Computer Science and Application 计算机科学与应用, 2020, 10(11), 2064-2077
Published Online November 2020 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2020.1011218

Hans汉斯

基于自注意力的门控卷积神经网络的要素类情感分类研究

张　颖，郑建国

东华大学旭日工商管理学院，上海
Email: 965279586@qq.com

收稿日期：2020年11月4日；录用日期：2020年11月19日；发布日期：2020年11月26日

摘　要

[目的/意义]近年来，将注意力机制与LSTM模型结合的方法常被用于要素类情感分类任务，但是该方法存在参数多、训练时间长的弊端。门控卷积神经网络模型不仅结构简单、参数少、运行时间短，还能分别提取要素特征和情感特征、具有较高的分类精度，但是该模型采用的要素类嵌入是预定义的、与上下文无关。对要素类情感分类任务来说，要素类的质量对预测文本在要素类上情感极性的准确率的重要性不言而喻。[方法/过程]本文关联要素类提取和要素类情感分类任务，提出融合自注意力机制下的要素类特征的门控卷积神经网络模型，通过结合自注意力机制的神经网络提取出基于上下文优化的要素类嵌入，然后将优化后的要素类向量和文本词向量通过门控卷积神经网络进行训练。[结果/结论]在2014年至2016年的SemEval数据集上的实验结果表明，本文提出的模型能有效改善要素类提取的效果和提高要素类情感分类的分类准确率。

关键词

要素类提取，要素类情感分类，自注意力机制，门控卷积神经网络

Research on Aspect Category Sentiment Classification Based on Gated Convolution Neural Network Combined with Self-Attention Mechanism

Ying Zhang, Jianguo Zheng

School of Management, Donghua University, Shanghai
Email: 965279586@qq.com

Received: Nov. 4th, 2020; accepted: Nov. 19th, 2020; published: Nov. 26th, 2020

文章引用: 张颖, 郑建国. 基于自注意力的门控卷积神经网络的要素类情感分类研究[J]. 计算机科学与应用, 2020, 10(11): 2064-2077. DOI: 10.12677/csa.2020.1011218

Abstract

[Purpose/Significance] In recent years, a common method for aspect category sentiment classification is to combine LSTM model with attention mechanism. Compared to that, the gated convolutional neural network model not only has a simple structure, fewer parameters and shorter training time, but also achieves higher classification accuracy being able to extract aspect features and emotion features. [Method/Process] Considering that the quality of aspect category is crucial for aspect category sentiment classification, this paper coupled aspect category extraction and aspect category sentiment classification, and put forward Gated Convolutional Neural Network with Self Attention-based Aspect Embedding (GCAE_SelfAtt) model to relate the aspect category embeddings to corresponding context, and to achieve a higher accuracy. [Result/Conclusion] The experiment on SemEval dataset shows that GCAE_SelfAtt model does help to extract more coherent aspect categories and achieve higher accuracy for sentiment classification.

Keywords

Aspect Category Extraction, Aspect Category Sentiment Classification, Self-Attention Mechanism, Gated Convolutional Nueral Network

Copyright © 2020 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

伴随着移动端互联网的快速发展，越来越多的人通过评论在社交平台和电商平台来表达自己的观点与情感，这种行为也越来越日常化，从而产生了海量的文本数据。如何利用这些评论文本，挖掘用户的观点和分析用户的情感，具有重要的实践意义与实用价值。根据评价文本的层次不同，情感分类任务可以分为文档级情感分类、句子级情感分类和要素级情感分类三种。相较于前者，面向细粒度的要素级情感分类属于更深层次的情感分类任务，需要提供要素从而预测句子在特定要素上所传达的情感倾向性。

如果根据要素的种类进行划分，要素级情感分类可分为要素项情感分类(Aspect Term Sentiment Classification, ATSC)和要素类情感分类(Aspect Category Sentiment Classification, ACSC)两种[1]。以句子“这家的面包很不错，就是太远了”为例，文本中出现的“面包”是要素项，“不错”反映出评价者对该要素项具有正向的情感极性；“面包”和“远”在语义上分别属于“食物”和“距离”这两个要素类，“不错”和“太远”反映出评价者对“食物”和“距离”分别具有正向和负向的评价。一般来说，要素类可以通过对提取出的要素项进行聚类来得到，如从语料库中提取出的“面包”、“蛋糕”、“牛排”等要素项会被聚类为“食物”这一要素类。不同于要素项情感分类的是，要素类情感分类所针对的要素类不一定会出现在评论文本，因此需要挖掘更深层的信息。

传统的文本情感分类技术主要分为基于规则的方法和基于统计的机器学习方法[2]，然而前者没有考虑上下文的语义信息、容易受限于人工制定的规则与提供的情感词典，后者存在表达能力弱、高度依赖特征工程、无法自动从数据中提取有意义的特征的缺点。深度学习凭借其强大的特征学习能力，大大减少了特征的设计与抽取工作量，近年来被广泛应用于情感分类任务，取得了比传统方法更优的分类效果[3] [4]。例如，Socher 等人[5]提出的递归神经网络模型，在正负情感倾向性分类和细粒度情感分类任务上都取得了优异表现；Kim [6]将文本转化为 word2vec 词向量并输入 CNN 模型，显著提高了分类精度；

Wang 等人[7]采用 LSTM 模型来保留文本中词语之间的依赖关系，实现文本情感分类；Mnih 等人[8]将注意力机制与 LSTM 模型结合，以解决远距离词的情感信息获取问题。

本文以深度学习相关网络模型作为研究方法，以要素类情感分类作为研究任务，提出融合自注意力机制下的要素类特征的门控卷积神经网络模型，先后进行要素类提取和要素类情感分类实验。具体地，本文首先在结合注意力机制的神经网络模型的基础上，引入自注意力机制提取要素类，然后将基于上下文优化后的要素类嵌入与文本词嵌入输入基于门控机制的卷积神经网络模型，对要素类进行情感极性的预测。该模型在 SemEval 的 Restaurant 数据集上的实验结果表明，自注意力机制的引入能有效提高结合注意力机制的神经网络模型的要素类提取效果，以及将基于上下文优化后的要素类嵌入与文本词嵌入引入门控卷积神经网络能有效提高要素类情感分类的精度。

