

文章编号: 1003-0077(2021)08-0098-09

基于图卷积记忆网络的方面级情感分类

王光¹, 李鸿宇^{1,2}, 邱云飞¹, 郁博文², 柳厅文²

(1. 辽宁工程技术大学 软件学院, 辽宁 葫芦岛 125105;

2. 中国科学院 信息工程研究所, 北京 100089)

摘要: 在方面级情感分类中,常用的方法是用卷积神经网络或循环神经网络提取特征,利用注意力权重获取序列中不同词汇的重要程度。但此类方法未能很好地利用文本的句法信息,导致模型不能准确地在评价词与方面词之间建立联系。该文提出一种图卷积神经记忆网络模型(MemGCN)来解决此依赖问题。首先通过记忆网络存储文本表示与辅助信息,然后利用基于依存句法树的图卷积神经网络获取文本的句法信息。最后,使用注意力机制融合句法信息与其他辅助信息。在 SemEval 2014 任务和 Twitter 数据集上的实验结果表明,MemGCN 显著提升了模型性能。

关键词: 句法信息;图卷积网络;注意力机制;辅助信息

中图分类号: TP391

文献标识码: A

Aspect-based Sentiment Classification via Memory Graph Convolutional Network

WANG Guang¹, LI Hongyu^{1,2}, QIU Yunfei¹, YU Bowen², LIU Tingwen²

(1. School of Software, Liaoning Technical University, Huludao, Liaoning 125105, China;

2. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100089, China)

Abstract: In aspect-based sentiment classification, the attention mechanism is often combined in recurrent neural network or convolutional neural network to obtain the importance of different words. However, such kind of methods fail to capture long-range syntactic relations that are obscure from the surface form, which would be beneficial to identify sentiment features directly related to the aspect target. In this paper, we propose a novel model named MemGCN to explicitly utilize the dependency relationship among words. Firstly, we employ the memory network to obtain the context-aware memory representation. After that, we apply graph convolutional network over the dependency tree to propagate sentiment features directly from the syntactic context of an aspect target. Finally, the attention mechanism is used to fuse memory and syntactic information. Experiment results on SemEval 2014 and Twitter datasets demonstrate our model outperforms baseline methods.

Keywords: syntactic information; graph convolutional network; attention mechanism; auxiliary information

0 引言

方面级情感分类是一种细粒度情感分析任务,目标是确定评论文本中所出现的不同方面的情感倾向^[1-2]。如图 1 中方面级情感分类实例图所示,已知“food”和“service”方面词,方面级情感分类模型需要通过文本评论句子和方面词“food”判断出评论者对食物的情感倾向是正向的,依据“service”方面词

推断出服务是不令人满意的。本文实例中方面词使用下划线标识,通过粗体词汇标识评价词,评论者对每个方面词的情感倾向取决于评价词与方面词的对应关系,由于评论者对不同方面词的情感态度存在差异,因此错误地匹配评价词与方面词可能会带来相反的判断。相对于篇章级别、句子级别和词汇级别情感分类任务而言,方面级情感分类的针对性更强,不但可以帮助消费者筛选出满足自身需求的商品,而且还能指导商家进行产品改进。

收稿日期: 2020-01-16 定稿日期: 2020-04-10

基金项目: 国家自然科学基金(71371091)



图1 方面级情感分类任务实例

虽然神经网络模型在方面级情感分类任务中已取得了良好的效果,但现有方法没有很好地利用句法信息来辅助分类。一般来说,方面级情感分类任务中的评价词往往出现在句法关系复杂的文本中,以文本评论“We didn't know if we should order a drink or leave ?”为例,在句法关系十分复杂的句子中,判断评论者对方面词“drink”的情感态度需要全局的语义理解和推断,而通过句法分析能够帮助模型找到“drink”相关的评价词,进而完成情感分类。受到文献[3-4]图卷积神经网络(GCN, Graph Convolutional Network)在文本分类领域成功应用的启发,本文提出了一种结合图神经网络和注意力机制的方面级情感分类方法 MemGCN,它能够有效利用句法依存信息辅助分类。在三个公开的方面级情感分类数据集上,证明了 MemGCN 模型的优越性能。

1 相关工作

方面级情感分类的研究工作基于统计机器学习方法开展。Kiritchenko 等^[5]利用情感词典和特征工程提取情感信息,训练支持向量机(Support Vector Machine, SVM)情感分类模型,在笔记本电脑和餐馆领域的方面级情感分析评测任务中取得当时最好的效果。

由于通过构建情感词典和手工提取特征的方法^[6],一方面需要投入高额的人工成本,另一方面机器学习算法很难突破性能瓶颈,无法学习到方面词与句子中上下文词汇之间的关系。随着神经网络算法的兴起,自动提取特征的卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN)等深度学习算法在方面级情感分类任务中显著提升了分类性能。Tang 等^[7]提出带有方面词依赖的 TD-LSTM 模型,利用两个不同方向的长短时记忆网络(Long Short Term Memory Network, LSTM)分别学习方面词上文信息和下文信息,提升了模型分类效果。

