

# 基于层次注意力机制和门机制的属性级别情感分析

冯超  
云南广电网络集团有限公司  
471573289@qq.com  
薛云  
华南师范大学  
xueyun@scnu.edu.cn

黎海辉  
深圳职业技术学院  
lihaihuisncnu@m.scnu.edu.cn

赵洪雅(✉)  
深圳职业技术学院  
hy.zhao@szpt.edu.cn  
唐婧尧  
华南师范大学  
manderous@foxmail.com

## 摘要

近年来, 作为细粒度的属性级别情感分析在商业界和学术界受到越来越多的关注, 其目的在于识别一个句子中多个属性词所对应的情感极性。目前, 在解决属性级别情感分析问题的绝大多数工作都集中在注意力机制的设计上, 以此突出上下文和属性词中不同词对于属性级别情感分析的贡献, 同时使上下文和属性词之间相互关联。本文提出使用层次注意力机制和门机制处理属性级别情感分析任务, 在得到属性词的隐藏状态之后, 通过注意力机制得到属性词新的表示, 然后利用属性词新的表示和注意力机制进一步得到上下文新的表示, 层次注意力机制的设计使得上下文和属性词的表达更加准确; 同时通过门机制选择对属性词而言上下文中有用的信息, 以此丰富上下文的表达, 在SemEval 2014 Task4和Twitter数据集上的实验结果表明本文提出模型的有效性。

**关键词:** 属性级别; 情感分析; 注意力机制; 门机制

## Aspect-level Sentiment Analysis Based on Hierarchical Attention and Gate Networks

FENG Chao  
Yunnan Radio Broadcasting  
& Television Networking  
Group CO., Ltd.  
471573289@qq.com

LI Haihui  
Shenzhen Polytechnic  
lihaihuisncnu@m.scnu.edu.cn

ZHAO Hongya(✉)  
Shenzhen Polytechnic  
hy.zhao@szpt.edu.cn

XUE Yun  
South China Normal University  
xueyun@scnu.edu.cn

TANG Jingyao  
South China Normal University  
manderous@foxmail.com

## Abstract

As a fine-grained task, aspect-level sentiment analysis whose purpose is to identify the sentiment polarity corresponding to the specific aspect in a sentence has received more and more attention in the business and academic. The most of the related works focus on the design of attention network to highlight the different contributions of words in context and make context and aspect interacted. In this paper, we put forward the hierarchical attention and gate networks to process aspect-level sentiment analysis task. we obtain the new representation of the context through the attention network between the context and the new representation of the aspect which weight by the context. At the same time, the useful information in the context is selected through the gate networks to enrich the representation of the context. The design of hierarchical attention gate networks make the representation of the context and the aspect more accurate. The experimental results on the SemEval 2014 Task4 and Twitter show the validity of the model.

**Keywords:** Aspect-level, sentiment analysis, attention networks, gate networks

## 1 引言

近年来,随着互联网的发展,大量的消费平台和社交网络平台逐渐走进人们的生活(Li et al., 2019b)。在消费平台上,消费者消费之后往往会留下一段关于产品优劣的评论;在社交网络平台上,网民也会留下关于某个事件的看法。这些含有情感信息的评论无论是对企业或者是政府而言都是极具价值的,企业可以通过分析消费者的评论,了解产品的不足,以达到改善产品性能的目的;政府也可以通过总结网民对事件的看法以引导事件的发展方向。因此,如何从富含情感信息的评论中挖掘出有用的信息成为自然语言处理的热点研究方向之一。

情感分析是自然语言处理中重要的子领域,可分为句子级别情感分析、篇章级别情感分析和属性级别情感分析(Shaikh and Deshpande, 2016)。句子级别情感分析为对整个句子的单一情感极性(积极、中性和消极)进行判断。篇章一般由多个句子构成,所以篇章级别情感分析在于根据其构成的多个句子判断篇章的整体情感极性(欧阳志凡, 2018)。由此可知,句子级别和篇章级别情感分析均为对评论文本的单一情感极性进行判断,而在产品以及产品属性日益多样化的今天,消费者一般会留下一段关于产品多个属性评论的评论文本,如:“*The food was extremely tasty, but the service was dreadful.*”,这是一段来自酒店的评论,通过分析可以看出属性词“*food*”和“*service*”的情感极性分别是积极和消极,而得到整体情感极性的句子和篇章级别情感分析明显不适合这种含有多个属性词且对每个属性词表达情感不一致的情况,所以,属性级别情感分析逐渐受到学术界和商业界的关注。