2. 相关工作

在之前的研究中，要素类提取的方法主要分为三种，分别是基于有监督学习的序列模型、基于语言规则的方法和基于无监督学习的主题模型[9]。一般来说，要素类提取，可以拆解成先提取要素项和再对要素项聚类的两个过程。无监督学习由于不需要提供带标注的数据，在要素类提取任务上的应用比较广泛。例如，LDA 主题模型在不需要单独执行要素项提取以及要素项聚类的情况下，就能直接获得要素类。基于 LDA 的模型识别出的要素类虽然能较好地描述语料库，但质量不佳、连贯性较差。考虑到主题的连贯性主要受词共现统计的影响[10]，Yan 等人[11]提出能够生成共现词对的双项主题模型。Wang 等人[12]提出基于受限玻尔兹曼机的模型，同时提取句子中的要素类和要素类对应的情感极性，但是存在依赖大量的先验知识的缺点。He 等人[13]指出可以通过神经网络模型中的词嵌入来统计词共现分布，通过注意力机制来降低对无关单词的关注，在提取要素类的同时还能得到要素类的代表词。

早期的要素级情感分类研究大多首先借助情感词典，得到与上下文无关的情感词所在句子的情感极性，然后再根据与要素有关的相应分句的情感极性来推断该要素的情感极性[14]。然而，许多情感词的情感倾向在不同的上下文中是不同的[15]。这种方法没有对上下文和要素间的语义相关性进行建模，而这恰恰是要素级情感分类的难点，与提高分类精度的关键。Mnih 等人[8]指出神经网络可以捕获要素和上下文词语之间的语义联系，即首先分别表示要素和上下文，然后通过模型为要素识别重要的与情感相关的上下文。表 1 总结了近年来一些具有代表性的基于深度学习的要素级情感分类研究。

Table 1. Summary on aspect-level sentiment classification research based on deep learning
表 1. 基于深度学习的要素级情感分类研究总结

研究工作	模型简称	区别上下文的方式	研究亮点	任务级别
Tang 等人[16]	TD-LSTM TC-LSTM	无	将上下文按照分成两部分； 将要素项表示和词嵌入表示合并作为输入	ATSC
Wang 等人[17]	AT-LSTM ATAE-LSTM	单层注意力	将注意力机制与 LSTM 模型结合	**ACSC**
Liu 和 Zhang [18]	BiLSTM-ATT BiLSTM-ATT-C BiLSTM-ATT-G	单层注意力	采用双向 LSTM 模型； 将左右上下文与整句结合； BiLSTM-ATT-G 的门控单元	ATSC
Tang 等人[19]	MemNet	多层注意力	采用多层注意力机制； 应用端对端记忆网络； 内容信息和距离信息结合	ATSC
Ma 等人[20]	IAN	单层注意力	要素项和上下文交互； 分别计算要素项和上下文注意力；	ATSC
Chen 等人[21]	RAM	多层注意力	加权记忆机制； 将注意力结果与 GRU 结合	ATSC

Continued

Li 等人[22]	TNet	卷积特性与距离特征	用 CNN 取代注意力机制； 用 CPT 模块捕捉要素项信息； 动态生成要素项向量	ATSC
Tang 等人[23]	/	/	提出渐进自监督注意力机制算法 且成功应用于现有模型	ATSC
Xue 和 Li [1]	GCAE	门控机制和卷积特性	提出的门控卷积神经网络能够有效、 高效的实现 ATSC 和 ACSC	ATSC & **ACSC**

上述文献反映出：第一，要素特征的有效提取是深度学习模型在要素级情感分类任务上的关键；第二，注意力机制能够使模型区分不同的上下文词对要素的重要性，从而将注意力集中在与特定要素相关的文本上；第三，CNN 模型的卷积特性和池化操作可在一定程度上起到注意力机制的作用，具有并行化训练优势；第四，要素级情感分类的研究集中在要素项情感分类任务上，要素类情感分类任务亟待研究。

3. 融合自注意力机制下的要素类特征的门控卷积神经网络模型

为了关注更多有价值的信息，本文在结合注意力机制的神经网络(Attention-based Aspect Extraction, ABAE)模型和基于门控机制的卷积神经网络(Gated Convolutional Neural Network with Aspect Embedding, GCAE)模型的基础上，联动要素类提取和要素类情感分类任务，提出了一种融合自注意力机制下的要素

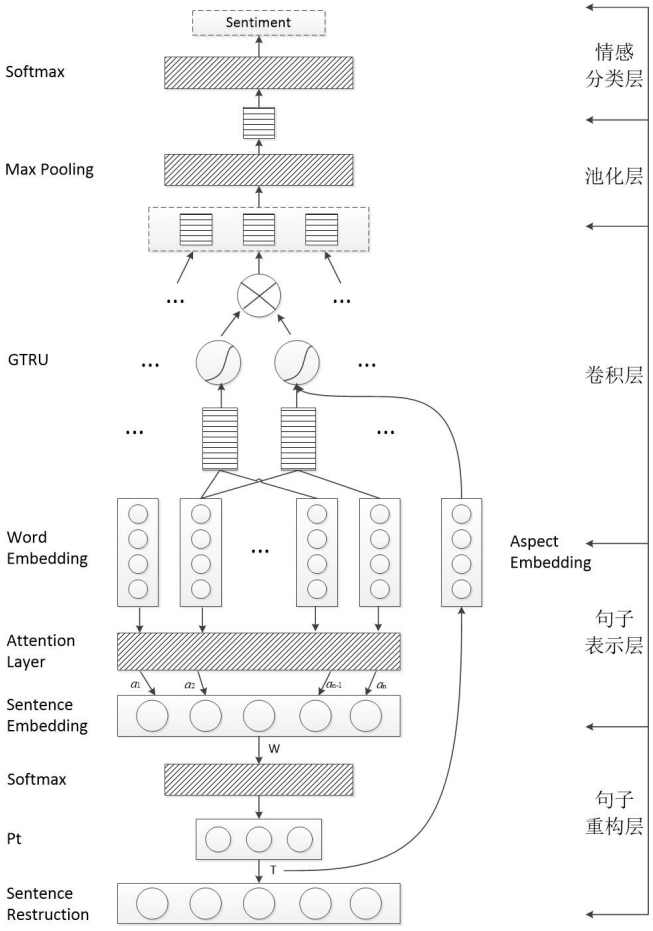

Figure 1. Schematic diagram of GCAE_SelfAtt model
图 1. 融合自注意力机制下的要素类特征的门控卷积神经网络模型示意图

