

文章编号:1008-1542(2020)06-0518-10

## 文本方面级情感分类方法综述

李胜旺, 杨 艺, 许云峰, 张 妍

(河北科技大学信息科学与工程学院, 河北石家庄 050018)

**摘 要:**随着深度学习的发展, 方面级情感分类已经在单领域和单一语言中取得了大量的研究成果, 但是在多领域的研究还有提升的空间。通过对近年来文本方面级情感分类方法进行归纳总结, 介绍了情感分类的具体应用场景, 整理了方面级情感分类常用的数据集, 并对方面级情感分类的发展进行了总结与展望, 提出未来可在以下领域开展深入研究: 1) 探索基于图神经网络的方法, 弥补深度学习方法存在的局限性; 2) 学习融合多模态数据, 丰富单一文本的情感信息; 3) 开展更多针对多语言文本和低资源语言的研究。

**关键词:**自然语言处理; 情感分类; 方面级别; 文本分类; 深度学习; 图神经网络; 图卷积网络

中图分类号: TP311.13 文献标识码: A doi: 10.7535/hbkd.2020yx06006

## A survey of text aspect-based sentiment classification

LI Shengwang, YANG Yi, XU Yunfeng, ZHANG Yan

(School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China)

**Abstract:** With the development of deep learning, aspect-based sentiment classification has achieved a lot of results in a single field and a single language, but there is room for improvement in multi-fields. By summarizing up the methods of text aspect-based sentiment classification in recent years, the specific application scenarios of sentiment classification were introduced, and the commonly used data sets of aspect-based sentiment classification were categorized. The development of aspect-based sentiment classification were summarized and prospected, and further research can be carried out in the following areas: exploring methods based on graph neural networks to make up for the limitations of deep learning methods; learning to fuse multi-modal data to enrich the emotional information of a single text; developing more targeted research work on multilingual texts and low-resource languages.

**Keywords:** natural language processing; sentiment classification; aspect-based; text classification; deep learning; graph neural network; graph convolutional network

收稿日期: 2020-10-02; 修回日期: 2020-11-06; 责任编辑: 王淑霞

基金项目: 中国留学基金委地方合作项目(201808130283); 中国教育部人工智能协同育人项目(201801003011); 河北科技大学校立课题(82/1182108)

第一作者简介: 李胜旺(1963—), 男, 河北邯郸人, 教授, 硕士, 主要从事计算机控制技术方面的研究。

通讯作者: 许云峰副教授。E-mail: hbkd\_xyf@hebust.edu.cn

李胜旺, 杨艺, 许云峰, 等. 文本方面级情感分类方法综述[J]. 河北科技大学学报, 2020, 41(6): 518-527.

LI Shengwang, YANG Yi, XU Yunfeng, et al. A survey of text aspect-based sentiment classification[J]. Journal of Hebei University of Science and Technology, 2020, 41(6): 518-527.

近年来,情感分析成为自然语言处理(NLP)领域中最活跃的研究方向之一,在信息检索、数据挖掘、文本挖掘等方面得到广泛研究。文本情感分析是对带有情感色彩的主观性文本进行分析、处理、归纳和推理的过程<sup>[1]</sup>。随着社交媒体和大数据的快速发展,情感分析已经从计算机科学蔓延到其他学科领域,其巨大的发展前景和商业价值引发了整个社会的共同关注。情感分类是情感分析技术的核心,其任务是判断文本的情感取向,根据文本所表达的含义和情感信息将文本划分成积极或消极2种或多种类别。按照情感的粒度可以分为3种分类问题:二分类(积极和消极)、三分类(积极、消极和中立)和多分类。例如对推特(Twitter)评论进行“开心”、“悲伤”、“愤怒”、“惊讶”的四元情感分类就属于多元分类问题。按照对文本研究粒度的不同可以分为3种不同级别:篇章级情感分类、句子级情感分类和方面级情感分类<sup>[2]</sup>,如图1所示。

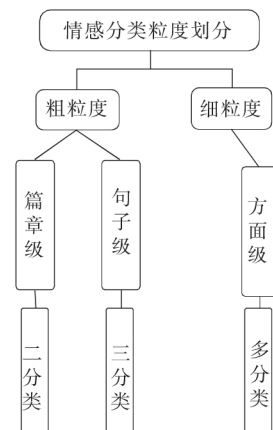


图1 情感分类粒度划分

Fig.1 Sentiment classification granularity

由于情感分类是一种具有较大实用价值的分类技术,在一些现实场景中有着广泛的应用,例如产品推荐、电影票房预测等,一直以来都受到研究人员的重点关注。现如今,生活中的方方面面变得愈加智能化,例如语音助手、智能客服等。在智能化道路上,机器的自然语言理解能力对智能化发展起着至关重要的作用。同时,对文本情感倾向的准确判断是更加准确地理解自然语言文本的重要环节。在3种不同级别的情感分类中,方面级情感分类任务需要考虑方面词和情感之间的交互,如果将其分离开,即在判断情感时不考虑方面信息,可能会导致分类不正确,因为同一个形容词对不同方面或在不同领域形容不同方面时,表达的情感是不同的。方面级情感分类可以更加细致地判断句子在特定方面的情感极性,对文本情感的判断更加准确,从而使得方面级情感分类成为当前情感分析领域的重点研究方向,在文本情感分类中起着举足轻重的作用。