近年来,注意力机制(Attention Mechanism)在机器翻译、文本摘要与问答系统等自然语言处理领

域均得到了成功应用^[8-9]。注意力机制的变体在方面级情感分类任务中也发挥了重要作用。Wang 等^[10]将长短时记忆网络和注意力机制进行结合,提出了方面词嵌入方法,使得 ABAE-LSTM 模型可以关注到方面词所对应的情感信息。曾锋等^[11]采用层次注意力机制对词汇维度和句子维度分层建立模型,在由多个句子组成的文本评论中,分类效果获得了显著提高。

Xue 等^[12]提出的 GCAE 模型结合了卷积神经网络和门控机制,利用并行计算大幅加快了方面级情感分类模型的训练速度;梁斌和刘全等^[13-14]证明卷积神经网络与注意力机制和长短期记忆网络结合均能改善模型;杜慧等^[15]将词向量、词性信息和注意力同时融入到卷积神经网络和循环神经网络中,增强了模型的性能。此外, Tang 等^[16]设计记忆网络(Memory Network)在 SemEval-2014, 方面级情感分析任务中取得最佳成绩; Ma 等^[17]训练得到对方面词和上下文信息同时建模的 IAN 模型提升分类效果; Chen 等^[18]提出的 RAM 模型是对 MemNet 模型的改进; Zhu 等^[19]提出带有辅助记忆网络的方面级情感分类方法,学习方面词和情感词之间的信息,进一步优化了方面级情感分类模型。

2 基于图卷积记忆网络的方面级情感分类

MemGCN 模型由文本表示层、语义学习层、位置记忆层、图卷积神经网络层、注意力机制和情感分类层组成,整体框架如图 2 所示。为便于理解和描述,文中文本评论句子使用 $S = \{w_1, w_2, \dots, w_{t-1}, w_t, \dots, w_{t+l}, w_{t+l+1}, \dots, w_n\}$ 表示,其中, n 表示句子中词的数目,方面词为 w_t, \dots, w_{t+l} ,可能由一个或多个词构成。 t 和 l 分别表示词的位置和方面词的长度, w_1, w_2, \dots, w_{t-1} 与 w_{t+l+1}, \dots, w_n 分别指代上文词和下文词。

2.1 文本表示层

Word2Vec、FastText、GloVe 等^[8,20]无监督词表示方法广泛应用于自然语言处理的各项任务中,能够表示文本的语义信息。将词向量集合表示为 $T \in \mathbf{R}^{d \times |V|}$,其中, d 为词向量维度, $|V|$ 是词汇个数,句子 S 映射到低维、连续、稠密空间后对应 $V = \{v_1, v_2, \dots, v_{t-1}, v_t, \dots, v_{t+l}, v_{t+l+1}, \dots, v_n\}$ 。借助词嵌入方法思想,本文将方面词和评论文本中的词汇的词性映射到低维、连续、稠密的空间。文本表

示层的输出包括词向量、词性向量和方面词向量， $x_t = \{[v_1; \text{pos}(w_1); aw], [v_2; \text{pos}(w_2); aw], \dots, [v_{t-1}; \text{pos}(w_{t-1}); aw], [v_t; \text{pos}(w_t); aw], \dots,$

$[v_{t+l}; \text{pos}(w_{t+l}); aw], [v_{t+l+1}; \text{pos}(w_{t+l+1}); aw], \dots, [v_n; \text{pos}(w_n); aw]\}$ 。若同一条评论中不止有一个方面词被提及，则分开处理。