类特征的门控卷积神经网络模型(Gated Convolutional Neural Network with Self Attention-based Aspect Embedding, GCAE_SelfAtt)，由结合自注意力机制的神经网络和门控卷积神经网络构成。其中，要素类提取任务的最终目的是学习一组要素类嵌入，使得每个要素类可以通过词嵌入空间附近的代表性词语来解释。

模型的示意图如图 1 所示，由词向量层、句子表示层、句子重构层、卷积层、池化层和情感分类层构成。其中，词向量层作为中间层连接要素类提取任务和要素类情感分类任务，句子表示层通过基于自注意力机制的编码器构建句子的向量表示，句子重构层根据要素类嵌入矩阵的线性组合重构句子的向量表示。

3.1. 任务定义

对于一则长度为 n 的句子 $s=\{w_1,w_2,\cdots,w_n\}$，文本中可能存在一个或多个要素类，但要素类不一定显式地以词语或短语的方式出现在句子中，因此，要素类情感分类任务的目标是决定句子中的要素类以及预测相关文本在对应要素类上所反映的情感极性。本文在预定义要素类的个数的情况下，首先通过无监督学习为每个句子提取出相应要素类的向量表示，再通过与代表词对照，关联句子的真实要素类标签，进一步预测句子在该要素类上的情感极性。

3.2. 构建词向量输入层

由于词嵌入会将经常共同出现在同一上下文语境的词语映射在词嵌入空间中的相近位置上，本文将语料库中句子的每个词语都映射为低维的、连续的实值向量，即词向量，实现词共现分布的编码。所有的词向量构成了一个词向量矩阵，矩阵的行数表示词向量纬度，矩阵的列数表示词典中词语个数。

3.3. 句子表示层

本文通过注意力编码为每个词语赋予相应的注意力权重，过滤句子中的与要素类无关的词语，从而提高所提取的要素类的连贯性，并得到句子的向量表示 z_s。

如果采用一般的注意力机制，则句子中每个词语的注意力权重 a_i 是通过计算每个词语的词向量和句子的平均词向量之间的相关性得到的。

$$y_s=\frac{1}{n}\sum_{i=1}^{n}e_{w_i} \tag{1}$$

$$d_i=e_{w_i}^{T}\cdot M\cdot y_s \tag{2}$$

$$a_i=\frac{\exp(d_i)}{\sum_{j=1}^{n}\exp(d_j)} \tag{3}$$

$$z_s=\sum_{i=1}^{n}a_i e_{w_i} \tag{4}$$

其中，n 表示句子长度，e_{w_i} 表示文本词向量，$M\in\mathbb{R}^{d\times d}$ 是待学习的转化矩阵。

如果采用自注意力机制，则句子中每个词语的注意力权重 a_i 是通过计算每个词语和句子中的其他词语之间的相关性得到的。

$$s(Q_i,K_j)=\frac{Q_iK_j^{\mathrm{T}}}{\sqrt{d_z}}=\frac{\left(e_{w_i}W^{Q}\right)\left(e_{w_j}W^{K}\right)^{\mathrm{T}}}{\sqrt{d_z}} \tag{5}$$

$$\alpha_{ij}=\frac{\exp\left(s(Q_i,K_j)\right)}{\sum_{j=1}^{n}\exp\left(s(Q_i,K_j)\right)} \tag{6}$$

$$z_i = \sum_{j=1}^{n} \alpha_{ij} V^j = \sum_{j=1}^{n} \alpha_{ij} \left(e_{wj} W^V \right) \tag{7}$$

由于查询状态和候选状态属于相同的序列，自注意力机制能够无视词语之间的距离直接计算词依赖关系，学习句子的内部结构。如图 2 所示，本文采用自注意力机制进行句子的向量表示，并以键值对的形式输出模块的示意图。

