

基于方面语义和门控过滤网络的方面级情感分析

何智豪, 陈红梅, 罗川

引用本文

何智豪, 陈红梅, 罗川. 基于方面语义和门控过滤网络的方面级情感分析[J]. 计算机科学, 2023, 50(10): 193-202.

HE Zhihao, CHEN Hongmei, LUO Chuan. [Aspect-based Sentiment Analysis Based on Aspect Semantic and Gated Filtering Network](#) [J]. Computer Science, 2023, 50(10): 193-202.

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)

[基于多粒度特征融合的新型图卷积网络用于方面级情感分析](#)

Novel Graph Convolutional Network Based on Multi-granularity Feature Fusion for Aspect-based Sentiment Analysis

计算机科学, 2023, 50(10): 80-87. <https://doi.org/10.11896/jsjcx.230600036>

[融合语义和句法图神经网络的实体关系联合抽取](#)

Fusion of Semantic and Syntactic Graph Convolutional Networks for Joint Entity and Relation Extraction

计算机科学, 2023, 50(9): 295-302. <https://doi.org/10.11896/jsjcx.220700041>

[基于并行卷积网络信息融合的层级多标签文本分类算法](#)

Hierarchical Multi-label Text Classification Algorithm Based on Parallel Convolutional Network Information Fusion

计算机科学, 2023, 50(9): 278-286. <https://doi.org/10.11896/jsjcx.221200133>

[基于深度学习的红外视频显著性目标检测](#)

Deep Learning Based Salient Object Detection in Infrared Video

计算机科学, 2023, 50(9): 227-234. <https://doi.org/10.11896/jsjcx.220700204>

[邻域双向聚合与全局感知的TKG链接预测模型](#)

Link Prediction Model on Temporal Knowledge Graph Based on Bidirectionally Aggregating Neighborhoods and Global Aware

计算机科学, 2023, 50(8): 177-183. <https://doi.org/10.11896/jsjcx.220900061>

基于方面语义和门控过滤网络的方面级情感分析

何智豪¹ 陈红梅² 罗 川³

1 西南交通大学唐山研究院 河北 唐山 063000

2 西南交通大学计算机与人工智能学院 成都 611756

3 四川大学计算机学院 成都 610065

(zhhe@my.swjtu.edu.cn)

摘 要 方面级情感分析(Asspect-based Sentiment Analysis, ABSA)是一项细粒度的情感分析任务,旨在预测文本中特定方面的情感极性。目前,鉴于循环神经网络在序列建模方面的卓越性能以及卷积神经网络学习局部模式的出色表现,部分工作将两者相结合来挖掘情感信息,并且取得了不错的效果。但是,少有工作在将两者结合后应用到方面级情感分析任务中的同时考虑方面信息。在方面级情感分析任务中,大部分工作将方面视作一个独立整体与上下文进行交互,但是对于方面的表示过于简单,缺乏真实语义。针对上述问题,文中提出了一种基于方面语义和门控过滤网络(Asspect Semantic and Gated Filtering Network, ASGFN)的神经网络模型,用于挖掘方面级情感信息。首先,设计了方面编码模块,用于捕捉特定语境下的方面语义信息,该模块基于全局上下文融合多头注意机制与图卷积神经网络构建包含特定语义的方面表示。其次,设计门控过滤网络连接循环神经网络和卷积神经网络,以此增强方面与上下文的交互,同时结合循环神经网络与卷积神经网络的优势,进而提取情感特征。最后,将情感特征与方面表示相结合,生成预测情感极性的语义表征。在 restaurant, laptop 和 twitter 这 3 个公用数据集上分别取得了 84.72%, 78.64%, 76.22% 的情感分类准确率。实验结果表明了所提模型的有效性,它能提高方面级情感分类任务的性能。

关键词: 方面级情感分析; 方面语义; 门控过滤; 循环神经网络; 卷积神经网络

中图法分类号 TP391

Aspect-based Sentiment Analysis Based on Aspect Semantic and Gated Filtering Network

HE Zhihao¹, CHEN Hongmei² and LUO Chuan³

1 Tangshan Research Institute, Southwest Jiaotong University, Tangshan, Hebei 063000, China

2 School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

3 College of Computer Science, Sichuan University, Chengdu 610065, China

Abstract Aspect-based sentiment analysis(ABSA) is a fine-grained sentiment analysis, which aims to predict sentiment polarity of text toward a specific aspect. Currently, given the excellent capabilities of recurrent neural networks(RNN) in sequence modeling and the outstanding performance of convolutional neural networks(CNN) in learning local patterns, some works have combined the two to mine sentiment information and achieved good results. However, few works consider aspect information while applying the combination of the two to ABSA. In aspect-based sentiment analysis tasks, most of the work treat aspect as an independent whole interacting with the contexts, but the representation of aspect is too simple and lacks real semantic. To address the above issues, this paper proposes a neural network model based on aspect semantic and gated filtering network(ASGFN) to better mine aspect-based sentiment information. First, an aspect encoding module is designed to capture context-specific aspect semantic information, which is based on a global context fusion multi-head attention mechanism with a graph convolutional neural network to construct aspect representation containing specific semantic. Second, a gated filtering network is designed to connect RNN and CNN as a way to enhance the interaction of aspect with the contexts, while combining the advantages of the RNN and the CNN, and then extracting the sentiment feature. Eventually, the sentiment feature is combined with aspect representation to generate semantic representation that predicts sentiment polarity. Sentiment classification accuracies of 84.72%, 78.64%, and 76.22% are achieved in three communal datasets, restaurant, laptop, and twitter, respectively. Experimental results demonstrate the effective-

到稿日期: 2022-09-20 返修日期: 2022-12-07

基金项目: 国家自然科学基金(61976182, 62076171); 四川省自然科学基金(2022NSFSC0898)

This work was supported by the National Natural Science Foundation of China(61976182, 62076171) and Natural Science Foundation of Sichuan Province, China(2022NSFSC0898).

通信作者: 陈红梅(hmchen@swjtu.edu.cn)

ness of the proposed model, which can improve the performance of ABSA.

Keywords Aspect-based sentiment analysis, Aspect semantic, Gated filtering, RNN, CNN

1 引言

情感分析是文本分类的一个细分领域,其模型构建于人工智能、语言文学、统计学等多领域的研究成果之上。根据分类对象粒度的不同,情感分析主要分为粗粒度情感分析和细粒度情感分析。其中方面级情感分析便是一种细粒度情感分析任务,其目的是挖掘文本中所涉及的特定方面的情感信息^[1]。方面级情感分析又包含多个子任务,其中最主要的是方面抽取、方面类别检测、观点词抽取、情感分类这4类任务^[2-5]。其中方面类别检测任务的目标是检测文本中所描述的实体或方面所属的类别,例如方面“pizza”的类别为食物类。部分研究者为了揭示更全面的方面情感层次结构,将上述4个子任务联合学习,从而抽取四元组(方面、方面类别、观点词、情感极性)^[6-7]。而本文主要在方面已知的前提下探讨特定方面的情感分类任务。如图1所示,在句子“Great food, but the service was dreadful”中可以明显看到,方面“food”由带有积极情感色彩的观点词“great”描述,因此将其情感极性归类为积极。但是,由于描述方面“service”的观点词为“dreadful”,因此将方面“service”的情感极性分类为消极。