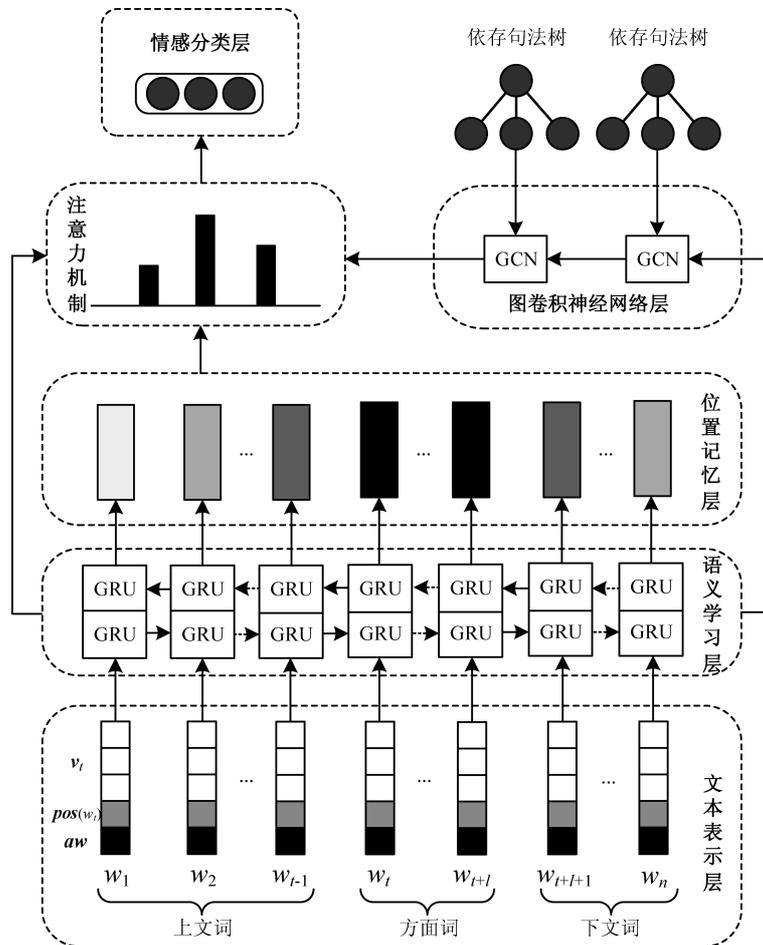


图2 MemGCN模型的整体框架

2.2 语义学习层

文本评论数据中词之间存在时间序列关系，循环神经网络能够学习到长距离的语义依赖信息。门控循环单元^[21](Gate Recurrent Unit, GRU)是一种循环神经网络，其结构如图3所示，门控机制能避免RNN梯度弥散问题，而且比LSTM拥有更少的参数量，所以训练模型的速度更快。

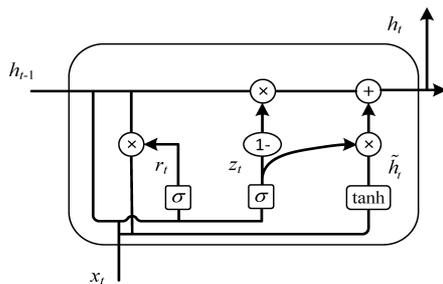


图3 门控循环单元的结构

图3中的 x_t 和初始化为 x_t 的 h_{t-1} 为门控循环单元的输入，分别表示 t 时刻文本表示层输出信息和上一时刻门控循环单元输出信息。更新门 z_t 和重置门 r_t 的计算方法如式(1)、式(2)所示。

$$z_t = \sigma(W_z [h_{t-1}, x_t] + b_z) \quad (1)$$

$$r_t = \sigma(W_r [h_{t-1}, x_t] + b_r) \quad (2)$$

其中， σ 是sigmoid函数， W_z 和 b_z 分别为更新门的权重和偏置，使用 W_r 和 b_r 分别代表重置门的权重和偏置。更新门的作用是限定上一时刻被保留到当前状态中的信息量，更新门的值越大表示上一时刻的状态信息留存越多。重置门用于控制上一时刻状态信息被写入到当前的候选集 \tilde{h}_t 上的程度，重置门的值越小说明前一状态的信息被写入得越少。

$$\tilde{h}_t = \tanh(W_h [r_t \times h_{t-1}, x_t] + b_h) \quad (3)$$

式(3)中 \tilde{h}_t 是将被添加到记忆模块的值。tanh

是双曲正切函数, W_h 与 b_h 分别为记忆模块候选集的权重和偏置。重置门信息辅助候选集决定上一时刻的记忆信息是否被写入当前状态。式(4)中的 h_t 即为 t 时刻循环门控单元的输出, 是综合了更新门和候选集信息计算得出的结果。

$$h_t = z_t \times \tilde{h}_t + (1 - z_t) \times h_{t-1} \quad (4)$$

Tom 等^[6]证明双向循环神经网络能够捕获文本序列正反双向的语义信息, 于是本文使用双向门控循环单元 (Bidirectional Gate Recurrent Unit, BGRU) 进行上下文语义特征表示, 语义学习层的输出 h_t^r 的计算如式(7)所示。

$$\vec{h}_t = \text{GRU}(x_t, \vec{h}_{t-1}) \quad (5)$$

$$\overleftarrow{h}_t = \text{GRU}(x_t, \overleftarrow{h}_{t-1}) \quad (6)$$

$$h_t^r = \vec{h}_t + \overleftarrow{h}_t \quad (7)$$

2.3 位置记忆层

在方面级情感分类任务中, 利用方面词与上下文词汇的相对位置信息可以有效地辅助模型判断^[22]。因为评价方面词的情感词一般存在于方面词周围, 所以方面词附近的词汇理应具备更高的重要性。位置记忆层中用于计算词汇相对位置权重的计算如式(8)所示。

$$w_i^{\text{pm}} = 1 - \frac{|p_i - p_a| - 1}{L} \quad (8)$$

其中, 位置记忆权重使用 w_i^{pm} 表示, 取值范围为 $[0, 1]$, p_i 和 p_a 分别为文本上下文词汇和方面词在句子中的绝对位置, 若方面词由多个词汇组成, 则只记为一个。 L 指文本评论词汇长度, 为了防止位置权重计算结果中出现零和负值的情况, 添加绝对值计算并在分子部分进行减 1 计算。位置记忆网络层的输出 h_i^{pm} 用式(9)计算得到, 词汇位置权重 w_i^{pm} 通过式(8)计算得到, h_i^r 是 2.2 节中语义学习层的输出结果。