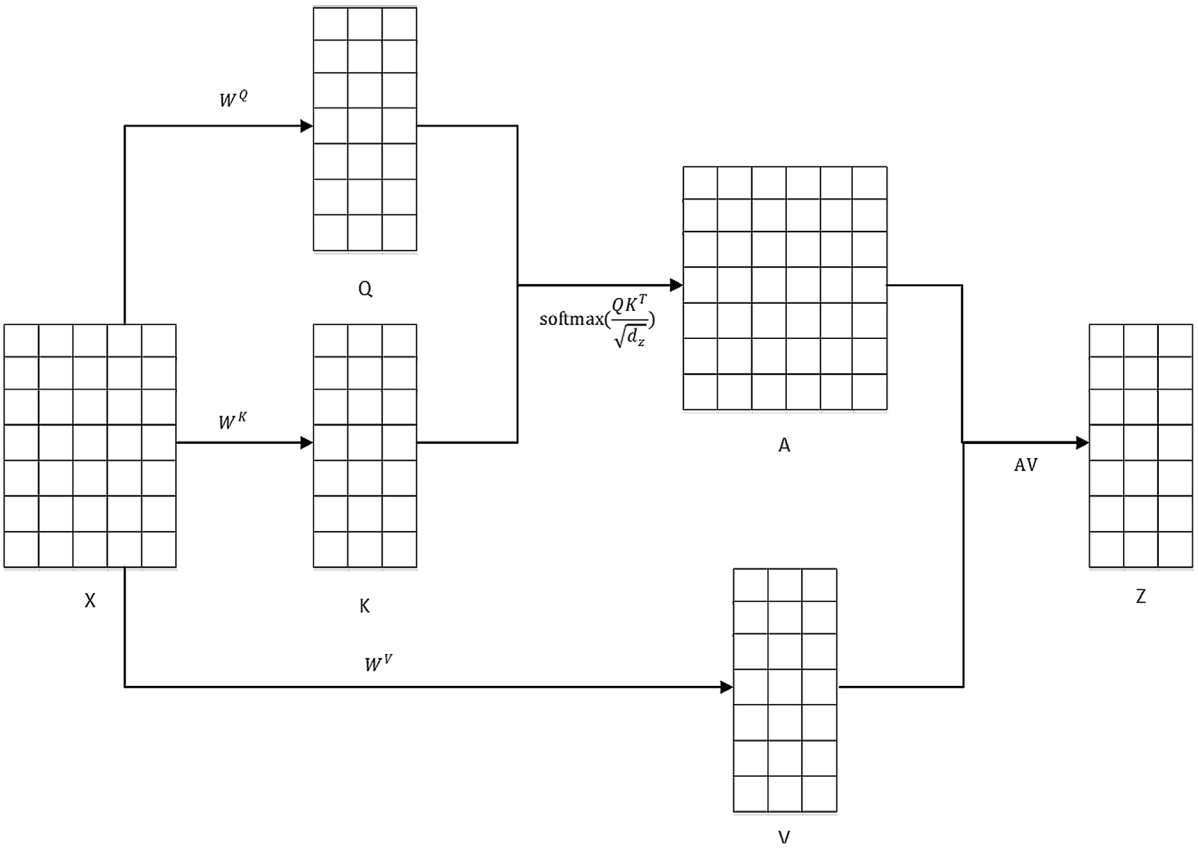

Figure 2. Schematic diagram of vector representation of sentences based on self-attention mechanism
图 2. 基于自注意力机制下的句子的向量表示示意图

具体地，首先使用大小均为 $d_x \times d_z$ 但数值不同的 3 个初始化矩阵 W^Q，W^K，W^V，对输入 $X \in \mathbb{R}^{n \times d_x}$ 进行线性变换，分别得到查询向量矩阵 Q、键向量矩阵 K、值向量矩阵 V 以增强模型的表示能力；然后以缩放点积函数作为对齐分数函数 $s(Q,K)$ 计算查询向量和键向量之间的相似度后，将对齐分数函数进行 softmax 归一化，得到注意力权重矩阵 A；最后计算句子中各个词语的注意力权重与对应词向量的加权和，得到句子的向量表示 $Z = [z_1, z_2, \cdots, z_n] \in \mathbb{R}^{n \times d_z}$，作为模块的输出。在本文中，令 $d_x = d_z$。

$$Q = XW^Q \tag{8}$$

$$K = XW^K \tag{9}$$

$$V = XW^V \tag{10}$$

$$Z = \text{Attention}(Q,K,V) = \text{softmax}\left(s(Q,K)\right)V \tag{11}$$

3.4. 句子重构层

本文首先通过对句子的向量表示 z_s 进行降维处理和 softmax 处理，得到要素类的归一化权重向量 p_t，

然后根据要素类的归一化权重向量 p_t 和要素类嵌入矩阵 T，将句子重新表示为要素类嵌入的线性组合，得到句子的重构表示 r_s 。

$$p_t = \text{softmax}\left(W \cdot z_s + b\right) \tag{12}$$

$$r_s = T^{\mathrm{T}} \cdot p_t \tag{13}$$

其中，W 是权重矩阵，b 是偏置向量，要素类的归一化权重向量 p_t 的维度 K 表示预定义的要素类的数量，K 个元素值表示句子分别属于这 K 个要素类的概率。

3.5. 卷积层与 GTRU

在经过要素类提取实验之后，句子重构层的要素类嵌入矩阵 T 形成了基于上下文优化后的要素类词向量。考虑到要素类情感分类和要素类提取采用同一套数据集，保证词嵌入空间的一致性有利于模型的效果，因此本文采用优化后的文本词向量作为要素类情感分类的输入层。

本文在词向量输入层上配置了两个独立的卷积层，每个卷积层通过多个卷积核提取输入的不同特征，并分别与 Tanh 门控单元和 ReLu 门控单元连接，然后对门控单元输出的结果进行组合。相比于 Tanh 门控单元仅接收卷积层的输出，ReLu 门控单元在接收卷积层的输出的同时，还接收特定要素类的词向量 v_a 。Tanh-ReLU 门控单元(Gated Tanh-ReLU Unit, GTRU)具有能够控制情感特征通往池化层的流动路径的特性，使得模型能根据给定的要素类有选择地输出情感特征。具体地，将激活函数为 ReLU 的卷积操作生成的要素类特征 a_i 和激活函数为 Tanh 的卷积操作生成的情感特征 s_i ，逐元素相乘可以得到特定要素类的情感特征 c_i 。

$$a_i = \text{relu}\left(X_{i:i+k} * W_a + V_a v_a + b_a\right) \tag{14}$$

$$s_i = \tanh\left(X_{i:i+k} * W_s + b_s\right) \tag{15}$$

$$c_i = s_i \times a_i \tag{16}$$

其中，W_a 、V_a 和 W_s 是权重矩阵，b_a 和 b_s 是偏置向量。

如果文本中含有多个要素类，GTRU 可以通过 ReLU 门自动忽略与给定要素类无关的其他要素类的情感信息。例如，在句子“这家的面包很不错，就是太远了”中，当给定的要素词为“食物”时，门控单元会自动忽略另一个分句中的要素词“距离”的情感极性，只输出“食物”所在分句的情感极性。这是因为，ReLU 门可以根据给定的要素类信息 v_a 和某位置处的要素类特征 a_i 的相关性计算相似度得分，从而决定该位置处的情感特征 s_i 是否能够通过 Relu 门控单元。