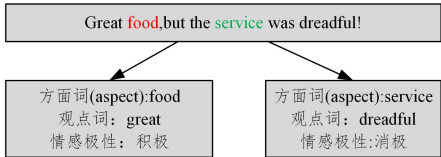


图1 方面级情感分析样例

Fig. 1 Example of aspect-based sentiment analysis

在早期的工作中,大多数方法都侧重于基于提取的特征构建有效的分类器^[8-11]。但是,这类方法高度依赖于特征工程的质量,存在特征构建的劳动密集型与抽取方面及其相关上下文词之间语义关系能力较弱的问题。

最近,随着深度学习的快速发展,很多研究者利用循环神经网络^[12](Recurrent Neural Network, RNN)、卷积神经网络^[13](Convolutional Neural Network, CNN)等深度学习模型处理NLP领域的各种问题,因为这些深度学习模型能够从上下文自动学习分布式表示^[14-15]。对于RNN来说,出色的序列建模能力使其能够在很多自然语言处理任务上发挥较大的作用。然而,为了缓解RNN的梯度消失等问题,长短期记忆网络(Long-Short Term Memory Neural Network, LSTM)被提出,用于序列建模^[16]。相比RNN, CNN具有更强的局部模式学习能力。CNN通过滑动窗口在同一时间内处理多个单词,从而捕获重要的局部特征。但是, CNN却忽略了上下文之间的时序关系。基于此,不少研究者将两者结合,用于解决文本情感分析任务,并且取得了较好的结果^[17-19]。但是,在ABSA任务中,却少有工作在融合两者的同时考虑到方面信息对情感极性判断的影响。前人的工作已经证实探讨方面词与上下文的联系有助于特定方面情感特征的提取^[5, 20]。

为了考虑方面信息,很多工作提出了各种模型来探索

方面和上下文的联系,但是这些模型仅通过固定的词嵌入(如Glove词向量^[21])表示方面,然后将其直接用于与上下文进行交互。由于固定的词向量是在固定语料下训练得到的,因此同一个词或者在固定词嵌入下语义相似的词,在上下文语境变化较大时,可能会有较大的语义差异。Liang等在其研究中指出,在很多前人的研究中,部分样本能够预测正确的原因是它们的情感极性是与方面无关的。但是对于需要根据特定语境下的方面语义去判断情感极性的样本表现得并不理想^[22]。因此,我们需要根据上下文语境先捕获方面语义信息,然后判断其情感极性。

针对上述问题,本文提出了一个基于方面语义和门控过滤网络(ASGFN)的网络模型,用于ABSA任务。首先,将上下文单词的词性信息与相对位置信息进行编码,并将其引入本文模型。另外,为了在利用RNN和CNN各自优势的同时考虑方面信息,我们设计了一个门控过滤组件对方面和上下文进行交互,过滤噪音,并将其用于连接RNN和CNN,通过RNN进行序列建模,由CNN抽取局部特征,进而提取情感特征。在确定文本中某个单词的语义时,我们需要考虑上下文信息。基于此,我们根据上下文信息对方面语义进行建模。鉴于图卷积神经网络(Graph Convolutional Networks, GCN)提取特征以及解决长距离依赖问题的优秀能力^[23],我们将整个文本视作一个网络,将单词视作网络中的节点,以方面为中心节点,利用GCN聚合上下文节点的信息,从而更新方面表示。但是传统GCN总是以相同的权值聚合邻居节点,这在一定程度上会引入噪音,因此我们在更新节点表示之前引入多头自注意机制来计算各个节点的初始权值。基于此,我们结合上下文对方面进行了语义建模,得到了具有特定语义的方面表示。最后将提取的情感特征与方面表示联合作为最终的语义表征,用于情感分类。为了评估本文方法,我们在不同的数据集上与相关算法进行了对比实验。实验结果表明,本文方法在所有数据集上的性能都优于对比方法。本文的主要贡献如下:

- 1)设计了一个门控过滤组件,将RNN和CNN融合引入ABSA任务的同时,还考虑了方面与上下文的交互。
- 2)首次提出先捕获特定语境下的方面语义信息,然后将其用于与上下文进行交互,从而抽取特定方面的情感特征。
- 3)本文方法在2014年SemEval的笔记本电脑和餐厅数据集以及twitter数据集上,与相关算法进行了比较,实验结果验证了所提方法的有效性。

本文第2章介绍了相关的工作;第3章介绍了所提模型的具体细节;第4章介绍了数据集、实验设置、实验结果、结果分析以及案例研究;最后总结全文并展望未来。

2 相关工作

情感分析是自然语言处理领域的一个基础性和实践性的研究分支。近年来,神经网络在自然语言处理领域得到了广泛的应用。神经网络自动提取任务相关特征信息的能力引起

了众多研究者的关注。本章主要分两部分来回顾与本文研究最相关的面向方面级情感分析的神经网络的研究工作。

2.1 基于注意力机制的方面级情感分析

在早期的研究中,Tang 等认为探索方面与上下文的关系有助于情感特征的提取,因此提出了基于长短期记忆网络的与目标相关的情感分类模型(Target-Dependent Long-Short Term Memeoy Networks,TD-LSTM),该模型将方面分别与左上下文和右上下文拼接,并将其馈入 LSTM 网络抽取方面对应情感特征,这是第一次考虑将方面与上下文联合建模来提高分类精度^[5]。但是,其对于方面项与上下文交互的方式过于简单。随着注意力机制被成功地应用于机器翻译和问答系统中^[24-25],注意力机制开始成为方面和上下文交互的主要工具。

Wang 等提出了基于注意力的长短期记忆网络(ATAE-LSTM),以方面为查询键,引入注意力机制查询与方面高度相关的上下文,其中对方面项内所有单词的词嵌入求平均,并将其作为方面表示^[20]。Ma 等提出了交互注意网络(Interactive Attention Networks,IAN),并行生成方面表示和上下文表示^[26]。Yang 等认为 Ma 等提出的交互机制比较简单,因此提出了一种对方面和上下文交替建模的协同注意机制(Coattention),串行生成方面和上下文表示^[27]。Jiang 等利用 Bi-LSTM 获取句子及其各自的方面词的语义依赖关系,然后通过互注意机制学习方面的情感极性^[28]。为了防止重要的语义信息随着网络层数加深而丢失,一部分研究者提出了记忆网络用于 ABSA 任务。Tang 等提出了一种由多个计算层组成的深度存储网络(Deep Memory Network,MemNet),每一层都是一个注意力模块,以方面信息为参考,明确揭示了句子中每个单词的重要性^[29]。Lv 等的工作与 Tang 等的工作相似,提出了上下文和方面记忆网络(Context and Aspect Memory Network,CAMN)模型,引入多头注意机制来获取文本在不同子空间中的特征表示,使用多跳机制更新方面表示^[30]。受残差网络的启发,Wu 等通过残差注意机制弱化方面与上下文交互过程中信息的损失^[31]。Xu 等提出了多注意网络(Multi-Attention Network,MAN),该网络使用全局注意模块和局部注意模块来捕捉上下文和方面之间不同粒度的交互信息^[32]。这些方法都是将方面单独抽取出来视作一个独立的整体,然后将其与上下文联合建模,从而指示上下文中重要的情感词,这也验证了方面信息的重要性。但是,上述大部分方法中存在的一个共同问题就是,用于与上下文交互的方面表示都过于简单,仅通过固定词嵌入表达其语义,并未考虑真实的上下文语境,从而可能会导致部分语义的丢失。