对于研究领域中主流的情感分类方法,可按照技术路线将其分为2大类:传统方法和深度学习方法。传统方法又分为基于词典的方法、基于机器学习的方法、基于弱标注信息的方法、词典与机器学习相混合的方法<sup>[3]</sup>。传统分类方法主要针对篇章级和句子级进行粗粒度的情感分类,只能假定文本只有单一的情感(例如正向和负向),不能对文本的多个方面(即用户从哪个角度评论,或者商品从哪个角度介绍,例如价格、性能、服务等)进行情感识别。

近年来由于神经网络的快速发展,基于深度学习的情感分析取得了良好的效果,该方法的最大优势是不依赖人工定义特征,可以进行端到端的自主学习。本文主要整理了当前比较热门的传统深度学习情感分类方法和基于图域的深度学习方法即图神经网络方法,如图2所示。

目前对包括图神经网络方法的情感分类综述数量有限。申影利等<sup>[4]</sup>在方面级情感分类综述中只整合了基于传统深度学习方法的相關研究,缺乏基于图神经网络方法的最新研究。图卷积网络GCN<sup>[5]</sup>是一种简单有效的基于图的卷积神经网络,可以从丰富的关系数据中捕捉相互依赖的信息,在当前基于图神经网络的情感分类方法研究中,大多数研究人员应用了经典的图卷积网络GCN提高分类模型的算法性能,表现出很大潜力。

## 1 定义和概念

**定义1** 方面级情感分类(aspect-based sentiment classification,简称ABSC)任务定义:给定一个由 $n$ 个词和方面词 $w_i$ 组成的句子 $s = \{w_1, w_2, \dots, w_i, \dots, w_n\}$ ,方面级情感分类旨在确定句子 $s$ 针对方面词 $w_i$ 的情感极性。简单来说,其任务是给定一句话和目标提及,预测出关于目标方面的句子的情感极性。如图3所示,在“The waiter is friendly while the pizza is very ordinary”这一句评论中,用户分别在“pizza”和

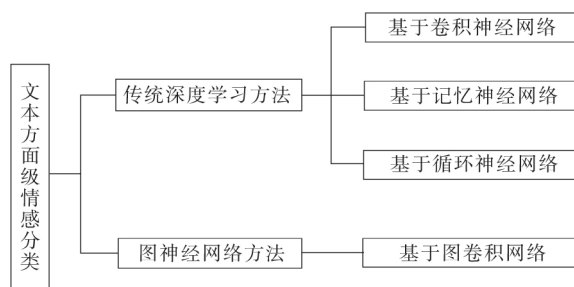


图2 文本方面级别情感分类方法

Fig.2 Text aspect-based sentiment classification methods

“waiter”2 个方面词项表达了负向和正向情感。同时, pizza 属于 food, waiter 对应 service, 所以在“food”和“service”2 个方面类型也表达了负向和正向情感。

**定义 2** 方面级情感分类的评估指标: Accuracy 和 Macro-F1。Accuracy 指正确预测样本数占总预测样本数的比值, 不考虑预测样本是正例还是负例, 表示对整个样本的判定能力。Macro-F1 也叫宏观 F1 分数, 用于多分类问题, 不受数据不平衡影响, 容易受到识别性高(高 recall 和高 precision)的类别影响。

**概念 1** 图(Graph, 如图 4 a)) 是一种对一组对象(Node)及其关系(Edge)进行建模的数据结构。由于图结构的强大表示能力, 近年来对图形的分析方法研究越来越受到关注, 例如社会科学中的社交网络、自然科学中的蛋白质交互网络以及知识图谱, 都可以用图结构进行表示。

**概念 2** 图神经网络(GNNs)是一种基于图域的深度学习方法。卷积神经网络(CNN)和循环神经网络(RNN)只能用于规则的欧式空间(如图 4 b)), 而 GNNs 可用来处理不规则的非欧式图数据, 通过图形节点之间的消息传递捕获图形的依赖性<sup>[6]</sup>。由于图神经网络的可推理性和高可解释性, 该方法已经被广泛应用于图分析当中。

## 2 情感分类方法介绍

传统的情感分析不仅耗费人力, 而且随着大数据时代的发展, 已经不能满足情感分类任务高效率和高质量的要求。一个文本序列通过 Word2Vec<sup>[7]</sup>等词嵌入(word embedding)模型转化为词向量(word vector)序列作为模型的输入, 使特征具有语义信息, 深度学习构建网络模型模拟人脑神经系统对文本进行逐步分析、特征抽取并且自动学习优化模型输出, 以提高分类的准确性。本文围绕传统深度学习方法和图神经网络方法 2 大类来介绍相关研究工作, 并重点介绍图神经网络方法的算法思想。

### 2.1 传统深度学习方法

#### 2.1.1 基于卷积神经网络的方法

卷积神经网络(CNN)<sup>[8]</sup>通过卷积和池化操作从原始数据中提取出高级特征, 在计算机视觉领域的发展中起到了很大作用, 并且不需要消耗太多的计算力就能捕获丰富的语义信息, 其有效性已经在许多领域得到验证。

HUANG 等<sup>[9]</sup>利用参数化的滤波器门控结构将方面词信息引入 CNN 中, 提出了卷积神经网络参数化滤波器(PF-CNN)和参数化门控卷积神经网络(PG-CNN)2 个模型变量。

FAN 等<sup>[10]</sup>提出了一种卷积记忆网络(CMN), 该模型融合了注意力机制, 可以同时捕捉句子中的单词和多词表达。为了帮助 CNN 特征抽取器更准确地定位情感指标, LI 等<sup>[11]</sup>在重新审视了注意力机制的缺陷和阻碍 CNN 在分类任务中发挥作用的障碍后, 提出了一种新的分类模型 TNet, 该模型采用了一种接近策略, 利用单词和方面词语之间的位置相关性来缩放卷积层的输入。