属性级别情感分析需要找出特定属性词在上下文(本文将评论中除属性词部分称作为属性词的上下文)中所关联部分以达到判断特定属性词情感倾向的目的(Li et al., 2019a)。因此,属性级别情感分析需要充分利用属性词和上下文的关系,即不同属性词在上下文中关注的词也有所不同(Ma et al., 2017),如:当识别属性词“*picture quality*”情感极性的时候,“*clear-cut*”自然能与属性词“*picture quality*”产生联系。除此之外,在属性级别情感分析中,不同词对于不同属性词情感极性判断的作用也是不一致的,如:情感词。因此,在近些年属性级别情感分析工作中,如何通过上下文和属性词的交互式学习,从而达到识别出上下文中对属性词重要词的目的已经成为了研究重点。常见的交互方式有如下几种:第一,将属性词向量与上下文向量进行拼接;第二,通过设计上下文和属性词之间的交互注意力机制。但是简单的拼接并不能达到识别出上下文中对属性词重要词的效果,因此,通过设计上下文和属性词之间的注意力机制成为了学术界研究热点方向之一。但是单层的注意力机制难以达到将上下文和属性词相互关联的目的。而且,在大多数属性级别情感分析工作中将上下文与属性词之间的注意力机制结果直接作为上下文和属性词的表达,本文认为该做法在一定程度上忽略了原上下文和属性词表示的作用。因此本文提出使用层次注意力机制和门机制(Aspect-level Sentiment Analysis Based on Hierarchical Attention and Gate Networks, HAG)来处理属性级别情感分析任务,贡献如下:

首先,本文认为单层交互注意力机制并不能很好的将上下文和属性词相互关联,而且难以识别出在上下文中对于属性词比较重要的词。如:在对“*The food is average and dreadful serving speed*”中的“*The food*”进行情感分析的时候,单层交互注意力机制会同时加大“*average*”和“*dreadful*”的权重,而“*dreadful*”的情感强度通常更大,“*dreadful*”的权重一般会大于“*average*”,而多层交互注意力机制将进一步对词语进行加权,调节“*average*”和“*dreadful*”的权重。本文使用层次注意力机制先获取上下文对属性词的影响矩阵,从而更新属性词的表示,进一步的得到上下文的表示。此方式利用上下文与属性词之间的注意力机制使属性词的表达更加准确,再将利用新的属性词表示与上下文的注意力机制,使的上下文的表达更加准确,通过层次注意力机制的设计将上下文和属性词之间进行了充分的关联,同时对不同词进行了加权以达到判断属性级别情感分类的目的。

其次,在以往属性级别情感分析的研究工作中,通常将属性词对上下文的影响矩阵当作是上下文的最终表示,将上下文对属性词的影响矩阵之间当作是属性词的表示,该做法忽略了上下文和属性词原表示的作用。本文认为影响矩阵只能作为对原表示调整的依据而不能直接作为上下文和属性词的直接表示,因此,本文将二者相加,一定程度上保留了原始信息。

最后,本文认为门机制的选择作用能够选择出对于不同属性词上下文中重要的词,因此本文通过门机制改变上下文的表达,丰富上下文的表示。实验结果表明门机制的选择作用在一定程度上提高了模型的准确率。

## 2 相关工作

从Minqing Hu (Hu and Liu, 2004)提出需要关注产品不同属性的情感表达起,便出现了大量属性级别情感分析工作,这些工作大致可以分为基于传统机器学习的属性级别情感分析方法、基于深度学习的情感分析方法、深度学习与注意力相结合的方法和其他混合的方法。本文提出的基于层次注意力机制和门机制的属性级别情感分析属于近几年成为研究热点的深度学习和注意力机制相结合的方法,本章将从注意力机制和门机制两部分具体介绍属性级别情感分析的相关工作。

