 2022, 55(8): 173:1-173:38
- [24] HOLTZMAN A, BUYS J, DU L, et al. The curious case of neural text degeneration//Proceedings of the 8th International Conference on Learning Representations (ICLR). Addis Ababa, Ethiopia: OpenReview.net, 2019
- [25] ARTHUR D, VASSILVITSKII S. K-means++: The advantages of careful seeding//Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA). New Orleans, USA, 2007: 1027-1035
- [26] RAFFEL C, SHAZEER N, ROBERTS A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. <http://arxiv.org/abs/1910.10683>
- [27] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL), Volume 1 (Long and Short Papers). Minneapolis, USA, 2019: 4171-4186
- [28] LOSCHILOV I, HUTTER F. Decoupled weight decay regularization. <http://arxiv.org/abs/1711.05101>
- [29] HU E J, SHEN Y, WALLIS P, et al. LoRA: Low-rank adaptation of large language models. <http://arxiv.org/abs/2106.09685>
- [30] PRABOWO R, THELWALL M. Sentiment analysis: A combined approach. *Journal of Informetrics*, 2009, 3(2): 143-157
- [31] YANG Y, LI K, QUAN X, et al. Constituency lattice encoding for aspect term extraction//Proceedings of the 28th International Conference on Computational Linguistics (COLING). Barcelona, Spain (Online), 2020: 844-855
- [32] XIONG H, YAN Z, WU C, et al. BART-based contrastive and retrospective network for aspect-category-opinion-sentiment quadruple extraction. *International Journal of Machine Learning and Cybernetics*, 2023, 14(9): 3243-3255
- [33] WANG Z, XIA R, YU J. UnifiedABSA: A unified ABSA framework based on multi-task instruction tuning. <http://arxiv.org/abs/2211.10986>

- [34] BAO X, ZHONGQING W, JIANG X, et al. Aspect-based sentiment analysis with opinion tree generation//Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI). Vienna, Austria, 2022; 4044-4050
- [35] XU X, ZHANG J D, XIONG L, et al. iACOS: Advancing implicit sentiment extraction with informative and adaptive negative examples. <http://arxiv.org/abs/2311.03896>
- [36] WANG A, JIANG J, MA Y, et al. Generative data augmentation for aspect sentiment quad prediction//Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023). Toronto, Canada, 2023; 128-140
- [37] ZHANG W, ZHANG X, CUI S, et al. Adaptive data augmentation for aspect sentiment quad prediction. <http://arxiv.org/abs/2401.06394>
- [38] SEO Y, SONG S, HEO R, et al. Make compound sentences simple to analyze: Learning to split sentences for aspect-based sentiment analysis. <http://arxiv.org/abs/2410.02297>
- [39] ZHOU J, YU D, AZIZ K, et al. Generative sentiment analysis via latent category distribution and constrained decoding//Proceedings of the Artificial Neural Networks and Machine Learning - ICANN 2024. Switzerland, 2024; 209-223
- [40] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 2022, 35; 27730-27744
- [41] PENG H, XU L, BING L, et al. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis; 05//Proceedings of the AAAI Conference on Artificial Intelligence. 2020,34; 8600-8607
- [42] WAN H, YANG Y, DU J, et al. Target-aspect-sentiment joint detection for aspect-based sentiment analysis; 05//Proceedings of the AAAI Conference on Artificial Intelligence (AAAI); Vol. 34. New York, USA, 2020; 9122-9129
- [43] LIN P, YANG M. A shared-private representation model with coarse-to-fine extraction for target sentiment analysis//Findings of the Association for Computational Linguistics; EMNLP 2020 (EMNLP). Online, 2020; 4280-4289
- [44] ZHANG Y, YANG Y, LI Y, et al. Boundary-driven table-filling for aspect sentiment triplet extraction//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP). Abu Dhabi, United Arab Emirates, 2022; 6485-6498
- [45] CHEN W, LIN P, ZHANG W, et al. Hierarchical interactive network for joint aspect extraction and sentiment classification. *Knowledge-Based Systems*, 2022, 256; 109825
- [46] LIANG S, WEI W, MAO X L, et al. STAGE: Span tagging and greedy inference scheme for aspect sentiment triplet extraction//Proceedings of the AAAI Conference on Artificial Intelligence. 2023,37(11);13174-13182
- [47] DONG L, WEI F, TAN C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL) (Volume 2: Short Papers). Baltimore, Maryland, 2014; 49-54
- [48] ZHANG H, CHEN Z, CHEN B, et al. Complete quadruple extraction using a two-stage neural model for aspect-based sentiment analysis. *Neurocomputing*, 2022, 492; 452-463
- [49] LEWIS M, LIU Y, GOYAL N, et al. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online, 2020; 7871-7880
- [50] YAN H, DAI J, JI T, et al. A unified generative framework for aspect-based sentiment analysis//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). Online, 2021; 2416-2429
- [51] GAO T, FANG J, LIU H, et al. LEGO-ABSA: A prompt-based task assemblable unified generative framework for multi-task aspect-based sentiment analysis//Proceedings of the 29th International Conference on Computational Linguistics (COLING). Gyeongju, Republic of Korea, 2022; 7002-7012
- [52] Peng Wen-Zhong, Xia Jia-Li, Wan Qi-Zhi, et al. Modeling aspect sentiment quad prediction as cloze task. *Chinese Journal of Computers*, 2024, 47(8); 1744-1762. (in Chinese)
(彭文忠, 夏家莉, 万齐智等. 基于完形填空的方面级情感四元组预测[J]. *计算机学报*, 2024, 47(8); 1744-1768)
- [53] ZHONG Q, LI H, ZHUANG L, et al. Iterative data generation with large language models for aspect-based sentiment analysis. <http://arxiv.org/abs/2407.00341>