3.6. 池化层

除了门控机制外，池化层中的最大池化操作也可以有选择性地从文本中提取和给定的要素类相关的情感特征，从而实现对特定要素类的情感极性的预测。

对一则由 n 个词语组成的句子来说，用一个尺寸为 k 的卷积核对其进行卷积操作会产生 $n-k+1$ 个特征，并构成该句子的卷积特征图 c。

$$c = \left[c_1, c_2, \cdots, c_{n-k+1}\right] \tag{17}$$

本文在池化层对卷积特征图 c 进行最大池化处理，形成一个维度固定的特征向量 $\hat{c}$ ，即通过保留卷积特征图中的最大值来保留整个句子中最显著的基于特定要素类的情感特征，从而移除不重要的情感特征。

$$\hat{c} = \max(c) \tag{18}$$

3.7. 情感分类层

本文将 N 个卷积核在卷积层中得到的 N 个卷积特征图分别输入池化层进行最大池化操作，然后将池化层输出的 N 个特征向量 $\hat{c}_i$ 拼接为一个维度固定的特征向量 $\hat{C}$，再将其输入全连接层并通过 softmax 分类器进行情感分类，输出文本在给定要素类上的预测情感极性 $\hat{y}$：

$$\hat{C} = [\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_N] \tag{19}$$

$$\hat{y} = \text{softmax}\left(W \cdot \hat{C} + B\right) \tag{20}$$

其中，W 是全连接层权重矩阵，B 是偏置。

3.8. 模型训练

本文采用端到端的反向传播算法对模型进行训练，以对比最大边际函数 $J(\theta)$ 作为第一目标函数，即最大化句子的重构表示 r_s 和句子的目标表示 z_s 之间的内积的同时，最小化句子的重构表示 r_s 和随机采样的 m 个负样本句子的向量表示 n_i 之间的内积，减少重构误差以在提取的要素类嵌入中尽可能多地保留要素类的信息。为鼓励要素类嵌入矩阵 $T \in \mathbb{R}^{K \times d}$ 的行正交性，本文还引入正则项 $U(\theta)$。

$$J(\theta) = \sum_{s \in D} \sum_{i=1}^{m} \max(0, 1 - r_s z_s + r_s n_i) \tag{21}$$

$$U(\theta) = \left\| T_n \cdot T_n^{\mathrm{T}} - I \right\| \tag{22}$$

$$L(\theta) = J(\theta) + \lambda U(\theta) \tag{23}$$

其中，D 表示数据集，s 表示句子，T_n 是对要素类嵌入矩阵 T 中的每一行标准化后的结果，λ 是正则项的权重参数，$\theta = \{E, T, M, W, b\}$ 表示待训练的参数。

本文以交叉熵损失函数 $\mathcal{L}$ 作为第二目标函数，即最小化在所有训练语料上的特定要素类的真实情感极性和预测情感极性之间的交叉熵损失和。

$$\mathcal{L} = -\sum_i \sum_j y_i^j \log \hat{y}_i^j \tag{24}$$

其中，i 表示句子的索引；j 表示情感类别的索引，共分成 positive、neutral 和 negative 三类；y 表示特定要素类的真实情感类别，$\hat{y}$ 表示特定要素类的预测情感类别。

4. 实验

本文将融合自注意力机制下的要素类特征的门控卷积神经网络模型先后应用在要素类提取和要素类情感分类任务上，首先在 keras 环境下通过结合自注意力机制的神经网络模型来确定文本的要素类及要素类的向量表示，然后在 pytorch 环境下运行带要素类嵌入的门控卷积神经网络模型来确定文本在某一要素类上反映出的情感极性。本文采用 word2vec 预训练的 300 维词向量对文本输入进行初始化，对词典中未出现的词则采用均匀分布 $U(-0.25, 0.25)$ 进行随机初始化。

4.1. 实验数据

本文实验采用的英文数据来自于 SemEvalABSA 任务下的 2014 年至 2016 年的 restaurant 评论下的训练集和带标注的测试集，以 XML 格式标记。每条评论都包含一系列的要素类和要素类相关的情感极性，<text>对应句子的文本内容、<aspectCategory>或是<Opinion>中的 category 和 polarity 分别表示句子的要

素类和在该要素类上的情感极性。所有评论一共包含 positive、neutral、negative 和 conflict 四种情感，{restaurant, food, drinks, ambience, service, price, anecdotes/miscellaneous, location}八种要素类。鉴于情感极性为“conflict”的文本数据量过少且“conflict”这一情感极性较难识别，相关文献都不考虑这类情感，因此本文将情感标签中的“conflict”用“neutral”替代。经过将 3 年的数据进行整理与整合，每一条评论对应存入一个字典，字典的三个键“sentence”、“aspect”和“sentiment”分别存放评论的文本内容、要素类和在该要素类上的情感极性。若一条评论文本中含有多个要素类，则句子中的文本内容以及不同的要素类和对应的情感极性将被分别存放进不同的字典。

原训练集和原测试集最终分别形成 3542 个字典和 2192 个字典，记作数据集 1 和数据集 2。出于评估本文模型在多要素类情感分类任务上的表现的需要，本文将两份数据集中的含有多个要素类的句子提取出来并单独存放进字典，分别创建了复杂数据集 1 和复杂数据集 2。经过整理，四份数据集的统计分布结果如表 2 所示。在要素类提取任务中，数据集 1 和数据集 2 将分别作为训练集和测试集；在要素类情感分类任务中，数据集 1、数据集 2 和复杂数据集 2 将分别作为训练集、验证集和测试集。

Table 2. Data distribution of the datasets
表 2. 数据集统计分布结果

	Positive	Negative	Neutral	总计
数据集 1	2173	999	370	3542
复杂数据集 1	924	325	105	1354
数据集 2	1378	639	175	2192
复杂数据集 2	545	201	50	796

在整合数据后，本文对评论文本进行了无词义符号替换、大写转小写、词性还原、要素类筛选等预处理操作。即便是人工对 anecdotes 和 miscellaneous 进行要素类标注，相关界定也很模糊，因此将含有“misc”要素类的字典在预处理阶段从数据中剔除。考虑到去停用词不可避免地会丢失文本的部分信息，众多研究也表明去停用词比保留停用词在情感分类任务上获得的精度更低，本文不进行去停用词的预处理操作。此外，本文还进行了建立词典、torchtext 分词、文本数值化、文本长度补全、数据分批等操作。

4.2. 实验参数设置

文本中词向量的维度、要素类词向量的维度都为 300。要素类提取实验以 keras 作为实验框架，以 Adam 作为优化器，负采样样本数为 6，dropout 为 0.5，初始学习率为 0.01，训练轮数为 50 轮，批数据大小为 256，正则项权重为 0.001。要素类情感分类实验以 pytorch 作为实验框架，以 Adam 作为优化器，设置窗口大小为 2、3、4 的滤波器各 100 个，dropout 为 0.5，初始学习率为 0.01，训练轮数为 30 轮，批数据大小为 256。