### 2.1 注意力机制

近些年来,神经网络的发展加快了属性级别情感分析的研究,CNN、RNN、LSTM、GRU等经典的神经网络更是在提取文本隐藏特征有着显著的作用;另一方面,在自然语言处理中,注意力机制能够选择出有助于特定自然语言处理任务的词,在减小计算量同时,可以大大减少噪声的引入。因此,在近些年的自然语言处理任务中,将神经网络与注意力机制相结合成为了主流方法。在属性级别情感分析上,注意力机制的使用一定程度上可以选择出有助不同属性情感分类的词,从而提高了实验准确率,因此,深度学习与注意力机制相结合的方法也成为了属性级别情感分析的主要研究方法。Wang (Wang et al., 2016)在处理属性识别情感分析任务时提出了两个模型。第一个模型通过平均池化将属性词矩阵表示转化成属性词词向量,然后通过属性词词向量与上下文隐藏表示之间的注意力机制对上下文表示进行加权。相对于第一个模型,第二个模型在上下文进入LSTM之前先将属性词向量与上下文矩阵表示进行拼接。由于使用了注意力机制,两个模型都取得了较高的实验结果,后者通过拼接进一步加大了上下文和属性词之间的关联性,因此实验效果也有一定的提升。自此,在处理属性级别情感分析任务的大多数工作中都引入了注意力机制。Ma (Ma et al., 2017)将上下文和属性词之间的关联定义为交互并认为不仅属性词对上下文具有选择作用,而且上下文对属性词也有选择作用,因此,提出使用二者之间的交互注意力机制来处理属性级别情感分析任务。Huang (Huang et al., 2018)提出使用Attention-over-Attention来增强上下文和属性词之间的交互程度。Zhang (Zhang et al., 2019)提出运用协同注意力机制处理属性级别情感分析任务,具体为:将评论句子分成属性词上文、属性词和属性词下文,然后将三部分两两进行协同注意力机制。Huang (Huang and Carley, 2019)通过语法感知和图注意力机制得到属性级别情感极性分布。以上工作都是通过单层注意力机制加深上下文和属性词之间的交互。Li (Li et al., 2019a)先通过协同注意机制在上下文和属性词的表示中分别融入彼此之间的影响,然后通过自注意机制进一步对自身表示进行加权。Li (Li et al., 2018)将上下文向量与位置向量进行拼接,经过GRU之后再与通过注意力机制得到的属性词向量进行拼接,最后再通过注意力机制得到含有属性词情感信息的最终表示。该工作通过多层注意机制使得上下文的表示能够更加准确的表达出属性词对应的情感信息。Song (Song et al., 2019)分别将上下文词嵌入与属性词词嵌入输入到多头注意力机制,分别得到上下文和属性词的表示,然后通过二者之间的注意力机制得到二者之间的交互向量,通过分类器进一步得到属性词对应情感极性分布。Zheng (Zheng and Xia, 2018)先将属性词隐藏表示与属性词上下文和属性词下文进行注意力机制得到属性词上文向量和属性词下文向量,然后将属性词上文向量和属性词下文向量分别与属性词隐藏表示进行注意力机制。该工作先通过注意力机制更新了属性词上文和属性词下文的表示,使得属性词上文和属性词下文的表示更加准确,然后进一步的获取属性词的表示,层次注意力机制的交互和互相选择作用使实验效果得到进一步的提升。

### 2.2 门机制

属性级别情感分析在于从上下文中选择出针对当前属性词情感分析有用的词从而改变上下文表示,以达到判断属性词情感极性分布的效果,而门机制在选择和提取信息有显著的效果。因此,也存在大量使用门机制处理属性级别情感分析任务。Xue (Xue and Li, 2018)先使用CNN提取属性词和上下文的信息,然后通过门机制(gating mechanisms)选择上下文和属性词中有用的信息。Liang (Liang et al., 2019)处理属性级别情感分析任务时设计了两种不同的门机制,使上下文的表示融入了属性词信息,同时提取和选择出了有用的上下文的信息,该方法在两个数据集上也取得了较好的实验效果。因此,本文使用了门机制选择出上下文中有用的信息,以此丰富了上下文的表达,实验结果证明门机制的设计在一定程度上丰富了上下文的表

达，对实验准确率的提高有一定的作用，进一步提高了模型的有效性。

### 3 模型

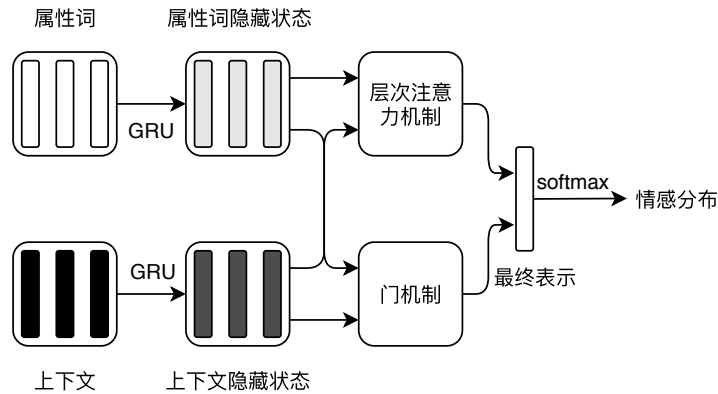


图 1: 总体模型结构

在属性级别情感分析中，常常给定一段对属性进行评论的评论文本，因此需要依次对评论文本中的属性进行情感极性的判断。本文将需要判断情感极性的属性称为属性词，而评论中除属性词之外的词称为上下文。给定一段长度为  $n + m$  的评论文本，其中上下文  $[w_c^1, w_c^2, \dots, w_c^n]$ 、属性词  $[w_t^1, w_t^2, \dots, w_t^m]$ ，根据 GloVe 得到对应上下文和属性词的词嵌入矩阵分别是  $C = [v_c^1, v_c^2, \dots, v_c^n] \in \mathbf{R}^{n \times d_w}$  和  $T = [v_t^1, v_t^2, \dots, v_t^m] \in \mathbf{R}^{m \times d_w}$ ，其中  $d_w$  是词向量维度的大小。为了进一步提取出文本的隐藏特征，本文使用了 GRU 提取出上下文和属性词隐藏信息。GRU 不仅可以缓解传统 RNN 在反向传播时出现的梯度消失和梯度爆炸的问题，而且在参数量上较少，因此 GRU 更适合短文本的属性级别情感分析任务。在通过 GRU 提取出上下文隐藏状态  $H_C = [h_c^1, h_c^2, \dots, h_c^n] \in \mathbf{R}^{n \times d_h}$  和属性词隐藏状态  $H_T = [h_t^1, h_t^2, \dots, h_t^m] \in \mathbf{R}^{m \times d_h}$  之后，为进一步在将上下文和属性词进行关联得到上下文和属性词更精准的表达，本文使用层次注意力机制和门机制在上下文和属性词的表示中融入了彼此影响之间的关系，拼接之后再通过分类器得到属性词在上下文中的情感极性分布，具体模型结构如图 1 所示。