$$h_i^{\text{pm}} = w_i^{\text{pm}} h_i^r \quad (9)$$

2.4 图卷积神经网络层

方面级情感分类任务中的评论大部分属于长难句, 因此能够辅助模型梳理句法结构信息会增强方面词与评价词之间的联系, 使得情感分类更加准确。文本评论的依存句法分析树借助 SpaCy 工具获得, 如图 4 所示实例“Great food but the service was dreadful!”的依存句法树结构中蕴含着丰富的句法信息。

本文将图卷积神经网络^[23-24]应用于解决方面

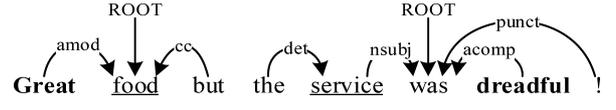


图 4 依存句法树实例图

级情感分析任务中, 充分利用依存句法分析树, 让模型具备句法感知的能力。融合句法信息的图卷积神经网络层实例结构, 如图 5 所示。

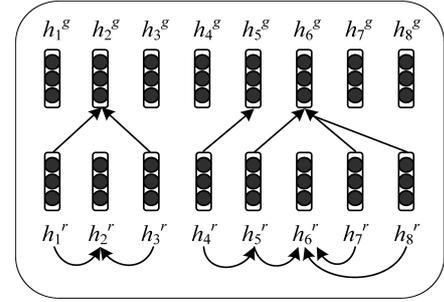


图 5 图卷积神经网络层实例结构

通过将词汇表示为节点, 词汇之间的关系抽象为边, 可以将评论文本使用图结构进行表示。利用图卷积神经网络具备提取相邻节点之间特征的能力, 学习评论文本中词汇之间的关系。邻接矩阵 A 归一化的计算如式(10)所示。

$$\tilde{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (10)$$

其中, D 表示依存句法树的度矩阵。

由文献[3]和文献[25]中证明单层图卷积神经网络未达到学习词汇之间信息的最佳效果, 而且随着图卷积神经网络层数增加, 模型效果不会得到持续提升, 因此本文使用双层 GCN 捕获句法信息, 图卷积神经网络层的输出结果 h_i^g 的计算如式(11)所示。

$$h_i^g = \tilde{A} \text{ReLU}(\tilde{A} h_i^r W_g^{(0)}) W_g^{(1)} \quad (11)$$

其中, ReLU 表示线性整流函数, $W_g^{(0)}$ 和 $W_g^{(1)}$ 分别表示双层图卷积神经网络的权重, h_i^r 是语义学习层的输出信息。

2.5 注意力机制

本文提出综合语义、词性、方面、位置和句法信息的注意力机制^[26], 将语义学习层、位置记忆层和图卷积神经网络层获取的文本信息进行有效利用。

$$v_i = h_i^{\text{pm}} h_i^g \quad (12)$$

$$a_i = \frac{\exp(v_i)}{\sum_{t=1}^n \exp(v_t)} \quad (13)$$

$$h_a = \sum_{i=1}^n a_i h_i^r \quad (14)$$

其中, h_i^r 、 h_i^{pm} 和 h_i^s 分别表示为语义学习层、位置记忆层和图卷积神经网络层学习到的隐藏状态信息, \exp 表示以 e 为底的指数函数, h_a 为注意力机制的输出。

2.6 情感分类层

方面级情感分类是一种多分类任务,由正向、中性和负向三种情感标签组成,情感分类层将注意力机制关注的信息 h_a 作为输入,每种情感标签概率 P_s 按照式(15)进行计算。

$$P_s = \text{softmax}(W_s h_a + b_s) \quad (15)$$

其中, softmax 为多分类任务中的激活函数, W_s 和 b_s 分别为情感分类层的权重和偏置。

3 实验

3.1 数据集

通过在三个公开数据集上的实验来验证本文提出模型的优越性,分别为 SemEval-2014 任务四^[18]的餐馆和笔记本电脑领域的方面级情感分类数据集,以及由 Dong 等^[6]所整理 Twitter 社交领域的方面级情感数据。该数据集中的每条样本都是由评价者真实生成的句子、句子中出现的方面词和方面词对应情感类别所组成,并且标签只包含正向的、中性的和负向的情感,数据集标签类别分布情况见表 1。

表 1 数据集标签类别分布统计表

数据集	正向的	中性的	负向的
Laptop-Train	994	464	870
Laptop-Test	341	169	128
Restaurant-Train	2 164	637	807
Restaurant-Test	728	196	196
Twitter-Train	1 561	3 127	1 560
Twitter-Test	173	346	173