4.3. 要素类提取实验结果与实验分析

本文通过结合注意力机制的神经网络(ABAE)模型或结合自注意力机制的神经网络(SABAE)模型提取要素类，学习要素类的词向量表示。要素类的代表性词语可以通过与该要素类经常一起出现在评论文本中的词语来表示，因此可以通过同一词嵌入空间中的要素类的邻近词语来解释所提取的要素类向量，从而与要素类的名称映射。例如，当代表性词主要是各种食物及与食物相关的形容词时，可以为该要素类向量指定“food”这一要素类名称。实验的理想状态是能将提取出的 8 个要素类向量一一映射至 restaurant、food、drinks、ambience、service、price、misc 和 location 这 8 个要素类名称上。

将 ABAE 模型和 SABAE 模型训练所得的 8 个要素类向量，根据在词嵌入空间中余弦距离最近的前 30 个词语的词向量进行降维，分别得到如图 3 与图 4 所示的可视化结果图。与大多数研究常用的主成分分析方法不同的是，本文采用的 t-SNE 属于非线性降维方法，首先将数据点之间的相似度转换为概率，然后通过原始空间和嵌入空间的联合概率的 KL 散度来评估可视化效果的好坏。t-SNE 是目前效果最好的数据降维与可视化方法之一，尤其适合不确定数据集是否有较好的可分性的情况。

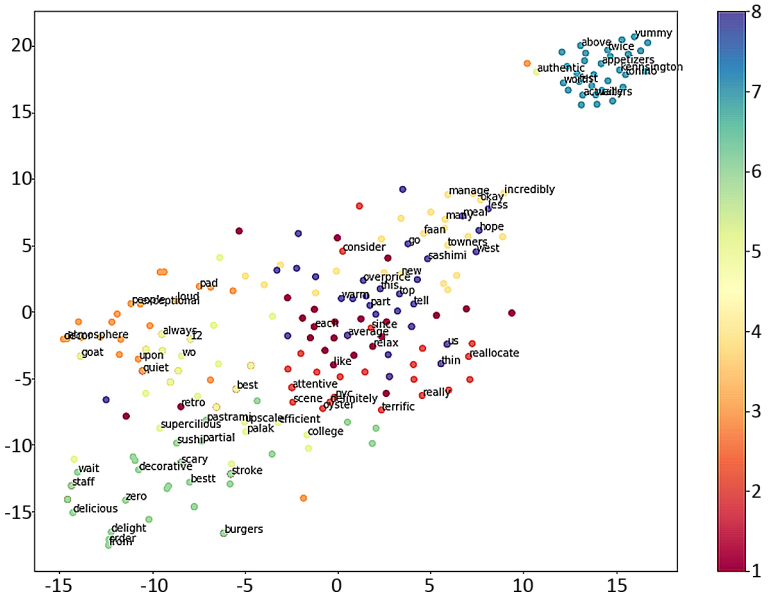

Figure 3. Visualization diagram of the representative words of the aspect categories of ABAE model

图 3. ABAE 模型的要素类代表词可视化图

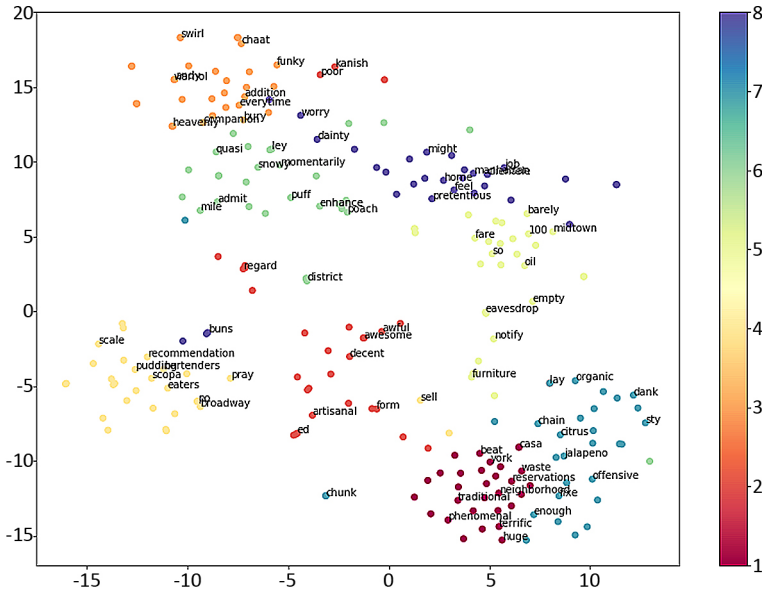

Figure 4. Visualization diagram of the representative words of the aspect categories of SABAE model

图 4. SABAE 模型的要素类代表词可视化图

就可视化的聚类效果而言，SABAE 模型提取的要素类的代表性词语的词向量聚类效果优于 ABAE

模型。以 SABAE 模型所在的图 4 的部分要素类及其代表词为例，色阶为 1 的代表词包括“traditional (传统的)”、“casual (休闲的)”、“cute (可爱的)”、“attractive (有吸引力的)”、“terrific (极好的)”等形容词和“lunch (午餐)”、“buffet (自助餐)”、“reservations (预订)”、“neighborhood (附近地区)”等名词，对应要素类 restaurant；色阶为 2 的代表词包括 kielbasa (波兰香肠)、apples (苹果)、seafood (海鲜)、stock (食物存货)、glaze (糖浆)、asparagus (芦笋)、dumplings (饺子)、omakase (日料——无菜单料理)等食物的要素项词语和“poor (糟糕的)”、“awesome (好吃的)”、“awful (难吃的)”、“aesthetic (有美感的)”、“decent (不错的)”、“artisanal (手工的)”等形容词，对应要素类 food。

鉴于要素类提取的目标之一是识别出的要素类能符合人类用户的标准，因此对模型识别出的要素类进行用户评估是有必要的，本文通过人工裁判来判定是否每个要素类的 TopN 代表词中的大多数都能较一致地表示该要素类。根据 Chen 等人[24]和 He 等人[13]的实验设置，本文招募了三名人类裁判，分别对 ABAE 模型和 SABAE 模型所识别出的 8 个要素类及其代表词进行用户评估。对单个要素类而言，如果有不少于 2 个人类裁判都认为该要素类的前 30 个代表词中的大多数都一致地表示同一个要素类，则将该要素类标记为具有一致性的要素类。ABAE 模型的一致性要素类的用户评估数量为 4，SABAE 模型的一致性要素类的用户评估数量为 6。