以前大部分工作使用了长短期记忆网络(LSTM)和注意力机制来预测有关目标的情感极性。XUE 等<sup>[12]</sup>认为该类模型很复杂并且在训练中需要花费很长的时间, 在此基础上, 提出了一个更有效的基于卷积神经网络和门机制(gating mechanisms)模型(GCAE)。该模型的计算在训练中很容易并行化, 并且门单元也能够独立工作。

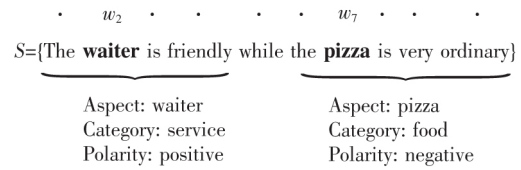
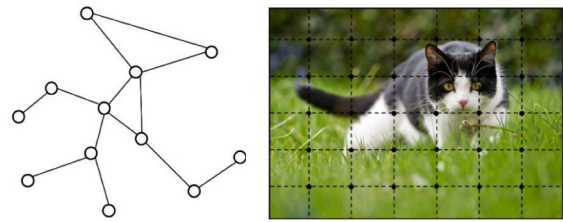


图 3 文本方面级情感分类任务示例

Fig.3 Example of text aspect-based sentiment classification task



a) 非欧式图

b) 欧式图

图 4 非欧式图和欧式图<sup>[6]</sup>

Fig.4 Graph in non-Euclidean space and image in Euclidean space<sup>[6]</sup>

以上研究结果表明,在方面级情感分类任务上,各种卷积神经网络变体比基本的卷积神经网络方法的分类性能更好,并且优于传统的机器学习方法。

### 2.1.2 基于记忆神经网络的方法

TANG 等<sup>[13]</sup>将深度记忆网络(DMN)应用到方面级情感分类任务中,该网络采用带有记忆网络存储的注意力机制捕获每个上下文词对于给定目标方面的重要性,性能大大优于基于 LSTM 和注意力 LSTM 的体系结构。

LI 等<sup>[14]</sup>将目标识别任务整合到情感分类任务中,以更好地模拟方面情感交互。在模型架构中,2 个子任务由一个深度记忆网络交错,其目标检测中产生的信号为极性分类提供线索,而预测的极性为目标的识别提供反馈。TAY 等<sup>[15]</sup>设计了一个二元记忆网络(DyMemN),该网络通过使用神经张量合成或全息合成进行记忆选择操作模拟方面和上下文之间的交互作用。

CHEN 等<sup>[16]</sup>提出了一种基于记忆网络的递归注意力机制(RAM),提取被长距离分隔的情感信息。该模型采用 Bi-LSTM(双向 LSTM)从输入中产生记忆,根据记忆切片与目标的相对位置对记忆切片进行加权,使得来自同一句子的不同目标有了自己定制的记忆。然后,对位置加权记忆进行多重关注,并将注意结果与一个递归网络 GRU<sup>[17]</sup>非线性结合。最后,将 softmax 应用于 GRU 网络的输出,对目标进行情感预测。

目前流行的方面级情感分类的内容注意力机制主要集中于识别情感词或移词,而没有考虑此类单词与句子中给定方面的相关性。为了解决此问题,LIU 等<sup>[18]</sup>提出一种由句子层面的内容注意力机制和上下文注意力机制 2 部分组成的方面级情感分类模型(CAM)。其中,句子层面的内容注意力机制能够从全局的角度捕捉到整个句子针对方面词语的重要信息,而上下文注意力机制则负责将单词的顺序及其相关性嵌入到一系列自定义中。

### 2.1.3 基于循环神经网络的方法

RNN 具有优秀的序列学习能力,许多方面级情感分类任务中的先进方法大多数都基于 RNN 进行建模。TANG 等<sup>[19]</sup>提出基于目标依赖的长短期记忆神经网络(TD-LSTM)和基于目标关联的长短期记忆神经网络(TC-LSTM)来扩展 LSTM。该模型把给定的目标视为一个特征,并将其与上下文特征连接起来进行方面级情感分类。ZHANG 等<sup>[20]</sup>利用门控神经网络结构,通过 Bi-RNN(双向 RNN)对句子中的句法和语义以及方面词与周围上下文词之间的相互作用进行建模,设计了 2 种门控神经网络模型(GRNN)。门控神经网络已经被证明可以通过更好地传播梯度来减少标准循环神经网络对序列末端的偏差。RUDER 等<sup>[21]</sup>提出基于方面情感分类的评论层次模型(H-LSTM),使用分层双向长短期记忆神经网络学习句内和句间的关系。模型将单词嵌入输入到句子级双向 LSTM 中,前向和后向 LSTM 的最终状态通过方面嵌入连接在一起,并反馈到双向评论级 LSTM 中。

注意力机制的核心目标是从众多信息中选择对当前任务目标更为关键的信息,利用注意力机制得到文本中关于某个方面类别的描述表示,从而进行更加准确的情感极性判别,与方面级情感分类的目标相一致。WANG 等<sup>[22]</sup>最早将注意力机制引入该任务,提出了一种使用目标嵌入的基于注意力的 LSTM 模型(ATAE-LSTM),在 LSTM 网络基础上为每个目标设置一个注意力向量,强制模型去注意句子中的重要部分,以响应特定的对象。该方法被证明是一种有效的强化神经模型来关注句子相关部分的方法。YANG 等<sup>[23]</sup>也提出了 2 种基于注意的双向 LSTM 来提高分类性能。LIU 等<sup>[24]</sup>通过区分给定目标的左上下文和右上下文获得的注意力来扩展注意力模型。HUANG 等<sup>[25]</sup>以 LSTM 为基础进行研究,用联合的方式为方面和句子建模,提出了注意力过度关注神经网络(AOA),通过 AOA 模块共同学习方面和句子的表示,并自动关注句子中的重要部分。其后,一些交互式注意力机制被应用于方面级情感分类任务上。MA 等<sup>[26]</sup>提出了一种既考虑目标注意力又考虑上下文注意力的交互式注意力网络(IAN),使用 2 个注意力网络交互式检测目标描述的重要单词和整个上下文的重要单词。