#### 3.1 层次注意力机制

层次注意力机制为本文的两个核心之一，其目的在于在通过层次注意力机制的加权，使上下文和属性词相关关联，表达更加准确。本文先通过上下文对属性词进行注意力机制，得到上下文对属性词的影响矩阵，从而更新属性词的表示；然后通过属性词新的表示，进一步对上下文进行注意力机制，得到属性词对上下文的影响矩阵，进一步更新上下文的表示。注意力机制的层次设计，先获取了属性词的表示，再获取上下文的表示，使属性词和上下文获取了更加准确的表示，同时加深了上下文和属性词之间的关联性；通过对原表示与影响矩阵进行相加然后再加权的结构设计，在一定程度保留原有信息，丰富了新表示的信息，通过多次加权结构，进一步的体现了上下文和属性词中不同词对于属性级别情感分析不同作用性的思想，具体结构如图 2 所示。

在通过 GRU 得到属性词隐藏状态  $H_T$  和上下文隐藏状态  $H_C$  之后，为了在属性词与上下文之间进行注意力机制，首先需要得到属性词与上下文之间的注意力权重矩阵，如公式 1 所示：

$$H = \text{relu}(H_T \cdot W \cdot H_C^T) = \begin{bmatrix} h_{1,1} & \cdots & h_{1,n} \\ \vdots & \ddots & \vdots \\ h_{m,1} & \cdots & h_{m,n} \end{bmatrix} \quad (1)$$

其中， $\text{relu}$  为线型激活函数， $W \in \mathbf{R}^{d_h \times d_h}$  为参数矩阵， $H_C^T$  为  $H_C$  的转置矩阵， $H \in \mathbf{R}^{m \times n}$  为注意力机制权重矩阵，为属性词中每个词与上下文中每个词之间的关联系数。