对于被裁定为一致性要素类来说，人类裁判同样需要对一致性要素类的代表词进行评估，如果有不少于 2 个人类裁判都认为该代表词与要素类是有关联的，则将该代表词标记为准确的代表词。在不同的 N 值下，本文分别统计了 ABAE 模型和 SABAE 模型中的各个一致性要素类的代表词的准确比例，然后联合所有一致性要素类，计算代表词的平均准确率。图 5 显示了 ABAE 模型及 SABAE 模型，在 N 值分别为 10、30 和 50 时，代表词的平均准确率。由图可知，无论代表词的数量为多少，SABAE 模型的代表词的平均准确率均高于 ABAE 模型。

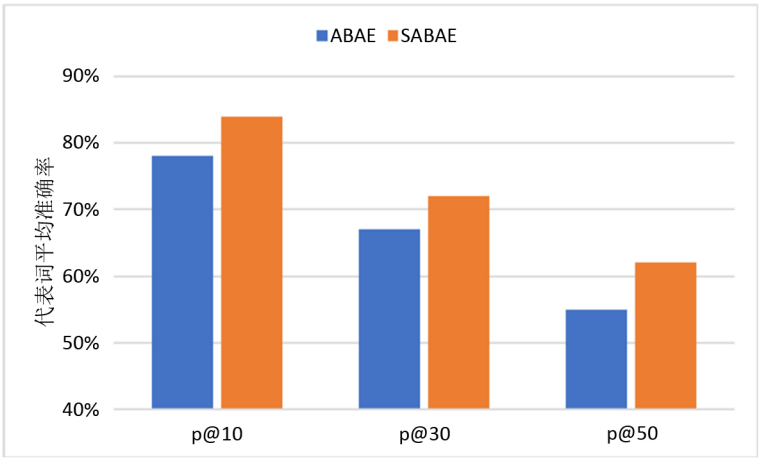

Figure 5. Average accuracy of representative words based on user evaluation
图 5. 基于用户评估的代表词平均准确率

4.4. 要素类情感分类实验结果

要素类情感分类采用准确率(accuracy)作为评价指标，准确率越高说明模型的分类精度越高。

消融实验是深度学习常用的一种控制变量实验，即通过单变量控制来验证某个条件或参数的改变对模型的实验结果的影响，最早出现在 Faster R-CNN 模型[25]的对比实验中。由于深度学习一般采用随机初始化的方法指定模型参数，固定随机种子使得实验运行时每次产生相同的随机数从而在一定程度上降低算法结果的随机性，因此，本文的消融实验固定了超参数和随机种子。出于了解基于上下文优化后的

要素类嵌入是否能提高要素类情感分类精度的目的，将本文提出的融合自注意力机制下的要素类特征的门控卷积神经网络模型与以下两种方法在 Restaurant 数据集上进行实验。

实验 1：GCAE 模型。尽管要素类词向量和文本词向量都由 word2vec 预训练的词向量进行初始化，但无论是在单要素类文本数据集上还是多要素类文本数据集上，其分类精度都优于将要素类词向量连接到文本词向量上的结合注意力机制的 LSTM (ATAE-LSTM)模型、标准的卷积神经网络(CNN)模型与不含要素类嵌入的门控卷机神经网络(GCN)模型[2]。

实验 2：GCAE_ATT 模型。输入门控卷积神经网络模型的要素类向量和文本词向量由结合注意力机制的神经网络模型提取出的要素类嵌入矩阵和优化后的文本词向量进行初始化。

实验 3：GCAE_SelfATT 模型。输入门控卷积神经网络模型的要素类向量和文本词向量由结合自注意力机制的神经网络模型提取出的要素类嵌入矩阵和优化后的文本词向量进行初始化。

表 3 展示了消融实验的结果，即 3 个模型分别在验证集和测试集上获得的分类准确率和耗费的训练时间。其中，验证集的数据即含有单个要素类及其情感极性的简单句子，也包含由多个要素类及其情感极性组成的复杂句子；测试集的数据则全部由复杂句子构成。

Table 3. Accuracy and training time of aspect category sentiment classification under the ablation experiment
表 3. 消融实验下的要素类情感分类准确率与训练时间

模型名称	验证集分类准确率	测试集分类准确率	训练时间
GCAE	79.4708%	81.1558%	569.0483s
GCAE_ATT	80.8850%	83.4171%	598.7766s
GCAE_SelfATT	**82.2080%**	**85.0503%**	**581.7217s**
提升幅度	2.74%	3.89%	/

实验结果说明，GCAE_SelfATT 模型在验证集和测试集上的分类精度均最高，其次是 GCAE_ATT 模型，最后是 GCAE 模型。要素类情感分类的精度高低与对应的要素类提取模型得到的代表词可视化聚类效果呈现正相关，也就是说要素类代表词的聚类效果越好，则将优化后的要素类向量和文本词向量对 GCAE 模型的输入进行初始化，得到的要素类情感分类精度越高。

5. 总结

针对要素类情感分类任务中，大多数方法没有对要素类信息和情感信息进行有效的结合，本文构建了融合自注意力机制下的要素类特征的门控卷积神经网络模型，通过引入自注意力机制获得基于上下文优化后的要素类向量表示，然后将优化后的要素类向量和文本词向量通过门控卷积神经网络模型进行训练，从而进行特性融合。当给出文本中不同的主题时，该模型能够获取到文本中的不同信息。通过在 SemEval 数据集上的消融实验表明，本文提出的模型所提取的要素类更一致、可解释性更强，对要素类进行情感分类的准确率更高。

本文以流水线的形式先实现要素类提取，再对提取的要素类进行情感极性预测，没有完全利用两个子任务的联合信息。因此，本文的下一步研究重点是多任务学习，同步提取要素类和进行要素类的情感分类。

基金项目

本文系上海市自然科学基金项目“基于自适应的多阶段差分进化计算研究”(项目编号：15ZR1401600)研究成果之一。

作者贡献

张颖：数据获取和论文撰写；郑建国：提出研究方向和方法。

参考文献

[1] Xue, W. and Li, T. (2018) Aspect Based Sentiment Analysis with Gated Convolutional Networks. *Proceedings of the* 56*th Annual Meeting of the Association for Computational Linguistics*, **1**, 2514-2523. https://doi.org/10.18653/v1/P18-1234

[2] 何炎祥, 孙松涛, 牛菲菲, 等. 用于微博情感分析的一种情感语义增强的深度学习模型[J]. 计算机学报, 2017, 40(4): 773-790.