2.2 基于 CNN 和 RNN 的方面级情感分析

自 Collobert 等的开创性工作以来,CNN 被成功应用到自然语言处理的各个领域,CNN 在情感分析中更善于捕捉 n 元语法特征,提取潜在的语义特征。Xue 等提出了一种基于卷积神经网络和门控机制(Gated Convolutional Network with Aspect Embedding,GCAE)的模型,该模型利用两个卷积层分别对方面和上下文进行编码,然后利用门控机制控制方面和上下文的权重^[33]。

考虑到 RNN 和 CNN 各自的优势和局限性,部分研究者

将 LSTM 与 CNN 结合,并将其应用到情感分析任务中。Bassiri 等认为 RNN 高维的输入空间增加了模型的复杂度,于是将 CNN 和 RNN 结合,从而降低特征空间的维度,然后引入注意力机制关注重要特征^[17]。Li 等针对句子级情感分析任务,提出了一个 CNN-LSTM 双通道的模型,并且融合情感词典信息获取分类结果^[18]。Hu 等将 RNN 和 CNN 结合,挖掘文本序列特征和局部特征,并且融入注意力机制来增强模型对特征的筛选和保留能力^[19]。这些方法都是 RNN 和 CNN 在句子级情感分析任务中的应用,并且取得了不错的效果,这也从一定程度上验证了 RNN 和 CNN 结合的有效性。ABSA 任务相比句子级情感分析任务来说,处理要更复杂一些,但是也有一部分工作将 RNN 和 CNN 结合引入 ABSA 任务。Piao 等提出了一种同时预测方面和情感的方法,将 CNN 和 RNN 以单通道的方式结合用于方面级情感分析,结合投票策略选取分类结果^[34]。Sun 等将区域卷积神经网络(Regional Convolution Neural Network,RCNN)与融入了高速公路网络的 RNN 结合,以此捕捉局部特征和时序特征^[35]。Tao 等的工作与 Sun 等的工作比较相似,通过双通道的方式结合 RCNN 和有序神经元长短时记忆来捕捉局部特征和时序特征,然后引入注意机制融合两个通道生成的语义信息^[36]。这些方法都是以单通道或者双通道的方式结合 CNN 和 RNN 来抽取局部和时序特征,但是它们都忽略了方面信息,削弱了提取特定方面对应情感特征的能力。

3 ASGFN 模型

本章将详细描述所提出的 ASGFN 模型的结构,模型的整体结构如图 2 所示。

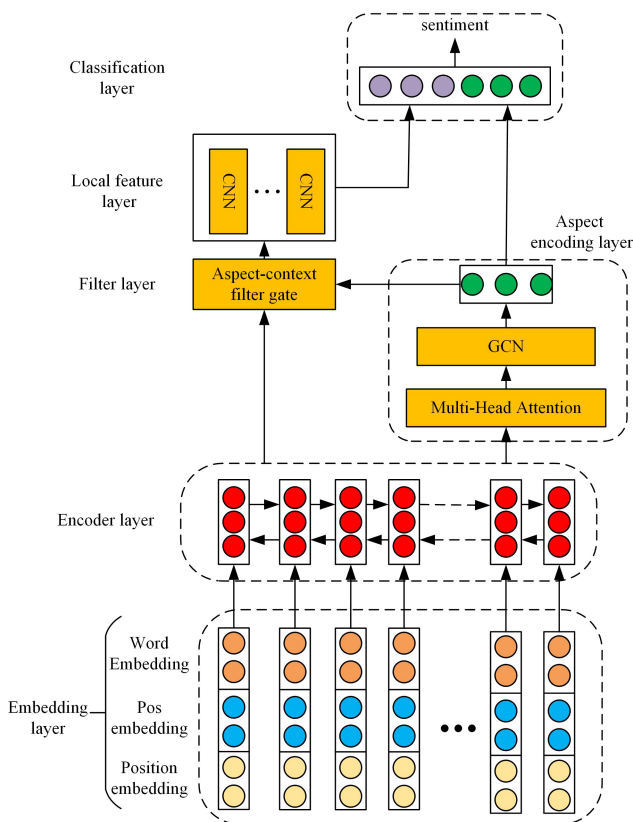


图2 ASGFN 的整体结构

Fig.2 Overall structure of ASGFN

模型由 6 部分组成,嵌入层包括词嵌入、词性嵌入和位置嵌入 3 部分。语义编码层应用 Bi-LSTM 对上下文进行编码,获取上下文中每个时间步的隐藏状态。方面编码模块引入多头注意机制和图卷积神经网络,联合上下文构建具有特定语义的方面表示。过滤层基于门控思想,设计了一个过滤门网络,用于过滤与当前方面无关的情感信息。局部特征提取层在过滤门的基础上,利用多粒度的 CNN 抽取 n 元语法特征,进而提取情感特征。输出层将情感特征和方面表示相结合,并将其作为最终的语义表征用于情感极性预测。

在数学上,我们定义了方面级情感分析任务,给定一个句子-方面对 (S, A) , 其中方面 $A = \{w_a^1, w_a^2, \dots, w_a^m\}$ 是句子 $S = \{w_c^1, w_c^2, \dots, w_c^n\}$ 的子序列。方面级情感分析任务的目标是预测句子 S 中方面 A 的情感极性,其中 0, 1, 2 分别代表积极、消极、中性 3 种情感极性。

3.1 嵌入层

本节简要描述嵌入层,包括词嵌入、词性嵌入和位置嵌入,旨在将每个上下文词本身、词的词性以及其相对于方面的相对位置分别映射到低维实值向量。

3.1.1 词嵌入层

将文本序列 $S = \{w_c^1, w_c^2, \dots, w_c^n\}$ 送入模型之前,我们需要利用词嵌入方法对文本进行处理。令 $E \in \mathbb{R}^{d \times |V|}$ 为预训练得到的查找词嵌入矩阵,该嵌入矩阵主要通过 Glove 获得,其中 d 为词嵌入向量的维度, $|V|$ 是词表的大小。通过嵌入矩阵将 S 中的每个单词映射到对应词向量,获得一个向量列表 $v_s = \{e_1, e_2, \dots, e_n\}$, e_i 表示第 i 个单词的词嵌入。

3.1.2 词性嵌入层

将文本 S 送入 Stanford Parse¹⁾ 解析器,解析获得文本中每个单词的词性,以此建立词性类别索引向量 $c = \{c_1, c_2, \dots, c_n\}$, 其中 c_i 代表第 i 个单词的词性类别的索引。由于 c_i 仅为一个独热向量,无法较好地表示其对应的词性信息,因此通过可训练参数矩阵 W_c 将 c_i 转换为低维实值向量,其对应的词性向量可表示为:

$$p_i = c_i W_c \quad (1)$$

其中, $W_c \in \mathbb{R}^{d_p \times |V_p|}$, d_p 为词性嵌入向量的维度, $|V_p|$ 为所有词性类别的数量。

3.1.3 位置嵌入层

一个文本中,对特定方面表达情感倾向的词通常分布在方面的附近,而与方面距离较远的单词通常包含很少的情感信息。例如文本“Great food but the service was dreadful”,很明显,距离方面“food”更近的观点词“great”对于判断方面“food”的情感极性起着更重要的作用。同时,距离方面“service”更近的观点词“dreadful”更有利于对方面“service”的情感极性的判断。因此将上下文单词到方面的相对位置信息作为额外特征引入模型中。利用上下文单词到特定方面的相对距离表示上下文单词的位置,具体计算式如下:

$$index_i = \begin{cases} i-a, & i < a \\ 0, & a \leq i \leq a+m \\ i-m, & i > m \end{cases} \quad (2)$$