首先利用上下文对属性词进行加权，进而得到新的属性词表示如公式 2 所示：

$$H_T^a = \text{relu}[(H_T + H H_C) W_T + b_T] \quad (2)$$

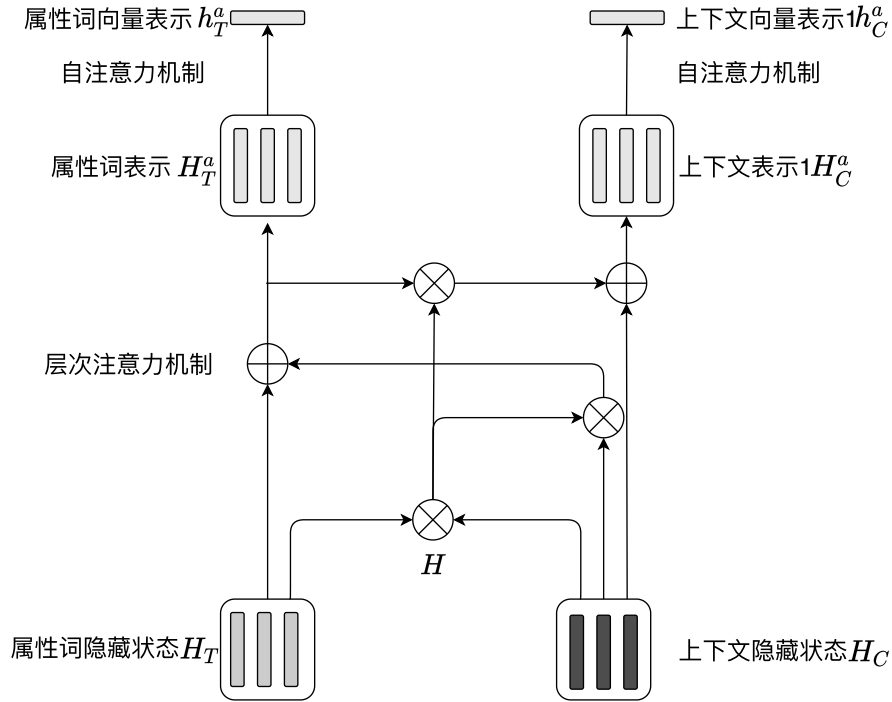


图 2: 层次注意力机制结构图

其中,  $HH_C$ 表示利用注意力机制权重与上下文相乘, 得到上下文对属性词的影响矩阵, 然后进一步的与属性词的隐藏状态进行相加, 再通过参数矩阵  $W_T \in \mathbf{R}^{d_h \times d_h}$  进行调整,  $b_T \in \mathbf{R}^{m \times d_h}$  为偏置矩阵, 得到基于层次注意力属性词表示  $H_T^a \in \mathbf{R}^{m \times d_h}$ 。进一步的根据属性词表示  $H_T^a \in \mathbf{R}^{m \times d_h}$ , 得到基于层次注意力上下文表示  $H_C^a \in \mathbf{R}^{n \times d_h}$ , 如公式3所示:

$$H_C^a = \text{relu}[(H_C + H^T H_T^a)W_C + b_C] \quad (3)$$

其中,  $H^T \in \mathbf{R}^{n \times m}$  为  $H$  的转置, 表示上下文中每个词与属性词中每个词之间的关联系数。 $H^T H_T^a$  表示利用注意力机制权重与属性词新的表示相乘, 得到属性词新表示对上下文的影响矩阵, 然后进一步的与上下文的隐藏状态进行相加, 再通过参数矩阵  $W_C \in \mathbf{R}^{d_h \times d_h}$  进行调整,  $b_C \in \mathbf{R}^{n \times d_h}$  为偏置矩阵, 得到基于层次注意力属性词表示  $H_C^a$ 。为进一步的突出属性词和上下文中不同词的重要性, 本文通过自注意力机制得到属性词向量表示和上下文向量表示1如公式4、5、6、7所示:

$$\alpha = \text{softmax}(H_T^a w_T) \quad (4)$$

$$\beta = \text{softmax}(H_C^a w_C) \quad (5)$$

$$h_T^a = \sum_{i=1}^m \alpha^i H_T^{a,i} \quad (6)$$

$$h_C^a = \sum_{j=1}^n \beta^j H_C^{a,j} \quad (7)$$

其中  $w_C \in \mathbf{R}^{d_h}$ 、 $w_T \in \mathbf{R}^{d_h}$  为参数向量,  $\alpha$ 、 $\beta$  分别是属性词和上下文自注意力机制系数,  $h_T^a \in \mathbf{R}^{d_h}$ 、 $h_C^a \in \mathbf{R}^{d_h}$  分别是属性词向量表示和上下文向量表示,  $\alpha^i$ 、 $\beta^j$  分别为  $\alpha$ 、 $\beta$  中的第  $i$ 、 $j$  个数,  $H_T^{a,i}$ 、 $H_C^{a,j}$  分别表示  $H_T^a$  和  $H_C^a$  中第  $i$ 、 $j$  个向量。

### 3.2 门机制

为进一步丰富上下文的表示, 本节通过使用门机制选择出针对属性词而言, 选择出上下文中重要词信息, 并把这一部分当作是属性词对上下文表示的影响矩阵。通过影响矩阵与上下文原表示相加, 并在参数矩阵调节下得到上下文表示  $2H_C^a$ , 具体结构如图3所示。