3.2 评价指标

本文采用准确率 Accuracy 和 Macro- F_1 值(简称为 F_1)作为评价指标,验证方面级情感分类任务中的模型性能。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (16)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (17)$$

其中, P 为由计算式(18)得出的 Macro-Precision, R 代表 Macro-Recall,由式(19)计算得到。

$$\text{Macro-Precision} = \frac{1}{C} \sum_{i=1}^C \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \quad (18)$$

$$\text{Macro-Recall} = \frac{1}{C} \sum_{i=1}^C \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (19)$$

其中, C 的值取 3,因为类别中包括正向的、中性的和负向的三种标签。真阳率 TP 和真阴率 TN 分别表示模型正确预测正向、中性和负向情感类别的样本数,假阳率 FP 和假阴率 FN 则分别表示模型错误预测正向、中性和负向情感类别的样本数。

3.3 目标函数

本文提出的端到端训练的模型,可以通过误差反向传播的方式完成参数学习。目标函数使用交叉熵损失函数,完成情感多分类任务。式(20)中, y 表示文本评论的真实情感类别, \hat{y} 为模型预测的情感类别。

$$\text{loss} = \sum_{i=1}^D \sum_{j=1}^C y_i^j \log \hat{y}_i^j + \lambda \|\theta\|^2 \quad (20)$$

其中, D 表示训练集数据样本量, C 与式(18)和式(19)中的 C 意义相同。为缓解模型训练时过度拟合训练集的问题,使用系数为 λ 的二范数作为正则化项。

3.4 参数设置

优化器使用 Adam 算法,模型使用 10^{-3} 学习率,连续 5 次性能不提升则提前停止学习。词嵌入方法分为两种,一种是 GloVe840B^[20] 预训练的 300 维词向量,另一种使用包含 12 层网络结构的 BERT^[27] 预训练语言模型输出作为词向量,维度是 768 维。文本中出现的未登录词和词性向量的维度均为 50,随机初始化矩阵中的参数满足 $(-0.01, 0.01)$ 的均匀分布。同时,为防止模型缺乏泛化能力,使用值为 0.5 的 Dropout 参数随机失活神经网络的神经元,正则化系数 λ 设置为 10^{-5} 。

3.5 对比模型

本文对比模型如下:

- SVM^[5-6]: 支持向量机分类器,通过人工构造大量特征,并引入外部情感词典等方式完成分类。
- LSTM: 长短时记忆网络模型,用于方面级情感分类。
- TD-LSTM^[7]: Target-Dependent LSTM 模

型,在方面词前后上下文分别建模,使用正向和反向两个 LSTM 进行细粒度情感分类。

- **ATAE-LSTM^[10]**: Attention-Based LSTM with Aspect Embedding 模型提出方面词嵌入概念,并采用注意力机制加深文本评论中方面词相关词汇的关注度。

- **GCAE^[12]**: Gated Convolutional Networks 模型,结合门控机制和卷积神经网络进行方面级情感分类,卷积核数目为 100,卷积核尺寸是[3, 4, 5]。

- **IAN^[17]**: Interactive Attention Network 模型,通过注意力机制分别对方面词和上下文建模进行方面级情感分类。

- **RAM^[18]**: Recurrent Attention Network on Memory 模型,提出循环注意力机制完成方面级情感分类任务,多跳数目设置为 3。

- **AOA^[28]**: Attention-over-Attention Network 模型。联合学习方面词和上下文表示,利用注意力机制自动关注文本评论中的重要词汇。

3.6 实验结果与分析

实验在搭载 Intel(R)Xeon(R)CPU E5-2620 处理器和型号为 Tesla P100 GPU,拥有 32GB 内存的服务器上进行。软件环境为 Ubuntu 16. 04. 6 LTS 操作系统,Python 3.6.2 开发平台,基于 PyTorch 深度学习框架构建模型。在三个领域的测试集上模型的指标对比结果如表 2 所示,评价指标分别是 3.2 节所介绍的准确率 Accuracy 和宏平均的 F_1 值,最好的结果加粗表示。由于 SVM 模型用到未知的特征无法进行复现,因此本文引用 Kiritchenko 等^[4]与 Dong 等^[18]论文中的结果进行实验对比,NA 表示没有获得的结果,表 2 中其余模型表现均为本文实验复现结果。此外,除 MemGCN-BERT 模型之外,均采用预训练的 GloVe 词向量进行文本表示,并且最终结果为 5 次实验的平均值^[29]。

表 2 测试集上模型的评测结果表

模型	Laptop		Restaurant		Twitter	
	准确率	F_1	准确率	F_1	准确率	F_1
SVM	70.49	NA	80.16	NA	63.40	63.30
LSTM	69.59	62.73	77.32	65.29	68.87	66.76
TD-LSTM	70.47	64.79	78.73	68.75	69.85	67.85
ATAE-LSTM	73.07	67.97	79.44	68.86	70.28	68.51
GCAE	70.43	64.75	79.16	68.75	70.11	67.83