其中, i 表示单词的原索引, a 和 m 表示方面在原始序列中的开始索引和结束索引。

获取上下文单词到特定方面的相对位置之后,需要将新的位置索引通过可训练参数矩阵 W_p 转换到低维实值向量,其对应的位置向量可表示为:

$$g_i = index_i W_p \quad (3)$$

其中, $W_p \in \mathbb{R}^{d_g \times |V_g|}$, d_g 为位置嵌入向量的维度, $|V_g|$ 设置为所有样本中最长样本的长度。

获取词嵌入、词性嵌入、位置嵌入之后,将三者拼接获得文本中每个单词的向量表示 x_i , 以此作为模型的初始输入。

$$x_i = [e_i; p_i; g_i] \quad (4)$$

其中, $i \in [1, n]$, $x_i \in \mathbb{R}^{1 \times (d + d_p + d_g)}$, ‘;’ 表示向量拼接。

3.2 语义编码层

通过嵌入层获取的词嵌入无法表达其真实语义,每个单词的语义需要联合上下文进行建模。因此,我们引入 BiLSTM 对文本进行语义编码, BiLSTM 由前向 LSTM 和反向 LSTM 组成。前向 LSTM 对当前时间步的左上下文进行建模,表达式如下:

$$\vec{h}_i = \overrightarrow{\text{LSTM}}(x_i, \theta_{\text{LSTM}}), i \in [1, n] \quad (5)$$

反向 LSTM 对当前时间步的右上下文进行建模,具体表达式如下:

$$\overleftarrow{h}_i = \overleftarrow{\text{LSTM}}(x_i, \theta_{\text{LSTM}}), i \in [1, n] \quad (6)$$

其中, θ_{LSTM} 是 LSTM 模型的参数。

将隐藏状态 $\vec{h}_i$ 与 $\overleftarrow{h}_i$ 拼接得到第 i 个时间步的隐藏状态(第 i 个单词的语义表示), 即 $h_i = [\vec{h}_i, \overleftarrow{h}_i]$ 。因此,我们可以获得词嵌入 $X = [x_1, x_2, \dots, x_n]$ 的上下文语义表示 $H = [h_1, h_2, \dots, h_n]$ 。其中, $H \in \mathbb{R}^{n \times 2d_h}$, d_h 为单向 LSTM 的隐藏状态的维度。

3.3 方面编码层

在获取每个单词的隐藏状态的基础上引入 GCN 对长距离依赖进行建模,获取更深层次的单词表示。在之前的大部分工作中,主要利用 GCN 对句法依赖树进行建模,从而提取特征,并取得了不错的性能。但是由于句法解析器存在误差,单层 GCN 可能无法捕捉特定方面的情感特征,而多层 GCN 虽然能够聚合高阶邻居节点的信息,但是可能导致更多的无关信息的引入。因此,受 Li 等的工作的启发^[37],我们引入多头注意力机制对上下文进行建模,计算注意力分数矩阵 A ,将该分数矩阵视作邻接矩阵。具体计算式如下:

$$score_i = \text{Attention}(QW_i^Q, KW_i^K) \quad (7)$$

$$\text{Attention}(Q, K) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad (8)$$

$$A = \frac{\sum_{i=1}^h score_i}{h} \quad (9)$$

其中, $W_i^Q \in \mathbb{R}^{2d_h \times d_k}$ 和 $W_i^K \in \mathbb{R}^{2d_h \times d_k}$ 是不同的可训练参数矩阵, $Q = K = H$, $A \in \mathbb{R}^{n \times n}$, d_k 为输入节点特征的维度, h 为多头注意中头的数量。

获取邻接矩阵 A 之后,引入 GCN,通过聚合邻接节点的

¹⁾ <https://stanfordnlp.github.io/CoreNLP/>

信息更新当前节点的特征向量:

$$\mathbf{h}_i^{(l)} = \sigma \left(\frac{\sum_{j=1}^n \mathbf{A}_{ij} \mathbf{h}_j^{(l-1)} \mathbf{W}^l + \mathbf{b}^l}{d_i + 1} \right) \quad (10)$$

$$d_i = \sum_{j=1}^n \mathbf{A}_{ij} \quad (11)$$

其中, $\mathbf{W}^l$ 为权重矩阵, $\mathbf{h}_j^{(l-1)}$ 是上一层卷积层的输出, $\mathbf{b}^l$ 为偏置项, d_i 为节点 i 的度, σ 为激活函数, 例如 relu 等。

利用 GCN 获取新的节点表示 $\mathbf{H}^{(l)} = [\mathbf{h}_1^{(l)}, \mathbf{h}_2^{(l)}, \dots, \mathbf{h}_n^{(l)}]$ 之后, 我们利用特定方面零掩码机制遮蔽非方面词的节点特征, 仅保留具有真实语义的特定方面的节点特征, 具体表达式如下:

$$\mathbf{M}_{\text{mask}} = [0, \dots, 1, \dots, 1, \dots, 0] \quad (12)$$

$$\mathbf{H}_{\text{mask}}^{(l)} = \mathbf{M}_{\text{mask}} \mathbf{H}^{(l)} \quad (13)$$

$$\mathbf{H}_{\text{mask}}^{(l)} = [0, \dots, \mathbf{h}_a^l, \dots, \mathbf{h}_{a+m}^l, \dots, 0] \quad (14)$$

对方面内所有单词的节点向量求和取平均得到方面项的特征表示, 计算式如下:

$$\mathbf{h}_a = \frac{\sum_{i=1}^m \mathbf{h}_i^{(l)}}{m} \quad (15)$$

3.4 过滤层

对于方面级情感分析任务来说, 提取携带情感信息的 n 元语法特征是重要的, 如“not bad”等, 因此我们设计了局部特征提取层来增强 n 元语法特征的学习。首先, 我们设计了门控过滤网络来控制情感信息的流动, 结构如图 3 所示, 包括一个门和两个非线性激活函数。

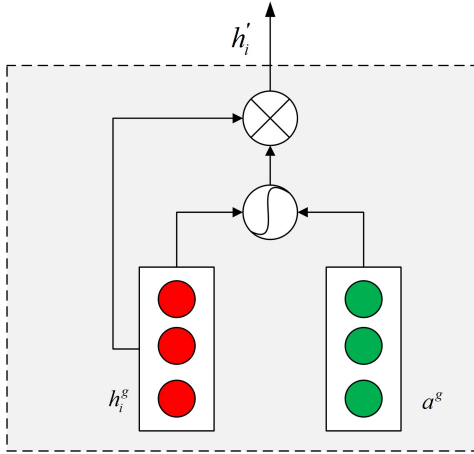


图3 过滤门

Fig. 3 Filter gate

利用门控思想设计过滤门对方面和上下文进行建模, 以此减轻噪音的影响, 增强相关情感词的贡献度。将 BiLSTM 的隐藏层的输出 H 和特定方面编码层的输出 $\mathbf{h}_a$ 输入门控过滤网络, 具体计算式如下:

$$\mathbf{h}_i^g = \mathbf{W}_h \cdot \mathbf{h}_i + \mathbf{b}_g \quad (16)$$

$$\mathbf{a}^g = \mathbf{W}_a \cdot \mathbf{h}_a + \mathbf{b}_a \quad (17)$$

$$\mathbf{f}_i = \sigma(\mathbf{h}_i^g + \mathbf{a}^g + \mathbf{b}_f) \quad (18)$$

$$\mathbf{h}_i' = \mathbf{h}_i^g \odot \mathbf{f}_i \quad (19)$$

其中, $\mathbf{W}_h$ 和 $\mathbf{W}_a$ 是权重矩阵, $\mathbf{b}_f$ 是偏置项, $\mathbf{h}_i$ 是 BiLSTM 输出的第 i 个单词的隐藏向量, $\mathbf{h}_a$ 是包含上下文语义的特征向量, σ 代表非线性激活函数 (例如 sigmoid 等), $\odot$ 表示向量元素