## 2.2 图神经网络方法

注意力机制和卷积神经网络因其在特定方面及上下文词语语义对齐的固有能力和被广泛应用于方面级情感分类。但是,此类模型缺乏解释相关句法约束和远距离单词依赖的机制,因此可能会错误地将与语法无关的上下文单词识别为判断方面情感的线索。为了解决这个问题,ZHANG 等<sup>[27]</sup>在句子的依赖树上构建图卷积网络(GCN),利用句法信息和单词依赖,提出一种特定方面的情感分类框架(ASGCN)。该模型从 Bi-LSTM 层开始,捕获有关单词顺序的上下文信息。为了获得方面特征,在 LSTM 输出的顶部加入多层图

卷积结构,随后用掩码机制过滤非特定方面的单词并保留高级特定方面的特征。特定方面的特征被反馈到 LSTM 输出,检索有关该方面的信息性特征,然后将其用于预测基于方面的情感。

现有的方面级情感分类方法大多是在一个句子中独立地对不同的方面进行建模,忽略了不同方面之间的情感依赖。然而,ZHAO 等<sup>[28]</sup>发现这种不同方面之间的依赖信息会带来额外的有价值的信息,并提出了一种基于注意力图卷积的方面级情感分类框架(SDGCN)。该工作将一个方面视为一个节点,一条边代表 2 个节点的情感依赖关系,是第一个在一句话中考虑方面之间情感依赖性的模型。

HOU 等<sup>[29]</sup>认为仅依赖树上的 GCN 模型容易出现解析错误,并且为了减轻由解析错误引起的问题,将自我注意序列模型与 GCN 模型相结合,设计了一个新的基于选择性注意的 GCN 模型(SA-GCN)。即在依赖树上的 GCN 模型基础上,使用多头注意力直接为句子中每个单词选择注意分数最高的  $k$  个单词。然后在生成的 top  $k$  注意图上应用另一个 GCN 模型整合上下文词的信息。

XIAO 等<sup>[30]</sup>指出,大多数方法不能完全捕捉上下文的语义信息,也缺乏解释相关句法约束和远程词汇依赖的机制。针对这些问题,提出了一种基于注意编码的图卷积网络模型(AEGCN),体系结构如图 5 所示。该模型主要由多头注意力和在句子依赖树上建立的改进图卷积网络组成。将预先训练好的 BERT<sup>[31]</sup>(一种从 transformer<sup>[32]</sup>模型得来的双向编码表征模型)应用于该任务,实现了更好的性能。

图 5 中,“embedding”表示 GloVe<sup>[33]</sup>嵌入或预训练的 BERT 嵌入;“hidden state”代表 Bi-LSTM;MHSA 指多头自注意力;MHIA 指多头互动注意力。首先,将 Bi-LSTM 用于上下文和特定目标的初步语义建模,在获得上下文的隐藏状态后,将 GCN 和 MHSA 组合在一起对句法信息进行编码。同时,利用 MHSA 对上下文的隐藏状态和特定目标进行注意编码,以获得更丰富的语义信息。然后,上下文语义编码和目标特定语义编码通过 MHIA 与句法信息编码相互作用,将平均池应用于交互式信息和上下文语义编码。最后,将它们合并以获得用于预测情感极性的最终特征表示,实验结果如表 1 所示。