续表

模型	Laptop		Restaurant		Twitter	
	准确率	F_1	准确率	F_1	准确率	F_1
IAN	71.47	65.78	78.80	68.14	70.52	68.93
RAM	74.16	69.95	80.23	70.35	70.08	68.47
AOA	73.19	68.57	80.07	69.75	71.18	68.84
MemGCN	74.38	70.30	81.55	72.87	72.19	70.29
MemGCN-BERT	78.68	73.93	85.18	77.32	72.98	72.08

从表 2 可以看出:首先,基于统计机器学习的 SVM 模型在 Restaurant 数据集上的效果优于 LSTM 模型,但它们都无法解决评论者对不同方面词情感态度不一致时的方面级情感分类问题。其次,TD-LSTM 模型借助双向 LSTM 对方面词与上下文建立联系,三个数据库中均取得了模型性能的提升。再次,ATAE-LSTM 模型的性能相对于 LSTM 模型获得了 1%~3% 幅度的性能提升,这证明了方面词向量和注意力机制的有效性^[30-31]。最后,基于 CNN 的 GCAE 模型,在全部评测数据集上均与 TD-LSTM 有着相近的表现,但是其具有可以并行计算、模型训练和预测速度更快的优势。

IAN、RAM 与 AOA 模型均采用了将注意力机制与 RNN 结合的方式改进模型,从表 2 的结果可以看出,IAN 模型在 Laptop 和 Restaurant 数据集上的性能略低于 RAM 模型,而在 Twitter 数据集上模型效果显著提升,AOA 模型在三个公开数据集上均取得优秀的方面级情感分类效果,这说明通过注意力机制与循环神经网络结合方式的设计改进,对于优化方面级情感分类模型性能有重要意义。

本文提出的融合注意力机制与词性、方面、位置和句法等辅助信息的 MemGCN 模型在 Laptop、Restaurant 和 Twitter 数据集上均取得优异的方面级情感分类效果。与 LSTM 相比,MemGCN 模型在 SemEval-2014 任务四中 F_1 值均提高超过 7%,Twitter 数据集上准确率取得了 3% 左右的提升。此外,实验中尝试采用预训练语言模型对用户评论进行文本表示,MemGCN-BERT 模型在全部公开数据集上均达到最佳的表现,这证明了 BERT 相比于 GloVe 更擅于表示文本,下游的图卷积记忆网络模型能够捕获到更多的情感信息。虽然使用预训练语言模型可以获得分类效果的进一步提升,但是训练模型和使用模型进行预测时需要花费更多的时间,而且模型训练也会占用大量内存、计算资源和存储空间。另

外,在样本量充足且均衡的社交领域数据上,通过两种文本表示方法训练出来的模型效果较为接近。

为验证词性、方面、位置和句法四种辅助信息对 MemGCN 模型分类准确率和 F_1 的影响,实验中对 MemGCN 模型进行了如图 6 和图 7 所示的消融实验。在柱状图中,使用黑色表示包括所有辅助信息的 MemGCN 模型性能,灰黑色、灰色、灰白色和白色柱用于区分去掉位置、句法、词性和方面信息的模型表现。