互乘, $\mathbf{h}_i'$ 是过滤噪音之后的单词表征。

3.5 局部特征提取层

经过门控过滤网络过滤无关情感信息之后, 我们引入 CNN 提取局部特征, 令 $\mathbf{W} \in \mathbb{R}^{k \times d}$ 为卷积核, 其中 k 为卷积核的宽度, d 为 $\mathbf{h}_i'$ 的维度。宽度为 k 的卷积核在输入序列上滑动, 提取局部特征。令 z_i 表示卷积核在单词序列 $\mathbf{h}_{i:i+k-1}'$ 上提取的局部特征, 具体计算式如下:

$$z_i = \sigma(\mathbf{W} * \mathbf{h}_{i:i+k-1}' + \mathbf{b}) \quad (20)$$

其中, σ 是非线性激活函数, $\mathbf{b}$ 是偏执项, $z_i \in \mathbb{R}$ 。卷积滤波器滑过整个文本序列后产生特征映射 $\mathbf{Z}_i = [z_1, z_2, \dots, z_{n-k+1}]$, 将输出通道数设置为 f , 则单个卷积核最终映射得到的特征空间如下:

$$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_f] \quad (21)$$

其中, $\mathbf{Z}_i \in \mathbb{R}^{1 \times (n-k+1)}$, $\mathbf{Z} \in \mathbb{R}^{(n-k+1) \times f}$ 。引入最大池化提取最显著局部特征, 池化层的滑动窗口的宽度为 $n-k+1$, 具体计算式如下:

$$\mathbf{h}_p = \text{MaxPool1d}(\mathbf{Z}) \quad (22)$$

其中, MaxPool1d 表示一维最大池化, $\mathbf{h}_p \in \mathbb{R}^{1 \times f}$ 。利用 l 个不同尺寸的卷积核提取多粒度的 n 元语法特征, 并将不同语义空间的局部特征向量拼接作为最终的局部特征, 具体表达式如下:

$$\mathbf{H}_l = [\mathbf{h}_1; \mathbf{h}_2; \dots; \mathbf{h}_l] \quad (23)$$

3.6 输出层

将情感特征和方面特征融合, 并将其作为最终的语义表征。此处再次利用门控机制, 以控制情感特征和方面特征的权重。具体计算式如下:

$$\alpha = \sigma(\mathbf{W}_g \mathbf{H}_a + \mathbf{W}_l \mathbf{H}_l + \mathbf{b}) \quad (24)$$

$$\mathbf{s} = \alpha \mathbf{H}_a + (1 - \alpha) \mathbf{H}_l \quad (25)$$

其中, $\mathbf{W}_h, \mathbf{W}_l$ 为权重参数, $\mathbf{H}_a$ 为全局方面表征, $\mathbf{H}_l$ 为局部特征, σ 为 sigmoid 非线性激活函数。

获取最终语义表征 $\mathbf{s}$ 之后, 将其送入全连接层计算每一类情感极性的概率分布:

$$\mathbf{p} = \text{softmax}(\mathbf{W}_s \mathbf{s} + \mathbf{b}_s) \quad (26)$$

其中, $\mathbf{W}_s \in \mathbb{R}^{3 \times d_s}$ 为可学习参数, d_s 为 $\mathbf{s}$ 的维度。

损失函数是由训练中预测的和真实的标签分布的交叉熵定义的:

$$\text{loss} = - \sum_{i=1}^n y_i \ln \bar{y}_i + \lambda \|\theta\|^2 \quad (27)$$

其中, y_i 是真实标签, $\bar{y}_i$ 是预测概率, λ 为 L_2 正则化的耦合系数。在模型训练期间, 我们使用 adam 优化器来计算和更新参数。在模型的训练过程中, 为了避免过拟合, 我们还采用了 dropout 策略。

4 实验结果及分析

为验证算法的有效性, 本文提出的 ASGFN 模型与相关算法在公用数据集上进行了比较分析。

4.1 数据集

本文在 3 个广泛使用的数据集上进行实验, 验证了模型的有效性。所有的数据集都包括句子、方面以及对应的情感极性。3 个数据集分别为 REST14, LAPTOP, TWITTER,

其中 REST14 和 LAPTOP 来自 SemEval 2014 任务四¹⁾, 分别由餐厅评论和笔记本评论组成。TWITTER 数据集来自 ACL14^[14], 由推特评论组成的推特数据集。3 个数据集的基本信息如表 1 所列。

表 1 数据集的基本信息

Table 1 Basic information of datasets

Dataset		# Pos	# Neu	# Neg
TWITTER	Train	1 561	3 127	1 560
	Test	173	346	173
LAPTOP	Train	994	464	870
	Test	341	169	128
REST14	Train	2 164	637	807
	Test	728	196	196

4.2 参数设置

我们在实验中采用基于 Glove 词嵌入和基于预训练模型 Bert^[38]的两类模型。对于基于 Glove 词嵌入的模型, 我们使用 300 维的 Glove 预训练词嵌入初始化数据集中所有单词的向量表示。BiLSTM 网络的隐藏层维度设置为 50, 使用 Adam 优化器优化模型, 学习率设置为 0.002, 采用小批量梯度下降策略训练网络参数, 批量训练样本数设置为 16, 词性嵌入和位置嵌入维度设置为 30, L₂ 正则化的系数设置为 0.0001。为了防止训练过程中出现过拟合, 我们还引入了随机失活策略 dropout, 将其设置为 0.7。对于基于 BERT 的模型, 我们采用 “[CLS]句子[SEP]方面[SEP]” 作为输入, 将词嵌入的维数设置为 768, 批量训练样本数设置为 16, 学习率设置为 0.00002。为评估模型的性能, 采用精度和宏 F1 指标作为模型的评价标准。

4.3 对比方法

为了更好地评估所提模型的有效性, 将其与 13 个相关方法进行比较。由于模型中方面编码层引入了 GCN, 因此比较了部分基于 GCN 的模型。比较方法如下:

- 1)TD-LSTM^[5], Tang 等在 2016 年将 LSTM 网络引入方面级情感分析任务, 应用前向 LSTM 和反向 LSTM 分别对拼接了方面的左上下文和右上下文进行编码。
- 2)ATAE-LSTM^[20], 在 2016 年, Wang 等将注意力机制结合 LSTM, 对方面和上下文联合建模, 进而提取与方面相关的情感特征用于分类。
- 3)IAN^[26], 首先对上下文和方面的隐藏向量进行平均, 然后使用交互注意机制生成方面和上下文的表示, 将这两个表示联合起来进行分类。
- 4)MemNet^[29], 受记忆网络的启发, Tang 等基于多头注意机制和多跳机制构建了深度记忆网络, 通过多头注意机制计算上下文单词的权重, 从而计算文本表征。
- 5)RAO^[31], 该模型基于残差网络的思想设计残差注意机制, 对特定方面与上下文进行交互, 从而弥补交互过程中信息的损失。
- 6)GCAE^[33], 该模型基于 CNN, 集成门控机制, 分别利用两个卷积层提取语义特征。
- 7)RCNN-BGRU^[35], 该模型将区域卷积神经网络与融入

了高速公路网络的 RNN 相结合, 以此捕捉局部特征和时序特征。

8)MACRO-A-S^[36], 以双通道的方式结合 RCNN 和 ON-LSTM, 捕捉局部特征和时序特征, 然后引入注意机制融合两个通道生成的语义信息。

9)Tnet^[39], 该模型设计目标特定的转换组件, 将方面信息集成到单词表示中, 然后利用 CNN 提取语义特征。

10)SPRN^[40], 该模型设计了两个模块, 分别是基于 CNN 的局部语义特征提取模块和基于门控机制的方面和上下文交互模块。

11)DualGCN^[37], 该模型利用 GCN 对句法结构进行建模, 利用多头自注意机制捕获语义关系, 通过两者的互补来构建对偶图卷积网络模型。

12)ASGCN-DT^[41], 为了利用句法约束和长距离依存关系, Zhang 等在 2019 年将 GCN 引入方面级情感分析任务, 应用 GCN 对句法信息和词的依存关系进行建模。

13)R-GAT^[42], 将句法依赖树中依赖关系的标签信息作为特征引入模型, 从而构建关系图注意网络。

14)DGEDT^[43], 通过 Transformer 学习语义表示, 应用双向图卷积网络学习图表示, 引入双仿射变换对两个模块学习到的特征进行交互, 选择 Transformer 模块的输出作为最终的语义表征。

4.4 实验结果及分析

为了验证本文模型的性能, 将所提模型在 3 个公用数据集上的实验结果与基于注意力机制、CNN-RNN/CNN 以及图网络的基线模型进行比较。实验结果如表 2 所列, 表 2 中最好的结果加粗显示。

表 2 ASGFN 与基线模型的对比结果

Table 2 Comparison between ASGFN and baselines

Models		Rest14		Laptop		Twitter	
		Acc	F1	Acc	F1	Acc	F1
Attention	TD-LSTM	77.32	65.07	68.34	61.43	67.33	64.84
	ATAE-LSTM	77.20	67.02	68.70	63.93	68.64	66.60
	IAN	79.26	70.09	72.05	67.38	71.82	69.11
	MemNet	79.89	69.05	74.37	68.63	74.74	72.81
	RAO	80.98	71.34	74.45	69.72	73.12	71.20
RNN-CNN/CNN	GCAE	77.28	70.77	69.14	64.36	74.89	73.55
	RCNN-BGRU		77.54		71.97		
	MACRO-A-S	81.69	72.66	70.92	68.14	73.96	73.36
	Tnet	80.79	70.84	76.54	71.75	74.97	73.60
	SPRN	80.77	71.28	73.20	68.96	72.10	69.68
GCN	DualGCN	84.27	78.08	78.48	74.74	75.92	74.29
	ASGCN-DT	80.86	72.19	74.14	69.24	71.53	69.68
	R-GAT	83.55	75.99	78.02	74.00	75.36	74.15
	DGEDT	83.90	75.10	76.80	72.30	74.80	73.40
ours	ASGFN	84.72	78.55	78.64	75.67	76.66	75.46
BERT	Bert	84.11	76.68	77.58	73.28	75.58	73.45
	SPRN(BERT)	85.03	76.97	79.31	76.61	75.70	73.50
	DualGCN+BERT	87.13	81.16	81.80	78.10	77.40	76.02
	R-GAT+BERT	86.60	81.35	78.21	74.07	76.15	74.88
	DGEDT-BERT	86.30	80.00	79.80	75.60	77.90	75.40
	ASGFN(BERT)	86.97	80.78	82.11	78.15	77.85	76.48

由表 2 可见, 基于 Glove 词嵌入的模型, 在不同语法结构

¹⁾ <https://alt.qcri.org/semeval2014/task4>

的数据集上,ASGFN 的表现优于大部分竞争对手,证明了所提模型的有效性。在所有基线模型中,TD-LSTM 表现最差,主要原因是该模型对方面项的处理过于简单,并且忽略了方面与上下文的关系。引入注意力机制之后,增强了方面与上下文的交互。基于 RNN 与注意力机制的模型的性能有了一定的提升,与 ATAE-LSTM,IAN,Memnet,RAO 相比,ASGFN 在 3 个数据集上的准确率平均提升了 6.70%,宏 F1 值平均提升了 9.40%。与基于 RNN-CNN 或者 CNN 的基线模型 GCAE,RCNN-BGRU,MACRO-A-S,Tnet,SPRN 相比,本文模型也取得了较好的表现,在 3 个数据集上准确率平均提升了 4.48%,宏 F1 值平均提升了 5.32%。相比基于 GCN 的模型(DualGCN,ASGCN,RGAT,DGEDT),本文模型在 3 个数据集上的准确率平均提升了 1.87%,宏 F1 值提升了 2.96%。相比 DualGCN,本文模型在 3 个数据集上的性能都有提升,主要是因为我们考虑了具有语义的方面与上下文的交互,以及局部特征的提取。特别是在 Twitter 数据集上,本文模型的性能提升较多,一个可能的原因是, Twitter 数据集的语法规范性相对较差,引入句法信息可能会引入更多的噪音,而我们选择门控过滤网络结合 CNN 提取情感特征,不仅过滤了噪音,还抽取了重要的局部特征。

对于基于 BERT 的模型,最基本的 BERT 就已经比大部分现有的 ABSA 模型表现得更好,证明了预训练模型在这项任务中的有效性。将 BERT 引入本文模型后,相比其他基于 BERT 的比较著名的模型,在 LAPTOP 和 TWITTER 数据集上本文模型具有较强的竞争力。相比基于 Glove 词嵌入的模型,ASGCN(BERT)模型的性能也有显著的提升,这也验证了本文模型与预训练模型结合的有效性。

从所有数据集的比较结果可以看出,本文模型获得了相对显著的性能优势,主要原因有 3 点:1)将上下文单词的词性信息与位置信息引入模型,丰富文本特征;2)根据特定语境下的方面语义抽取情感特征,判断情感极性,提高了需要根据方面语义判断情感极性的样本的分类准确性;3)利用门控过滤网络将 RNN 和 CNN 融合进本文模型,这不但利用了 RNN 和 CNN 各自的优势,还考虑了方面信息,增强了方面与上下文的交互。