[3] Zhang, L., Wang, S. and Liu, B. (2018) Deep Learning for Sentiment Analysis: A Survey. *WIREs Data Mining and Knowledge Discovery*, **8**, e1253. https://doi.org/10.1002/widm.1253

[4] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. https://doi.org/10.1038/nature14539

[5] Socher, R., Perelygin, A., Wu, J., *et al.* (2013) Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank. Empirical Methods in Natural Language Processing, 1631-1642.

[6] Kim, Y. (2014) Convolutional Neural Networks for Sentence Classification. *Proceedings of the* 2014 *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), Doha, October 2014, 1746-1751. https://doi.org/10.3115/v1/D14-1181

[7] Wang, X., Liu, Y., Sun, C., *et al.* (2015) Predicting Polarities of Tweets by Composing Word Embeddings with Long Short-Term Memory. *Proceedings of the* 53*rd Annual Meeting of the Association for Computational Linguistics and the* 7*th International Joint Conference on Natural Language Processing*, **1**, 1343-1353. https://doi.org/10.3115/v1/P15-1130

[8] Mnih, V., Heess, N., Graves, A., *et al.* (2014) Recurrent Models of Visual Attention. arXiv: Learning.

[9] Zhang, L. and Liu, B. (2014) Aspect and Entity Extraction for Opinion Mining. In: Chu, W., Ed., *Data Mining and Knowledge Discovery for Big Data*. *Studies in Big Data*, Vol. 1. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40837-3_1

[10] Mimno, D., Wallach, H., Talley, E.M., *et al.* (2011) Optimizing Semantic Coherence in Topic Models. *Proceedings of the* 2011 *Conference on Empirical Methods in Natural Language Proceedings*, Edinburgh, 27-31 July, 262-272.

[11] Yan, X., Guo, J., Lan, Y., *et al.* (2013) A Biterm Topic Model for Short Texts. *Proceedings of the* 22*nd International Conference on World Wide Web*, May 2013, 1445-1456. https://doi.org/10.1145/2488388.2488514

[12] Wang, L., Liu, K., Cao, Z., *et al.* (2015) Sentiment-Aspect Extraction Based on Restricted Boltzmann Machines. *Proceedings of the* 53*rd Annual Meeting of the Association for Computational Linguistics and the* 7*th International Joint Conference on Natural Language Processing*, Beijing, 26-31 July, 616-625.

[13] He, R., Lee, W.S., Ng, H.T., *et al.* (2017) An Unsupervised Neural Attention Model for Aspect Extraction. *Proceedings of the* 55*th Annual Meeting of the Association for Computational Linguistics*, **1**, 388-397. https://doi.org/10.18653/v1/P17-1036

[14] Wilson, T., Wiebe, J. and Hoffmann, P. (2015) Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. *International Journal of Computer Applications*, **7**, 347-354.

[15] Wu, Y., Wang, M. and Jin, P. (2008) Disambiguating Sentiment Ambiguous Adjectives. *International Conference on Natural Language Processing and Knowledge Engineering*, Beijing, 19-22 October 2008, 1-8. https://doi.org/10.1109/NLPKE.2008.4906816

[16] Tang, D., Qin, B., Feng, X., *et al.* (2015) Effective LSTMs for Target-Dependent Sentiment Classification. *Proceedings of COLING* 2016, *the* 26*th International Conference on Computational Linguistics*: *Technical Paper*, Osaka, 11-17 December 2017, 3298-3307.

[17] Wang, Y., Huang, M., Zhu, X., *et al.* (2016) Attention-Based LSTM for Aspect-Level Sentiment Classification. *Proceedings of the* 2016 *Conference on Empirical Methods in Natural Language Processing*, Austin, November 2016, 606-615. https://doi.org/10.18653/v1/D16-1058

[18] Liu, J. and Zhang, Y. (2017) Attention Modeling for Targeted Sentiment. *Proceedings of the* 15*th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, April 2017, 572-577. https://doi.org/10.18653/v1/E17-2091

[19] Tang, D., Qin, B. and Liu, T. (2016) Aspect Level Sentiment Classification with Deep Memory Network. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, November 2016, 214-224.

https://doi.org/10.18653/v1/D16-1021

[20] Ma, D., Li, S., Zhang, X., *et al.* (2017) Interactive Attention Networks for Aspect-Level Sentiment Classification. *Proceedings of the* 26*th International Joint Conference on Artificial Intelligence*, 4068-4074. https://doi.org/10.24963/ijcai.2017/568

[21] Chen, P., Sun, Z., Bing, L., *et al.* (2017) Recurrent Attention Network on Memory for Aspect Sentiment Analysis. *Proceedings of the* 2017 *Conference on Empirical Methods in Natural Language Processing*, Copenhagen, September 2017, 452-461. https://doi.org/10.18653/v1/D17-1047

[22] Li, X., Bing, L., Lam, W., *et al.* (2018) Transformation Networks for Target-Oriented Sentiment Classification. *Proceedings of the* 56*th Annual Meeting of the Association for Computational Linguistics*, **1**, 946-956. https://doi.org/10.18653/v1/P18-1087

[23] Tang, J., Lu, Z., Su, J., *et al.* (2019) Progressive Self-Supervised Attention Learning for Aspect-Level Sentiment Analysis. *Proceedings of the* 57*th Annual Meeting of the Association for Computational Linguistics*, Florence, July 2019, 557-566. https://doi.org/10.18653/v1/P19-1053

[24] Chen, Z., Mukherjee, A., Liu, B., *et al.* (2014) Aspect Extraction with Automated Prior Knowledge Learning. *Proceedings of the* 52*nd Annual Meeting of the Association for Computational Linguistics*, **1**, 347-358. https://doi.org/10.3115/v1/P14-1033

[25] Ren, S., He, K., Girshick, R., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Neural Information Processing Systems*, 91-99.