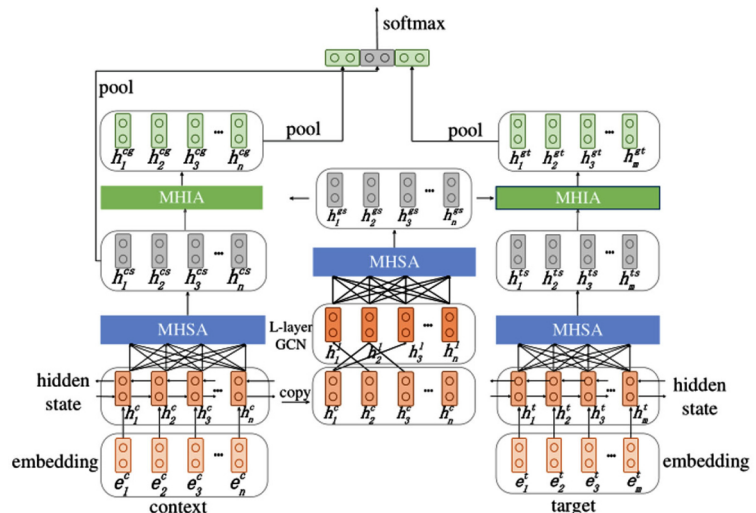


图 5 AEGCN 框架体系结构<sup>[30]</sup>

Fig.5 Framework of AEGCN<sup>[30]</sup>

表 1 AEGCN 实验结果<sup>[30]</sup>

Tab 1 Experimental results of AEGCN<sup>[30]</sup>

Model	TWITTER		LAP14		REST14		REST15		REST16	
	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1
SVM	63.40	63.30	70.49	N/A	80.16	N/A	N/A	N/A	N/A	N/A
LSTM	69.56	67.70	69.28	63.09	78.13	67.47	77.37	55.17	86.80	63.88
MemNet	71.48	69.90	70.64	65.17	79.61	69.64	77.31	58.28	85.44	65.99
AOA	72.30	70.20	72.62	67.52	79.97	70.42	78.17	57.02	87.50	66.21
IAN	72.50	70.81	72.05	67.38	79.26	70.09	78.54	52.65	84.74	55.21
TNet-LF	72.98	71.43	74.61	70.14	80.42	71.03	78.47	59.47	<b>89.07</b>	<b>70.43</b>
AEN	72.83	69.81	73.51	69.04	80.98	72.14	N/A	N/A	N/A	N/A
ASGCN-DT	71.53	69.68	74.14	69.24	80.86	<b>72.19</b>	79.34	60.78	88.69	66.64
ASGCN-DG	72.15	70.40	75.55	71.05	80.77	72.02	79.89	<b>61.89</b>	88.99	67.48
AEGCN-GloVe	<b>73.16</b>	<b>71.82</b>	<b>75.91</b>	<b>71.63</b>	<b>81.04</b>	71.32	<b>79.95</b>	60.87	87.39	68.22
AEGCN-BERT	<b>75.04</b>	<b>73.68</b>	<b>78.73</b>	<b>74.22</b>	<b>82.58</b>	<b>73.40</b>	<b>82.71</b>	<b>69.00</b>	<b>89.61</b>	<b>73.93</b>



实验选取了 8 个基线模型,在 5 个基准数据集上评估了 AEGCN 的有效性。通过实验结果中的粗体数据可以看出,采用图卷积网络 GCN 的模型,在大多数情况下表现出更好的效果;同时,基于 BERT 的模型比基于 GloVe 的模型表现出更大的优势。

### 2.3 方法归纳与比较

对以上研究工作进行总结归纳,详见表 2。

表 2 研究方法归纳与特点比较

Tab.2 Research summary and feature comparison

类别	研究者	算法模型	注意力机制	算法特点
卷积神经网络 (CNN)	HUANG 等 <sup>[9]</sup>	PF/PG-CNN	NO	利用参数化的滤波器门控结构将方面词信息引入 CNN。
	FAN 等 <sup>[10]</sup>	CMN	YES	融合了注意力机制,可以同时捕捉句子中的单词和多词表达。
	LI 等 <sup>[11]</sup>	TNet	NO	利用单词和方面词语之间的位置相关性缩放卷积层的输入。
	XUE 等 <sup>[12]</sup>	GCAE	NO	卷积神经网络和门机制模型的计算在训练中并行化,训练时间更少。
记忆神经网络 (MNN)	TANG 等 <sup>[13]</sup>	DMN	YES	采用带有记忆网络的注意力机制捕获每个上下文词对于给定目标的重要性。
	LI 等 <sup>[14]</sup>	AttNet	YES	将目标识别任务整合到情感分类任务中,更好地模拟方面情感交互。
	TAY 等 <sup>[15]</sup>	DyMemN	NO	使用神经张量合成进行记忆选择操作,模拟方面和上下文之间的交互作用。
	CHEN 等 <sup>[16]</sup>	RAM	YES	采用 Bi-LSTM 和注意力机制提取长距离分隔的情感信息。
循环神经网络 (RNN)	LIU 等 <sup>[18]</sup>	CAM	YES	句子级内容注意力机制和上下文注意力机制同时考虑单词的顺序及其相关性。
	TANG 等 <sup>[19]</sup>	TC/TD-LSTM	NO	给定的目标视为特征与上下文特征连接起来进行方面级情感分类。
	ZHANG 等 <sup>[20]</sup>	GRNN	NO	利用门控神经网络结构,通过 Bi-RNN 对方面词与上下文词的相互作用建模。
	RUDER 等 <sup>[21]</sup>	H-LSTM	NO	使用分层双向长短期记忆神经网络学习句内和句间的关系。
	WANG 等 <sup>[22]</sup>	ATAE-LSTM	YES	在 LSTM 基础上为每个目标设置一个注意力向量,以响应特定的对象。
	YANG 等 <sup>[23]</sup>	AB-LSTM	YES	基于注意的双向 LSTM 来提高分类性能。
	LIU 等 <sup>[24]</sup>	BILSTM-ATT	YES	通过区分给定目标的左上下文和右上下文获得的注意力来扩展注意力模型。
	HUANG 等 <sup>[25]</sup>	AOA	YES	用联合的方式为方面和句子建模,明确地捕捉方面和上下文句子之间的交互。
图卷积网络 (GCN)	MA 等 <sup>[26]</sup>	IAN	YES	使用 2 个注意力网络来交互式地检测目标和上下文描述的重要单词。
	ZHANG 等 <sup>[27]</sup>	ASGCN	NO	在句子的依赖树上构建图卷积网络(GCN),以利用句法信息和单词依赖。
	ZHAO 等 <sup>[28]</sup>	SDGCN	YES	在位置编码的双向注意力机制上,用 GCN 捕捉句子中不同方面之间的情感依赖。
	HOU 等 <sup>[29]</sup>	SA-GCN	YES	在依赖树上的 GCN 模型的基础上,使用多头注意力为句子中每个单词选择注意。
	XIAO 等 <sup>[30]</sup>	AEGCN	YES	由多头注意力和在句子依赖树上建立的改进图卷积网络组成。