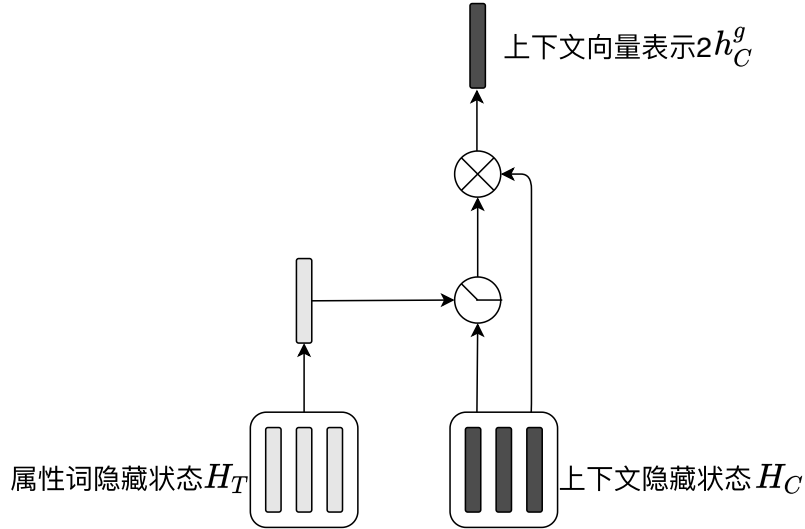


图 3: 门机制结构图

首先，通过将属性词隐藏状态 $H_T$ 进行平均池化如公式8所示，得到属性词的平均向量表示 $h_T^g \in \mathbf{R}^{d_h}$ 。

$$h_T^g = \sum_{i=1}^m H_T / m \quad (8)$$

然后通过门机制对上下文的表示进行调整，并把调整后的上下文表示当作是属性词对上下文表示的影响矩阵，具体门机制计算如公式9、10、11所示。

$$A^i = \text{relu}(H_C^i w_A + h_T^g w_g) \quad (9)$$

$$S^i = \text{tanh}(H_C^i w_s) \quad (10)$$

$$P^i = A^i \times S^i \quad (11)$$

其中， $w_A \in \mathbf{R}^{d_h}$ 、 $w_g \in \mathbf{R}^{d_h}$ 、 $w_s \in \mathbf{R}^{d_h}$ 为参数向量， $\text{relu}$ 、 $\text{tanh}$ 为激活函数， $P^i \in \mathbf{R}^{d_h}$ 为属性词向量通过门机制对上下文中第 $i$ 个向量进行选择后的表示，因此可以得到属性词向量对上下文中每个向量进行选择后的表示 $P \in \mathbf{R}^{n \times d_h}$ 。 $P$ 为经过属性词经过门机制对上下文进行选择后表示，本文这部分看作是属性词对上下文的影响矩阵。因此将 $P$ 与上下文原表示进行相加，并通过参数矩阵进行调整如公式12所示：

$$H_C^g = \text{relu}[(H_C + P)W_C] \quad (12)$$

其中， $W_P \in \mathbf{R}^{d_h \times d_h}$ 为参数矩阵，并通过自注意力机制得到上下文向量表示 $2h_C^g \in \mathbf{R}^{d_h}$ ，如公式13、14所示：

$$\gamma = \text{softmax}(H_C^g w_C) \quad (13)$$

$$h_C^g = \sum_{i=1}^n \gamma^i H_C^{g,i} \quad (14)$$

其中， $\gamma$ 为上下文注意力机制系数2， $\gamma^i$ 为上下文注意力机制系数2中的第 $i$ 个数， $H_C^{g,i}$ 为 $H_C^g$ 中的第 $i$ 个向量，通过自注意力机制后的 $h_C^g$ 为考虑了属性词对上下文这部分影响后的上下文向量表示2。

将上下文向量表示1、属性词向量表示和上下文向量表示2进行拼接之后得到 $r \in \mathbf{R}^{3d_h}$ ,并通过分类器得到属性词在上下文对应的情感极性分布,具体如以下公式所示:

$$J = - \sum_{i=1}^C g_i \log y_i + \lambda_r \left( \sum_{\theta \in \Theta} \theta^2 \right) \quad (15)$$

$$\Theta = \Theta - \lambda_l \frac{\partial J(\Theta)}{\partial \Theta} \quad (16)$$

其中 $C$ 为情感分类的总数,在本文中 $C = 3$ ,即积极、中性和消极, $g_i$ 为属性词在评论中真实的情感分布, $y_i$ 为通过模型对属性词情感的预测, $\Theta$ 为所有参数的集合, $\lambda_r$ 为 $L_2$ 正则化的参数, $\lambda_l$ 为更新参数的学习率。

## 4 实验

为验证基于层次注意力机制和门机制的模型的性能,本章将主要对实验设置、实验结果与分析 and 案例分析进行详细的介绍。

### 4.1 实验设置

**数据集:** 本文使用的数据集为公开的英文属性级别情感分析数据集SemEval 2014 Task 4 (Pontiki et al., 2014)和Twitter (Dong et al., 2014)数据集,其中SemEval 2014 Task 4由两部分组成,分别是笔记本领域和餐厅领域。数据集的具体分布情况如表1所示。

表 1: 数据集统计

数据集	积极	中立	消极
Laptop-Training	994	464	870
Laptop-Testing	341	169	128
Restaurant-Training	2164	637	807
Restaurant-Testing	728	196	196
Twitter-Training	1561	1560	3127
Twitter-Testing	173	173	346

**评价指标:** 本文使用平均准确率来衡量基于层次注意力机制和门机制的属性级别情感分析模型的性能,具体平均准确的公式如下所示。

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

其中 $TP$ 表示为真实类别为正例,预测类别也为正例; $TN$ 表示为真实类别为反例,预测类别也为反例; $FP$ 表示为真实类别为反例,预测类别也为正例; $FN$ 表示为真实类别为正例,预测类别也为反例。

**超参数:** 本文通过使用GloVe词向量对词进行初始化,词向量维度为200,同时GRU隐藏层的维度也设置为200;对于不在GloVe词典中的词均在 $[-0.1, 0.1]$ 之间随机取值,所有参数矩阵和向量的初始值均在 $[-0.1, 0.1]$ 之间随机选取,偏置的初始值均设置为0;为对参数进行调整,本文使用的优化器是Adam,学习率设置为0.01,同时为了防止过拟合,dropout设置为0.5。

### 4.2 实验结果与分析

为了衡量本文提出的基于层次注意力机制和门机制模型(HAG)的性能,本选取了7个基准模型与之进行比较,具体如下:

**SVM:** 通过简单的特征工程提取特征之后,再使用SVM分类器进行分类。

**TD-LSTM:** 将评论分成属性词上文与属性词、属性词与属性词下文两部分,然后将两部分分别经过两个LSTM之后的最后一个隐藏向量进行拼接,再通过softmax分类器分类 (Tang et al., 2016a)。

**ATAE-LSTM:** 将平均池化后的属性词向量与上下文中每个词向量进行拼接得到新的上下文矩阵, 然后将新的上下文矩阵通过LSTM得到上下文隐藏状态再与属性词向量拼接, 接下来通过注意力机制选择出对于属性词而言上下文中较为重要的词, 最后通过softmax分类器分类 (Wang et al., 2016)。

**MemNet:** 通过上下文和属性词之间的多层注意力机制加大有助于属性级别情感分析词的权重 (Tang et al., 2016b)。

**IAN:** 通过上下文和属性词之间的交互注意力机制加大对于属性词而言上下文中重要词的权重和加大对于上下文而言属性词中重要词的权重 (Ma et al., 2017)。

**RAM:**将评论句子输入双向LSTM之后, 利用多层注意机制综合句子中的重要特征 (Chen et al., 2017)。

**Co-attention:** 首先将评论分为属性词上文、属性词和属性词下文, 然后通过属性词上文与属性词、属性词与属性词下文和属性词上文与属性词下文之间的协同注意力机制得到6个向量表示, 将这6个向量进行拼接之后输入到softmax分类器得到属性词的情感分布 (Zhang et al., 2019)。

7个基准模型与HAG的对比实验结果如下表所示。

表 2: 实验结果

模型	Restaurant(%)	Laptop(%)	Twitter(%)
SVM	73.20	66.50	66.50
TD-LSTM (Tang et al., 2016a)	75.60	68.10	70.80
AEAT-LSTM (Wang et al., 2016)	77.20	68.70	68.60*
基准模型 MemNet (Tang et al., 2016b)	78.16	70.33	68.50
IAN (Ma et al., 2017)	78.60	72.10	70.40*
RAM (Chen et al., 2017)	80.23	<b>74.49</b>	69.36
Co-attention (Zhang et al., 2019)	80.40	73.20	71.70
Self-attention	79.03	71.63	70.08
本文模型 Hierarchical attention + Self-attention	80.26	72.10	71.39
HAG	<b>80.50</b>	72.50	<b>72.70</b>

在表2中“\*”表示作者根据原论文描述复现后的实验结果, 根据实验结果可以看出, 传统机器学习的SVM在三个数据集上的效果较差。相对于SVM, TD-LSTM由于使用了LSTM提取特征同时考虑了属性词信息, 因此在三个数据集上的实验效果有所提高。ATAE-LSTM通过注意力机制进一步选择出针对不同属性词而言, 上下文中不同词的重要性, 因此在一定程度上提高了实验的准确率。MemNet通过上下文和属性词之间的多层注意力机制进一步调整了上下文的表示。IAN认为上下文和属性词是相互选择, 因此利用交互注意力机, 通过上下文对属性词的表示进行加权、属性词对上下文进行加权。与MemNet相比, IAN更加充分的利用了属性词信息, 并通过上下文对属性词的表示进行调整, 使得实验结果更加准确。Co-attention通过属性词上文与属性词、属性词与属性词下文和属性词上文与属性词下文之间的协同注意力机制, 使得属性词上文、属性词和属性词下文三者之间相互选择。本文使用层次注意力机制, 在属性词的表示中加入了上下文的影响, 进一步得到属性词的表示, 使得属性词的表示更加准确; 在更新属性词表示之后再更新上下文的表示, 层次性注意力机制结果设计使得上下文的表示更加准确, 相对于使用自注意力机制, 层次注意力机制的增加在一定程度上提高了准确率。其次本文利用了门机制进一步的丰富了上下文的表示。实验结果表明HAG在Restaurant和Twitter数据集上取得了较好的实验室结果, 在Laptop数据集低于RAM, 这是由于Laptop数据集中含有如否定转移、转折等结构较为复杂的评论, RAM运用多重注意机制综合难句结构中的重要特征, 提高了实验的准确率。

### 4.3 案例分析

本节从餐厅领域选取了一条评论作为案例进行分析, 通过注意力机制系数的可视化以验证模型的有效性。如图4所示, 其中图的左侧为属性词, 右侧为属性词对应的上下文, 图中颜色的深浅分别代表了注意力机制系数 $\alpha$ 和 $\beta + \gamma$ 的大小。通过分析可知, 当属性词是“places”在上下文



中所关注的词为“cool”，而在案例中“so cool and”三个词的系数较大，当属性词变为“service”的时候，“is prompt and curious”几个词的权重加大，“cool”的权重降低，由此可见模型的合理性。