4.5 消融实验

为了验证模型的各个模块的有效性,本节将在 3 个公用数据集上比较 ASGFN 及其 5 个变体的性能。5 个变体分别表示为 ASGFN w/o pos,ASGFN w/o post,ASGFN w/o aspect semantic,ASGFN w/o gate 和 ASGFN w/o CNN。其中,“w/o”代表将对应模块移除模型,pos 代表词性编码模块,post 代表位置编码模块,aspect semantic 代表方面编码模块,gate 代表门控过滤模块,CNN 代表卷积层。具体实验结果如表 3 所列。由表 3 可知,未引入词性信息或位置信息的模型在所有数据集上的表现都不令人满意,这表明词性特征和位置特征有助于模型更多地关注到与方面相关的情感词。移除方面编码模块后,模型的性能急剧下降。基于此,可以得出结论,结合上下文构建具有真实语义的方面表示是重要的,其有助于模型根据不同语境判断不同方面的情感极性。与 ASGFN w/o gate 模型相比,ASGFN 的性能有一定程度的提升,

这表明本文设计的过滤门对噪音的过滤起到了一定的作用,提升了模型的性能。我们还注意到,移除卷积层之后,模型的性能下降相对较多,这是因为模型从一定程度上丧失了对 n 元语义特征的提取能力,导致模型的性能下降。

表 3 消融实验
Table 3 Ablation study

Models	Rest14		Laptop		Twitter	
	Acc	F1	Acc	F1	Acc	F1
ASGFN w/o pos	82.57	74.58	76.90	73.62	74.74	73.00
ASGFN w/o post	83.65	76.68	77.22	73.61	75.78	74.58
ASGFN w/o aspect semantic	82.31	73.23	75.63	71.10	74.30	72.61
ASGFN w/o gate	83.38	75.71	77.22	72.69	75.92	74.68
ASGFN w/o CNN	82.66	74.77	76.11	72.43	75.04	73.49
ASGFN	84.72	78.55	78.64	75.67	76.66	75.46

4.6 GCN 层数分析

为了探究 GCN 层数对实验结果的影响,将 GCN 的层数分别设置为 1~12,分析不同层数下精度和宏 F1 值的变化。实验结果分别如图 4 和图 5 所示。从图中可以看出,当 GCN 的层数设置为 2 时,相比其他层数设置,模型取得了最佳的效果。当 GCN 层数设置为 1 时,在 3 个数据集上精度和宏 F1 值的表现都不理想,这说明一层 GCN 无法很好地捕获特定方面的上下文依赖,无法获取更具代表性的方面语义。另外,随着 GCN 层数的增加,模型的深度也会加深,当 GCN 层数超过 2 时,模型性能随之下降。其主要原因可能是随着模型深度的增加,模型的参数量增大可能会导致模型难以学习,从而影响模型性能。

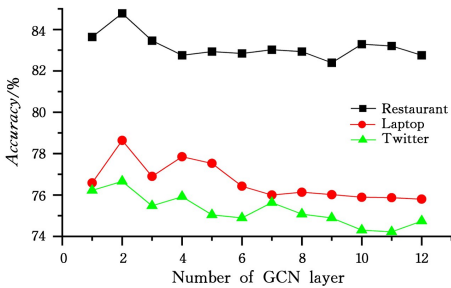


图 4 GCN 层数对模型精度的影响
Fig. 4 Impact of GCN layers on accuracy

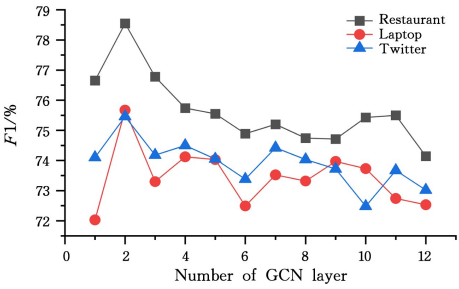


图 5 GCN 层数对模型宏 F1 值的影响
Fig. 5 Impact of GCN layers on F1

4.7 方面语义探究

为了探究方面语义对实验结果的影响,在该部分通过 3 种方法构建方面表示。第一种方法是本文模型中所采取的方法,通过多头自注意机制构建权重矩阵,然后利用 GCN

聚合上下文信息。第二种方法是利用句法依赖树构建邻接矩阵,然后利用 GCN 聚合邻接节点的信息,提取方面表征,并将其表示为 Aspect with syntax。第三种方法为直接利用方面词的词嵌入作为方面表示,将其表示为 Aspect with embedding。对比实验结果如表 4 所列。

表 4 方面表示及其变体实验结果

Models	Table 4 Experimental results of aspect representation and its variants					
	Rest14		Laptop		Twitter	
	Acc	F1	Acc	F1	Acc	F1
Aspect with syntax	82.04	73.46	76.42	72.71	75.48	74.27
Aspect with embedding	82.22	74.30	76.74	73.58	74.89	73.65
ASGFN	84.72	78.55	78.64	75.67	76.66	75.46

表 5 ASGFN w/o gate,Aspect with embedding,ASGFN 预测样本结果

Table 5 Predicted sample results of ASGFN w/o gate,Aspect with embedding and ASGFN					
Sentence	aspect	label	ASGFN w/o gate	Aspect with embedding	ASGFN
The battery life is amazingly long at 7 h and 5 h if you use it.	(battery life)	(P)	(P)	(P)	(P)
The reality was,it heated up very quickly,and took way too long to do simple things,likeopening my Documents folder	(opening my Documents folder)	(N)	(N)	(P)	(N)
Well,it happened because of a graceless manager and a rude bartender who had us waiting 20 min for drinks, and then tells us to chill out.	(manager,bartender, waiting,drinks)	(N,N,N,O)	(N,N,N,N)	(N,N,N,O)	(N,N,N,O)

第一个案例中,对于方面“battery life”,3 个模型都能正确预测,主要原因是,在笔记本电脑领域,“long batery life”这样的描述是普遍的。经过训练,模型能够很容易捕捉到情感词“long”,并得到一个“long”对应积极情感的模式。因此,当训练好的模型遇到包含“long”的样本时,大概率会将该样本分类为积极样本,从而导致 Aspect with embedding 模型将案例二分类为积极样本。Aspect with embedding 简单地使用方面词的固定词嵌入表示方面,导致其缺乏真实语义,从而无法正确判断当前语境下方面的情感极性。而 ASGFN 能够正确判断案例二的情感极性,主要是因为我们考虑特定语境下的方面语义。在案例三中,ASGFN w/o gate 未能正确预测方面“drinks”的情感极性,主要是因为缺乏方面与上下文的交互机制,模型无法探索方面和上下文的联系,从而导致其他方面的情感词影响其判断。

结束语 本文提出了一种基于方面语义和门控过滤网络的网络模型,用于方面级情感分析任务。首先将词嵌入、位置嵌入以及词性嵌入融合,通过 BiLSTM 对上下文进行编码;然后通过多头自注意机制生成权重矩阵,结合 GCN 对长距离依赖进行建模,从而提取更深层次的方面语义,生成方面表示。为了在利用 RNN 和 CNN 的优势的同时考虑方面信息,我们在 BiLSTM 的基础上设计门控过滤网络,用于探索具有语义的方面与上下文的联系,从而过滤无关情感信息,然后通过不同大小的卷积核提取多粒度的情感特征。最终将情感特征与方面语义表示相结合作为最终的语义表征。在 3 个公用数据集上的实验结果表明,本文模型在方面级情感分析任务上取得了优异的性能。由于方面级情感分析数据集的大小的限制,模型无法更好地发挥性能。在未来,我们将探讨如何从

从表 4 可以看出,无论是基于句法依赖构建方面表示还是直接使用方面词的词嵌入代替方面表示,模型的性能都有所下降,主要原因可能是因为句法解析带来的误差以及简单的词嵌入无法表达方面的深层语义。结合上下文信息能够获取更深层次的方面语义,从而使模型能够生成特定语境下的情感特征,减少了方面无关的情感特征带来的误差。

4.8 案例分析

为了更好地分析本文模型的性能,我们从数据集中选取一部分典型案例进行实例分析。我们选取了 3 个案例,包括它们的方面集合和对应的真实标签,P,N,O 分别表示积极、消极以及中性 3 种情感极性。实验结果如表 5 所列。

粗粒度数据集或者情感知识库中学习更多的情感表达模式,以提升模型的性能。

参 考 文 献

[1] SCHOUTEN K,FRASINCAR F. Survey on aspect-level sentiment analysis [J]. IEEE Transactions on Knowledge and Data Engineering,2015,28(3):813-830.