由表 2 可以看出,已经有大量基于传统深度学习的方面级情感分类方法。大部分研究工作针对需要解决的问题和模型的特性,将不同的模型和机制相互融合。但是,由于传统深度学习本身所具有的不可推理和不可解释的局限性,导致基于传统深度学习的方法进入缓慢发展时期。同时,业界研究人员认为,图神经网络的可推理和高可解释的特性,恰好弥补了传统深度学习的局限,从而在近两年变得愈加活跃。

通过图神经网络方法研究给出的部分实验结果可知,利用图卷积网络 GCN 捕获远程单词之间的依赖关系比其他一些深度学习方法更胜一筹,表现出很大的潜力。因此,在今后的研究中,会有更多的图神经网络方法逐渐登上舞台,为方面级情感分类做出贡献。尽管如此,图神经网络本身仍存在着一些有待解决的问题,例如叠加多层图卷积反而会使训练效果变差,处理大规模数据会使算法复杂度成指数增长等。方面级情感分类在近几年发展迅速,但是在深度学习领域的发展仍面临着巨大的挑战,且需要一个漫长的过程。

## 3 常用数据集与应用概述

### 3.1 常用数据集

方面级别情感分析(aspect-based sentiment analysis,简称 ABSA)领域最常用的是 SemEval 数据集,其基本任务是对 Twitter 文本进行情感分析。SemEval 数据集的官方链接如表 3 所示。

表 3 SemEval 数据集链接  
Tab.3 SemEval dataset official link

Dataset	official link
SemEval-2014 Task 4 <sup>[34]</sup>	<a href="http://alt.qcri.org/semEval2014/task4/">http://alt.qcri.org/semEval2014/task4/</a>
SemEval-2015 Task 12 <sup>[35]</sup>	<a href="http://alt.qcri.org/semEval2015/task12/">http://alt.qcri.org/semEval2015/task12/</a>
SemEval-2016 Task 5 <sup>[36]</sup>	<a href="http://alt.qcri.org/semEval2016/task5/">http://alt.qcri.org/semEval2016/task5/</a>
SemEval-2017 Task 4 <sup>[37]</sup>	<a href="http://alt.qcri.org/semEval2017/task4/">http://alt.qcri.org/semEval2017/task4/</a>

### 3.1.1 SemEval-2014 Task 4 数据集的子任务<sup>[34]</sup>

- 1) 方面词抽取(aspect term extraction)。给定针对某个实体(比如餐厅)的一些句子,识别其中的方面词。
- 2) 方面词(aspect term)的极性分类。给定一个句子和这个句子里的所有方面词,判定每一个方面词的情感极性。可能的极性包括正面(positive)、负面(negative)、中性(neutral)和冲突(conflict)。
- 3) 方面类别(aspect category)识别。给定一个句子,识别出其中的类别(一个句子可能包含多个类别)。
- 4) 方面类别的情感分类。给定一个句子以及句子里的一个或者多个方面类别,输出每个类别的情感分类。和子任务 2 的分类一样,这里的分类也分为正面、负面、中性和冲突。

### 3.1.2 SemEval-2015 Task 12 数据集的子任务<sup>[35]</sup>

- 1) 域内(in-domain)任务。给定一个完整的评论,需要完成以下 3 个子任务。
  - ① 方面类别识别。识别评论里所有的实体(E)和属性(A)对。E 和 A 都是预定义集合中的某一个值。
  - ② 想法意图表示(opinion target expression,简称 OTE)识别。只有餐馆数据集上有标注数据,OTE 任务的输入是所有的实体和属性对,需要识别实体和属性对里实体 E 对应的字符串。
  - ③ 情感分类。给定一个有上下文的句子和所有的实体和属性对,对其情感进行分类,可能的分类为正面、负面和中性。
- 2) 域外(out-of-domain)任务。增加了一个酒店的测试数据集(没有训练数据),然后考察模型在不同领域的泛化能力。

### 3.1.3 SemEval-2016 Task 5 数据集的子任务<sup>[36]</sup>

- 1) 句子级别的 ABSA。给定某个实体(餐馆或酒店)的一篇评论中的一个句子,需要确定所有观点的以下 3 方面内容。
  - ① 方面类别检测(aspect category detection)。确定文本里出现的所有实体和属性对,其中实体 E 来自预定义的实体类列表,属性 A 来自预定义的属性标签列表。
  - ② 想法意图表示(OTE)。只有餐馆数据集有该任务,需要确定每个实体和属性对里实体所对应的字符串的开始下标和结束下标,如果找不到则输出" NULL"。
  - ③ 情感极性。给定一个有上下文的句子和所有的实体和属性对,对其情感进行分类,可能的分类为正面、负面和中性。
- 2) 篇章级别的 ABSA。句子级别的问题是模型不能参考上下文,因此还有一个篇章级别的 ABSA 任务。它的任务和句子级别的任务是一样的,输入的是整个评论文本。
- 3) 域外(out-of-domain)ABSA。通过无训练数据的测试数据集考察模型在不同领域的泛化能力。

### 3.1.4 SemEval-2017 Task 4 数据集的子任务<sup>[37]</sup>

- 1) 对 Twitter 文本进行情感分类,分为积极、消极和中立。
- 2) 给定一个 Twitter 文本和一个 topic,判断 Twitter 文本内容关于这个 topic 的情感倾向,分为积极和消极。
- 3) 在子任务 2 的基础上,更加精细地分类,分为非常积极、弱倾向积极、中立、弱倾向消极、非常消极 5 个程度。
- 4) 给定一个 topic 和一组 Twitter 文本,估计文本在积极和消极 2 个情感程度上的分布。
- 5) 给定一个 topic 和一组 Twitter 文本,估计文本在积极、消极、非常积极、弱倾向积极、中立、弱倾向消极、非常消极 5 个情感程度的分布。