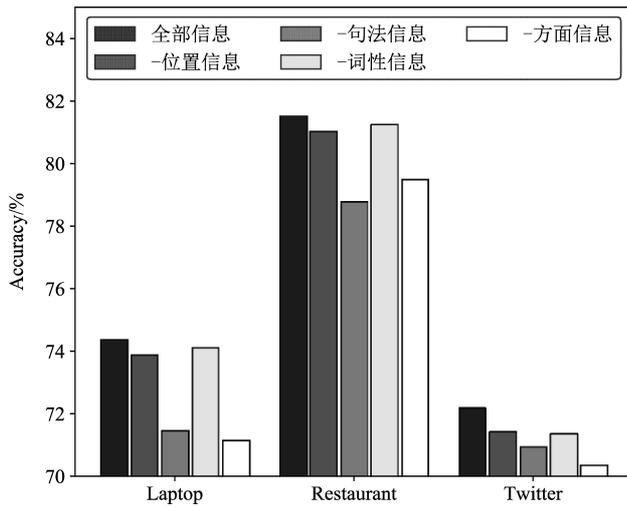


图 6 句法与辅助信息对分类准确率的影响

首先,在图 6 和图 7 中均可以看出,词性、方面、位置和句法四种辅助信息都有益于 MemGCN 模型进行方面级情感分类。其次,本文提出的句法信息

对于模型性能的影响是最大的。最后,其余辅助信息对 MemGCN 模型性能影响的重要程度依次为方面信息、位置信息和词性信息。

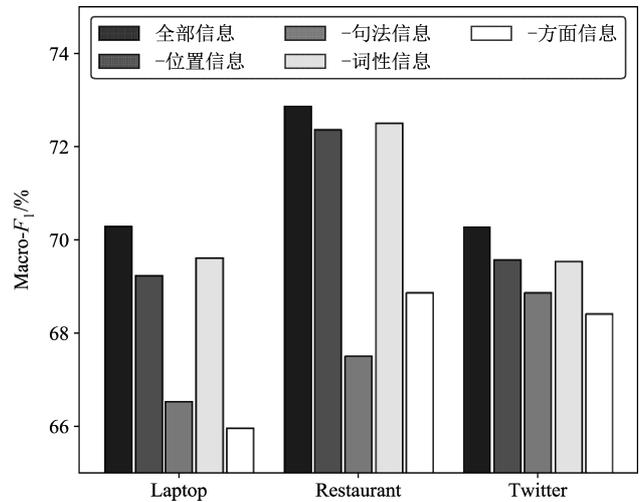


图 7 句法与辅助信息对分类 F_1 影响

为进一步探索不同模型的方面级情感分类效果,本文分别将 GCAE、IAN、RAM 和 AOA 模型与本文所提出的 MemGCN 模型实例分类效果进行对比分析。各模型在两个实例、三个方面词上的分类结果见表 3,其中,“真实”列表示评论者的正确情感类别。在对比中可以发现,MemGCN 模型具备根据特定方面词进行情感分类的能力,而在句法结构复杂的句子中,对比模型难于给出正确答案。

表 3 实例分析

实例	方面词	GCAE	IAN	RAM	AOA	MemGCN	真实
Great <u>food</u> but the <u>service</u> was dreadful !	<u>food</u>	负向	负向	正向	正向	正向	正向
	<u>service</u>	负向	正向	正向	负向	负向	负向
We didn't know if we <u>should</u> order a <u>drink</u> or <u>leave</u> ?	<u>drink</u>	负向	负向	负向	负向	中性	中性

图 8~图 10 分别为 MemGCN 在餐馆领域正向、负向和中性情感实例的可视化分析,如图所示,颜色越深则代表注意力权重越大。

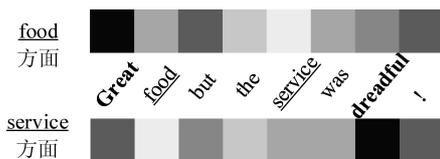


图 8 “food”“service”方面注意力权重示意图

在图 8 中,实例的两个评价词分别修饰两个方面词。对于“food”方面词来说,本文提出的

MemGCN 模型对带有明显正向情感的形容词“Great”注意力机制的权重最大,于是能够将方面词“food”归于正向评价。同样地,当方面词为“service”时,MemGCN 模型的注意力集中在“dreadful”评价词上,因此将服务方面评论的情感态度分类为负向。

通过图 9 对以上评论文本依存句法分析树的可视化,证明了在一句话中同时出现多个方面词与多个评价词时,MemGCN 可以正确匹配方面词对应的评价词,这样可以保证记忆网络在看到新的评价词“dreadful”时,不会遗忘“Great”对“food”方面的修饰,也可以解决模型过于关注评价词“Great”而与

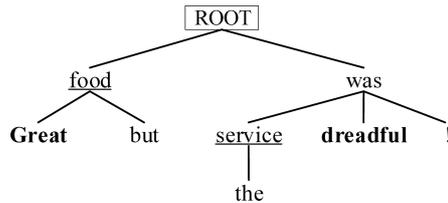


图9 实例的依存句法树可视化示意图

“service”方面词错误匹配的问题。



图10 “drink”方面注意力权重示意图

从图10可以看出,在句法结构较为复杂的实例中,MemGCN模型的注意力集中在否定词“n't”“should”和“leave”词汇上,将此实例的情感判别为中性。由于训练数据中中性情感的样本量较小,而且有离开之意的“leave”词距离方面词很近,可能将这条数据错误预测为负向情感。

4 总结

为解决复杂句法结构中的方面级情感分类方面词与评价词之间的依赖问题,本文使用注意力机制,结合记忆文本语义、词性、方面与位置信息的记忆网络,以及基于依存句法分析树的图卷积神经网络,提出了一种句法信息感知的方面级情感分类模型。实验结果证明,通过加入句法信息,可以增强评价词与方面词之间的联系,能够明显改善方面级情感分类的效果。在未来的工作中,我们将尝试更先进的预训练语言模型优化语义信息,并计划将本文提出的模型在其他细粒度情感分析任务,如评价对象提取中进行尝试。

参考文献

- [1] Liu B. Sentiment analysis and opinion mining[M]. San Rafael: Morgan and Claypool Publishers, 2012: 1-167.
- [2] Zhou J, Huang J X, Chen Q, et al. Deep learning for aspect-level sentiment classification: survey, vision, and challenges [J]. IEEE Access, 2019, 7: 78454-78483.
- [3] Thomas N, Max W. Semi-supervised classification with graph convolutional networks[C]//Proceedings of the Conference on International Conference on Learning Representations, 2017.