[2] MA D,LI S,WU F,et al. Exploring sequence-to-sequence learning in aspect term extraction[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019:3538-3547.

[3] HU M,ZHAO S,GUO H,et al. Multi-Label Few-Shot Learning for Aspect Category Detection[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing(Volume 1:Long Papers). 2021:6330-6340.

[4] WU M,WANG W,PAN S J. Deep weighted MaxSAT for aspect-based opinion extraction[C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing(EMNLP). 2020:5618-5628.

[5] TANG D,QIN B,FENG X,et al. Effective lstms for target-dependent sentiment classification[C]// Proceedings of COLING 2016,the 26th International Conference on Computational Linguistics:Technical Papers. 2016:3298-3307.

[6] ZHANG W,DENG Y,LI X,et al. Aspect Sentiment Quad Prediction as Paraphrase Generation[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. 2021:9209-9219.

[7] CAI H,XIA R,YU J. Aspect-category-opinion-sentiment quad-

- ruple extraction with implicit aspects and opinions[C] // Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), 2021:340-350.
- [8] DENG S, SINHA A P, ZHAO H. Adapting sentiment lexicons to domain-specific social media texts [J]. *Decision Support Systems*, 2017, 94: 65-76.
- [9] PORIA S, CAMBRIA E, WINTERSTEIN G, et al. Sentic patterns: dependency-based rules for concept-level sentiment analysis [J]. *Knowledge-Based Systems*, 2014, 69: 45-63.
- [10] 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 (SemEval 2014). 2014: 437-442.
- [11] JIANG L, YU M, ZHOU M, et al. Target-Dependent twitter sentiment classification [C] // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 2011: 151-160.
- [12] ELMAN J L. FINDING structure in time [J]. *Cognitive Science*, 1990, 14(2): 179-211.
- [13] KALCHBRENNER N, GREFFENSTETTE E, BLUNSOM P. A convolutional neural network for modelling sentences [C] // Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 655-665.
- [14] 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 (Volume 2: Short papers). 2014: 49-54.
- [15] NGUYEN T H, SHIRAI K. Phrasernn: phrase recursive neural network for aspect-based sentiment analysis [C] // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 2509-2514.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [17] BASIRI M E, NEMATIS S, ABDAR M, et al. ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis [J]. *Future Generation Computer Systems*, 2021, 115: 279-294.
- [18] LI W, ZHU L, SHI Y, et al. User reviews: Sentiment analysis using lexicon integrated two-channel CNN-LSTM family models [J]. *Applied Soft Computing*, 2020, 94: 106435.
- [19] HU Y L, TONG T Q, ZHANG X Y, et al. Self-attention-based BGRU and CNN for sentiment analysis [J]. *Computer Science*, 2022, 49(1): 252-258.
- [20] WANG Y, HUAN G M, ZHAO L, et al. Attention-based lstm for aspect-level sentiment classification [C] // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 606-615.
- [21] 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 (EMNLP). 2014: 1532-1543.
- [22] LIANG B, LUO W, LI X, et al. Enhancing aspect-based sentiment analysis with supervised contrastive learning [C] // Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021: 3242-3247.
- [23] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks [C] // 5th International Conference on Learning Representations. 2016.
- [24] BAHDANAU D, CHO K H, BENGIO Y. Neural machine translation by jointly learning to align and translate [C] // 3rd International Conference on Learning Representations. 2015.
- [25] KUMAR A, IRSOY O, ONDRUSKA P, et al. Ask me anything: Dynamic memory networks for natural language processing [C] // International Conference on Machine Learning. PMLR, 2016: 1378-1387.
- [26] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification [C] // Proceedings of the 26th International Joint Conference on Artificial Intelligence. 2017: 4068-4074.
- [27] YANG C, ZHANG H, JIANG B, et al. Aspect-based sentiment analysis with alternating coattention networks [J]. *Information Processing & Management*, 2019, 56(3): 463-478.
- [28] JIANG N, TIAN F, LI J, et al. MAN: mutual attention neural networks model for aspect-level sentiment classification in SIoT [J]. *IEEE Internet of Things Journal*, 2020, 7(4): 2901-2913.
- [29] TANG D, QIN B, LIU T. Aspect level sentiment classification with deep memory Network [C] // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 214-224.
- [30] LV Y, WEI F, CAO L, et al. Aspect-level sentiment analysis using context and aspect memory network [J]. *Neurocomputing*, 2021, 428: 195-205.
- [31] WU C, XIONG Q, YANG Z, et al. Residual attention and other aspects module for aspect-based sentiment analysis [J]. *Neurocomputing*, 2021, 435: 42-52.
- [32] XU Q, ZHU L, DAI T, et al. Aspect-based sentiment classification with multi-attention network [J]. *Neurocomputing*, 2020, 388: 135-143.
- [33] XUE W, LI T. Aspect based sentiment analysis with gated convolutional networks [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018: 2514-2523.
- [34] PIAO G, BRESLIN J G. FINANCIAL aspect and sentiment predictions with deep neural networks: an ensemble approach [C] // Companion Proceedings of the The Web Conference 2018. 2018: 1973-1977.
- [35] SUN Z F, WANG J. RCNN-BGRU-HN network model for aspect-based sentiment analysis [J]. *Computer Science*, 2019, 46(9): 223-228.
- [36] TAO Y C, WU W L, WU L, et al. Aspect-based Sentiment

Analysis of Multi-channel Based on Attention Mechanism and Sentiment Information [J]. Journal of Chinese Computer Systems, 2022, 43(7):1406-1412.

[37] 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 (Volume1: Long Papers). 2021:6319-6329.

[38] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pretraining of deep bidirectional transformers for language understanding [J]. arXiv:1810.04805, 2018.

[39] LI X, BING L, LAM W, et al. Transformation networks for target-oriented sentiment classification [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018:946-956.

[40] SONG W, WEN Z, XIAO Z, et al. Semantics perception and refinement network for aspect-based sentiment analysis [J]. Knowledge-Based Systems, 2021, 214:106755.

[41] 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 (EMNLP-IJCNLP). 2019:4568-4578.

[42] 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. 2020:3229-3238.

[43] TANG H, JI D, LI C, et al. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification [C] // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020:6578-6588.



HE Zhihao, born in 1998, postgraduate. His main research interests include intelligent information processing and sentiment analysis.



CHEN Hongmei, born in 1971, Ph.D., professor, Ph.D. supervisor, is a member of China Computer Federation. Her main research interests include intelligent information processing, pattern recognition, etc.

(责任编辑:喻黎)