### 3.2 实际应用概述

情感分类是一种具有很大实际应用价值的文本分类技术<sup>[38]</sup>,笔者从个人角度的商品评价和企业角度的民意测评来简单阐述情感分类的应用。

#### 3.2.1 商品评价

当个人试图去做某个决定并转化为行为时,通常会考虑他人的意见、建议、观点及看法。或者说,他人的意见和看法会对一个人的决定和行为产生显著影响。例如,当某人在购买一款商品时,就会主动考虑朋友、同事等对这个商品的意见。在大数据时代,更多的还会从互联网上搜集与此商品有关的热点评价或各种形式的开放信息,根据这些信息来决定是否购买商品<sup>[39]</sup>。通过情感分类技术将商品评价分类,获取可以让人们更好地判断商品的价值信息。

#### 3.2.2 民意测评

一个企业如果想要了解自己的产品或服务在大众心中的品质形象和评价等信息,过去只能通过面对面交流或采取调查问卷的方式获取。在如今的大数据时代,可以直接从互联网上获取大量的相关信息。对于更加高效地了解大众的情感导向,情感分类提供了关键的技术支持。

## 4 研究与展望

深度学习已经在情感分类领域做出很大贡献。利用深度学习模型间相互结合并加入一定的机制,可以更精确地识别句子的语义表达,从而判断出细节以及整个文档的情感极性,展现出对不同类型情感分类的适应能力。本文介绍了各种深度学习体系结构应用在方面级情感分类中的研究,将其分为2大类并细分为4个类别,分别是基于卷积神经网络、基于记忆神经网络、基于循环神经网络以及基于图卷积网络的方法。

在情感分类研究领域,方面级情感分类发展迅速,在单领域和单一语言上不断取得成果,但在一些方面仍面临挑战。例如由于语言的多样性与复杂性,不同领域间的情感表达差别较大。特别是当领域差异过大时,如将商品评论领域训练的情感分析模型应用到新闻领域上,情感分析的效率和性能则会下降。因此,训练高效的通用领域情感分析模型非常值得研究。同时,在近两年开始崭露头角的基于图神经网络的方法所展现出的巨大潜力,预示着在今后的研究中,图神经网络可能成为解决现有问题的一种有效途径,使得方面级情感分类得到进一步提升。

情感分析在技术上具有广阔的发展空间,在深度学习领域还有许多亟待解决的难点、痛点问题,今后应围绕以下问题开展进一步研究。

1) 多种不同的图神经网络模型与不同的机制相融合。在当前基于图神经网络方法的研究中,大多数都使用了经典的图卷积网络 GCN,然而,一些研究人员提出了同样优异甚至表现更好的 GCN 变体和一些其他图神经网络模型,例如 AM-GCN<sup>[40]</sup>,P-GNN<sup>[41]</sup>等。如何结合这些图神经网络模型的特性用于方面级情感分类任务中是一个值得探索的研究方向。

2) 对多语言文本和低资源语言的情感分析。目前由于低资源语言缺乏大量的权威标注数据集,并且单一语言的成果无法泛化到多语言环境,使得对多语言文本和低资源语言的研究工作较少。因此,针对多语言文本和低资源语言的情感分析将是一个新的研究方向。

3) 多模态数据的情感分析。多模态数据是指包含文本、视觉和听觉信息的多类型数据,可以为传统的文本特征提供额外的情感信息。由于深度学习模型可以将输入映射到某个潜在空间进行特征表示,因此也可以同时映射多模态数据的输入,来学习融合多模态数据提供的情感特征信息,例如使用特征串联、联合潜在空间或其他更复杂的融合方法。将多模态数据与深度学习技术结合,用多模态数据帮助情感分析是一个具有潜力的方向。

随着社交媒体数据的海量兴起,研究人员可以在大数据的基础上开展实验。期待在不久的将来,研究人员可以搭建出更加完善的情感分析系统,为自然语言处理领域的发展做贡献。

### 参考文献/References:

- [1] ZHANG L, WANG S, LIU B. Deep learning for sentiment analysis: A survey[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018, 8(4): 1253.