ZHONG Jiang, Ph. D., professor.

His research interests include big data analysis and mining, natural language processing, cloud network convergence, network security and so on.

LIU Yu-Xuan, M. S. His research

interests include sentiment analysis and information extraction.



DAI Qi-Zhu, Ph. D. His research interest is relation extraction.

WANG Jia-Qi, M. S. Her research interest is machine reading comprehension.

LAI Xin-Yi, M. S. Her research interest is image caption.

HU Wen-Yue, M. S. Her research interest is relation extraction.

Background

This paper delves into the domain of Aspect-Based Sentiment Analysis (ABSA), focusing on a particularly challenging issue-Sentiment Quadruple Extraction. ABSA is a subfield of natural language processing (NLP) that goes beyond traditional sentiment analysis by identifying and extracting specific aspects within text that contribute to an overall sentiment. The extraction of sentiment quadruples, which includes the aspect, category, opinion, and sentiment itself, provides a comprehensive understanding of the fine-grained sentiments expressed in textual data.

Internationally, the research on ABSA has made significant strides, with various methods being proposed to address the complexity of extracting sentiment quadruples. However, existing solutions often grapple with the limitations of prompt design in sequence-to-sequence methods and the challenge of imbalanced implicit sentiment data, which can hinder performance when processing such data.

In this paper, we advance the state-of-the-art by introducing a novel Dynamic Prefix Prompt method coupled with data augmentation strategies. Our approach dynamically optimizes prompts using adjustable prefixes and attention mechanisms, addressing the suboptimal prompt issue. Fur-

thermore, we leverage large language models for data augmentation, effectively balancing implicit sentiment data through fine-tuning. Our experiments on two real-world datasets demonstrate a notable improvement in $F1$ scores, achieving state-of-the-art performance and showcasing the effectiveness of our method.

This research was supported in part by the National Natural Science Foundation of China (No. 62176029) and the Special Program for Scientific and Technological Innovation and Applied Development of Chongqing Municipality (CSTB2023TIAD-KPX0064, CSTB2022TIAD-KPX0206). The project aims to push the boundaries of ABSA, facilitating better understanding and processing of user-generated content, which is vital in today's data-driven decision-making processes.

The significance of this project lies in its potential to transform the way businesses and researchers understand and utilize customer feedback and opinions expressed in text form. By enhancing the extraction of sentiment quadruples, our work paves the way for more accurate sentiment analysis tools that can be applied across various industries, from e-commerce product reviews to social media monitoring.