- [4] Liang Y, Mao C S, Luo Y. Graph convolutional networks for text classification[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019: 7370-7377.
- [5] 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, 2014: 437-442.
- [6] Dong L, Wei F R, Tan C Q, 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. Stroudsburg, USA: ACL, 2014, II: 49-54.
- [7] Tang D Y, Qin B, Feng X C, et al. Effective LSTMs for target-dependent sentiment classification[C]//Proceedings of the 26th International Conference on Computational Linguistics. Berlin, Germany: Springer, 2016: 3298-3307.
- [8] Tom Y, Devamanyu H, Soujanya P, et al. Recent trends in deep learning based natural language processing[J]. IEEE Computational Intelligence Magazine 2018, 13(3): 55-75.
- [9] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[C]//Proceedings of the 3rd International Conference on Learning Representations, 2015: 1-15.
- [10] Wang Y Q, Huang M L, Zhao L, et al. Attention-based LSTM for aspect-level sentiment classification [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2016: 606-615.
- [11] 曾锋, 曾碧卿, 韩旭丽, 等. 基于双层注意力循环神经网络的方面级情感分析[J]. 中文信息学报, 2019, 33(6): 108-115.
- [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. Melbourne, Australia: ACL, 2018: 2514-2523.
- [13] 梁斌, 刘全, 徐进, 等. 基于多注意力卷积神经网络的特定目标情感分析[J]. 计算机研究与发展, 2017, 54(8): 1724-1735.
- [14] 刘全, 梁斌, 徐进, 等. 一种用于基于方面情感分析的深度分层网络模型[J]. 计算机学报, 2018, 41(12): 2637-2652.
- [15] 杜慧, 俞晓明, 刘悦, 等. 融合词性和注意力的卷积神经网络对象级情感分类方法[J]. 模式识别与人工智能, 2018, 31(12): 1120-1126.
- [16] Tang D Y, Qin B, Liu T. Aspect level sentiment classification with deep memory network[C]//Proceedings of the 2016 Conference on Empirical Meth-

- ods in Natural Language Processing. Stroudsburg, USA: ACL, 2016: 214-224.
- [17] Ma D H, Li S J, Zhang X D, et al. Interactive attention networks for aspect-level sentiment classification [C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Berlin, Germany: Springer, 2017: 4068-4074.
- [18] Chen P, Sun Z Q, Bing L D, et al. Recurrent attention network on memory for aspect sentiment analysis [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2017: 452-461.
- [19] Zhu P S, Qian T Y. Enhanced aspect level sentiment classification with auxiliary memory [C]//Proceedings of the 27th International Conference on Computational Linguistics. Stroudsburg, USA: ACL, 2018: 1077-1087.
- [20] Pennington J, Socher R, Manning C D. Glove: Global vectors for word representation [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2014: 1532-1543.
- [21] 贾川,方睿,浦东,等. 基于循环实体网络的细粒度情感分析[J]. 中文信息学报, 2019, 33(9): 123-128.
- [22] 廖祥文,林威,吴运兵,等. 基于辅助记忆循环神经网络的视角级情感分析[J]. 模式识别与人工智能, 2019, 32(11): 987-996.
- [23] Diego M, Ivan T. Encoding sentences with graph convolutional networks for semantic role labeling [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2017: 1506-1515.
- [24] Xu Y, Yang J. Look again at the syntax: Relational graph convolutional network for gendered ambiguous pronoun resolution [J]. arXiv preprint arXiv: 1905.08868, 2019.
- [25] 李鸿宇. 基于图卷积记忆网络的细粒度情感分析研究[D]. 葫芦岛: 辽宁工程技术大学硕士学位论文, 2020.
- [26] He R D, Lee W S, Ng H T, et al. Effective attention modeling for aspect-level sentiment classification [C]//Proceedings of the 27th International Conference on Computational Linguistics. Stroudsburg, USA: ACL, 2018: 1121-1131.
- [27] Devlin J, Chang M, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: NAACL-HLT, 2019: 4171-4186.
- [28] Huang B X, Ou Y L, Kathleen M. Aspect level sentiment classification with attention-over-attention neural networks [C]//Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, 2018: 197-206.
- [29] He R D, Lee W S, Ng H T, et al. Exploiting document knowledge for aspect-level sentiment classification [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2018: 579-585.
- [30] Majumder N, Poria S, Gelbukh A F, et al. IARM: inter aspect relation modeling with memory networks in aspect-based sentiment analysis [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2018: 3402-3411.
- [31] Liu Q, Zhang H B, Zeng Y F, et al. Content attention model for aspect based sentiment analysis [C]//Proceedings of the World Wide Web Conference 2018. Washington, USA: IEEE, 2018: 1023-1032.



王光(1979—), 硕士, 副教授, 硕士生导师, 主要研究领域为自然语言处理、情感分析。
E-mail: 275469783@qq.com



邱云飞(1976—), 博士, 教授, 博士生导师, 主要研究领域为数据挖掘、自然语言处理。
E-mail: 7415575@qq.com



李鸿宇(1994—), 通信作者, 硕士研究生, 主要研究领域为自然语言处理、情感分析。
E-mail: 1757732597@qq.com