- [2] 洪巍,李敏.文本情感分析方法研究综述[J].计算机工程与科学,2019,41(4):750-757.  
HONG Wei,LI Min.A review:Text sentiment analysis methods[J].Computer Engineering and Science,2019,41(4):750-757.
- [3] 陈龙,管子玉,何金红,等.情感分类研究进展[J].计算机研究与发展,2017,54(6):1150-1170.  
CHEN Long,GUAN Ziyu,HE Jinhong,et al.A survey on sentiment classification[J].Journal of Computer Research and Development,2017,54(6):1150-1170.
- [4] 申影利,赵小兵.基于深度学习的方面级情感分析综述[J].信息技术与标准化,2020(1):50-53.  
SHEN Yingli,ZHAO Xiaobing.A review of research on aspect-based sentiment analysis based on deep learning[J].Information Technology & Standardization,2020(1):50-53.
- [5] KIPF T N,WELLING M.Semi-supervised Classification with Graph Convolutional Networks[EB/OL].[2017-02-22].<https://arxiv.org/pdf/1609.02907v4.pdf>.
- [6] ZHOU J,CUI G Q,ZHANG Z Y,et al.Graph Neural Networks:A Review of Methods and Applications[EB/OL].[2019-07-10].<https://arxiv.org/pdf/1812.08434.pdf>.
- [7] NGUYEN H T, NGUYEN L M. Effective attention networks for aspect-level sentiment classification[C]//2018 10th International Conference on Knowledge and Systems Engineering (KSE). Ho Chi Minh City: Institute of Electrical and Electronics Engineers, 2018: 25-30.
- [8] KRIZHEVSKY A,SUTSKEVER I,HINTON G E.ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017,60(6): 84-90.
- [9] HUANG B X,CARLEY K M.Parameterized convolutional neural networks for aspect level sentiment classification[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018: 1091-1096.
- [10] FAN C, GAO Q, DU J, et al. Convolution-based memory network for aspect-based sentiment analysis[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. New York: ACM, 2018: 1161-1164.
- [11] LI X, BING L, LAM W, et al. Transformation networks for target-oriented sentiment classification[C]//Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics, 2018:1805-1815.
- [12] 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 (Long Papers). Melbourne: Association for Computational Linguistics, 2018: 2514-2523.
- [13] TANG D Y, QIN B, LIU T. Aspect level sentiment classification with deep memory network[C]//Proceedings of the 2016 Conference on Empirical Methods. Austin: Association for Computational Linguistics, 2016:214-224.
- [14] LI C, GUO X, MEI Q. Deep memory networks for attitude identification[C]//Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. New York: ACM, 2017: 671-680.
- [15] TAY Y, TUAN L A, HUI S C. Dyadic memory networks for aspect-based sentiment analysis[C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. New York: ACM, 2017: 107-116.
- [16] CHEN P, SUN Z, BING L, et al. Recurrent attention network on memory for aspect sentiment analysis[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen: Association for Computational Linguistics, 2017: 452-461.
- [17] CHUNG J Y, GULCEHRE C, CHO K H, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling[EB/OL].[2014-12-11].<https://arxiv.org/pdf/1412.3555v1.pdf>.
- [18] WU X, HE R, SUN Z N, et al. A light CNN for deep face representation with noisy labels[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(11):2884-2896.
- [19] TANG D Y, QIN B, FENG X C, et al. Effective LSTMs for target-dependent sentiment classification[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: Association for Computational Linguistics, 2016:606-615.
- [20] ZHANG M, ZHANG Y, VO D T. Gated neural networks for targeted sentiment analysis[C]//Thirtieth AAAI Conference on Artificial Intelligence. Phoenix: AAAI Press, 2016: 3087-3093.
- [21] RUDER S,GHAFFARI P,BRESLIN J G.A hierarchical model of reviews for aspect-based sentiment analysis[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: Association for Computational Linguistics, 2016: 999-1005.
- [22] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: Association for Computational Linguistics, 2016: 606-615.
- [23] YANG M, TU W, WANG J, et al. Attention-based LSTM for target-dependent sentiment classification[C]//Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence. San Francisco: AAAI Press, 2017: 5013-5014.
- [24] LIU J, ZHANG Y. Attention modeling for targeted sentiment[C]//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers. Valencia: Association for Computational Linguistics, 2017: 572-577.
- [25] HUANG B, OU Y, CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks[C]//International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and

- Simulation. Washington: Springer, 2018: 197-206.
- [26] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification[C]//Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne: International Joint Conferences on Artificial Intelligence, 2017: 4068-4074.
- [27] ZHANG C, LI Q C, SONG D W. 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 (EMNIP-IJNLP). Hongkong Association for Computational Linguistics, 2019: 4568-4578.
- [28] ZHAO P L, HOU L L, WU O. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification[J]. Knowledge-Based Systems, 2020, 193: 105443.
- [29] HOU X C, HUANG J, WANG G T, et al. Selective Attention Based Graph Convolutional Networks for Aspect-level Sentiment Classification[EB/OL]. [2020-04-07]. <https://arxiv.org/abs/1910.10857.pdf>.
- [30] XIAO L W, HU X H, CHEN Y N, et al. Targeted sentiment classification based on attentional encoding and graph convolutional networks[J]. Applied Sciences, 2020, 10(3): 957.
- [31] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[EB/OL]. [2019-05-24]. <https://arxiv.org/pdf/1810.04805.pdf>.
- [32] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. Long Beach: NIPS, 2017: 5998-6008.
- [33] 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). Doha: Association for Computational Linguistics, 2014: 1532-1543.
- [34] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. Semeval-2014 task 4: Aspect based sentiment analysis[C]//Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). Dublin: Association for Computational Linguistics, 2014: 27-35.
- [35] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. Semeval-2015 task 12: Aspect based sentiment analysis[C]//Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval-2015). Denver: Association for Computational Linguistics, 2015: 486-495.
- [36] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. Semeval-2016 task 5: Aspect based sentiment analysis[C]//Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016). San Diego: Association for Computational Linguistics, 2016: 19-30.
- [37] ROSENTHAL S, FARRA N, NAKOV P. SemEval-2017 task 4: Sentiment analysis in Twitter[C]//Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017). Vancouver: Association for Computational Linguistics, 2017: 502-518.
- [38] LI Q, PENG H, LI J X, et al. A survey on text classification: From shallow to deep learning[J]. Association for Computing Machinery, 2020, 37(4): 35-69.
- [39] 江红. 情感分析研究综述[J]. 智能计算机与应用, 2018, 8(5): 103-105.  
JIANG Hong. Research on sentiment analysis[J]. Intelligent Computer and Applications, 2018, 8(5): 103-105.
- [40] WANG X, ZHU M, BO D, et al. AM-GCN: Adaptive multi-channel graph convolutional networks[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 1243-1253.
- [41] YOU J X, YING R, LESKOVEC J. Position-aware graph neural networks[C]// Proceedings of the 36th International Conference on Machine Learning. Long Beach: Computer Science, 2019: 1-10.