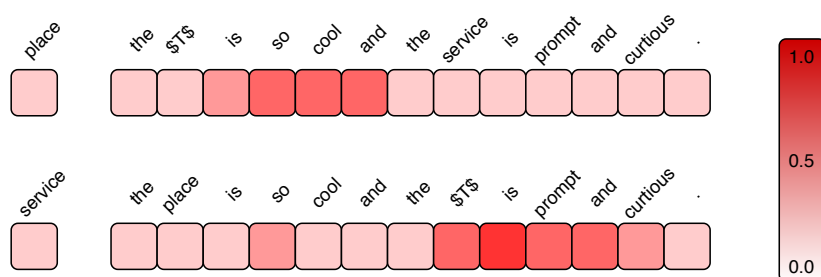


图 4: 案例分析

## 5 结论

本文提出了一种基于层次注意力机制和门机制的属性级别情感分析方法，通过层次注意力机制先利用上下文对属性词进行选择作用，并将选择后的属性词表示认为是上下文对属性词的影响矩阵从而更新属性词的表示，在得到属性词新的表示之后，再更新上下文的表示，以此在属性词的表示中融入上下文，在上下文的表示中融入属性词信息，层次性的注意力机制结构设计使得上下文和属性词的表示更加准确。其次，本文通过门机制选择出针对于属性词而言上下文中有用的信息，进一步的丰富了上下文的表示，实验结果证明本文提出的HAG模型的有效性。在未来的工作中将会考虑利用更多的信息以丰富词的嵌入表示，如加入一些词性信息、情感信息和位置信息等，在词向量的使用上将会考虑信息更加丰富的BERT词向量，同时会对复杂句式的属性级别情感作进一步研究。

## 致谢

基金项目:深圳市科创委知识创新基础研究项目(JCYJ20160527172144272); 广东省教育厅特色创新项目(2018GKTSCX069)

## 参考文献

- Peng Chen, Zhongqian Sun, Lidong Bing, and Wei Yang. 2017. Recurrent attention network on memory for aspect sentiment analysis. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017*, pages 452–461. Association for Computational Linguistics.
- Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou, and Ke Xu. 2014. Adaptive recursive neural network for target-dependent twitter sentiment classification. volume 2, pages 49–54, 06.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177.
- Binxuan Huang and Kathleen M Carley. 2019. Syntax-aware aspect level sentiment classification with graph attention networks. *arXiv preprint arXiv:1909.02606*.
- Binxuan Huang, Yanglan Ou, and Kathleen M Carley. 2018. Aspect level sentiment classification with attention-over-attention neural networks. pages 197–206.
- Lishuang Li, Yang Liu, and AnQiao Zhou. 2018. Hierarchical attention based position-aware network for aspect-level sentiment analysis. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 181–189.
- Haihui Li, Yun Xue, Hongya Zhao, Xiaohui Hu, and Sancheng Peng. 2019a. Co-attention networks for aspect-level sentiment analysis. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pages 200–209. Springer.

- Haihui Li, Ting Yuan, Haiming Wu, and Yun Xue. 2019b. Granular computing-based multi-level interactive attention networks for targeted sentiment analysis. *Granular Computing*, 04.
- Yunlong Liang, Fandong Meng, Jinchao Zhang, Jinan Xu, Yufeng Chen, and Jie Zhou. 2019. A novel aspect-guided deep transition model for aspect based sentiment analysis. pages 5568–5579.
- Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. Interactive attention networks for aspect-level sentiment classification. *arXiv preprint arXiv:1709.00893*.
- Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*, pages 27–35, 01.
- Tahura Shaikh and Deepa Deshpande. 2016. A review on opinion mining and sentiment analysis. *International Journal of Computer Applications*, 975:8887.
- Youwei Song, Jiahai Wang, Tao Jiang, Zhiyue Liu, and Yanghui Rao. 2019. Attentional encoder network for targeted sentiment classification. *arXiv preprint arXiv:1902.09314*.
- Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2016a. Effective LSTMs for target-dependent sentiment classification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 3298–3307, Osaka, Japan, December. The COLING 2016 Organizing Committee.
- Duyu Tang, Bing Qin, and Ting Liu. 2016b. Aspect level sentiment classification with deep memory network. pages 214–224.
- Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. 2016. Attention-based lstm for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 606–615.
- Wei Xue and Tao Li. 2018. Aspect based sentiment analysis with gated convolutional networks. *arXiv preprint arXiv:1805.07043*.
- Peiran Zhang, Hongbo Zhu, Tao Xiong, and Yihui Yang. 2019. Co-attention network and low-rank bilinear pooling for aspect based sentiment analysis. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6725–6729. IEEE.
- Shiliang Zheng and Rui Xia. 2018. Left-center-right separated neural network for aspect-based sentiment analysis with rotatory attention. *arXiv preprint arXiv:1802.00892*.
- 欧阳志凡. 2018. 基于依存树和注意力的属性级别情感分类研究. Master's thesis, 华南理工大学.