

融合位置信息的卷积门控网络实现与应用

时昭丽, 范红, 陈佳伟, 董亚博, 张子薇, 许武军

(东华大学 信息科学与技术学院, 上海 201620)

摘要:基于方面的情感分析 (Aspect-Based Sentiment Analysis) 通常使用长短期记忆网络和注意力机制方法, 这两种模型结构复杂, 运行时间长。现有的卷积神经网络结构简单, 具有代表性的是 GCAE (Gated Convolutional Networks with Aspect Embedding) 模型。但其由于未充分地利用词语的顺序信息, 不能准确快速地关注到关键词。因此提出了一种融合位置信息的卷积门控网络方法。采用 SemEval 数据集进行实验, 并与采用 GCAE 模型的实验结果进行对比, 结果表明, 所提模型迭代一次约用时 5.96 s, 优于长短期记忆模型的 81 s。该模型对句子中有多个方面的情感极性判断准确率为 55.00%, 高于 GCAE 模型的 53.00%。该研究对于提高基于方面的情感分析的迭代时间和准确度有一定的参考意义。

关键词:方面情感分析; 位置信息; 卷积神经网络; 门控机制

中图分类号: TP391

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.01.013

引用格式: 时昭丽, 范红, 陈佳伟, 等. 融合位置信息的卷积门控网络实现与应用 [J]. 信息技术与网络安全, 2020, 39 (1): 70-73.

Implementation and application of convolution gating network model combined with position information

Shi Zhaoli, Fan Hong, Chen Jiawei, Dong Yabo, Zhang Ziwei, Xu Wujun

(College of Information Science and Technology, Donghua University, Shanghai 201620, China)

Abstract: Aspect-based sentiment analysis usually uses long-short-term network and the attention mechanism method, the two models are complicated in structure and long time in running time. The existing convolutional neural network has a simple structure, and is representative of a GCAE (Gated Convolutional Networks with Aspect Embedding) model. Since the position information of the words does not fully utilized, the keywords cannot be accurately and quickly focused. Therefore, this paper proposed a convolutional gating network method that combines position information. The SemEval dataset was used for experiments and compared with the experimental results using the GCAE model. The results show that the iteration time of this model is about 5.96 s, which is better than the LSTM model of 81 s. The accuracy of this model for multiple aspects of sentences and multiple emotional polarity is 55.00%, which is higher than 53.00% of the GCAE model. This paper has certain reference significance for improving the iterative time and accuracy of aspect-based sentiment analysis.

Key words: aspect-based sentiment analysis; position information; convolutional neural network; gating mechanism

0 引言

情感分析分为基于句子的情感分析和基于方面的情感分析。如“这家餐厅菜品好吃, 但交通不便利”, 文本的实体为“这家餐厅”, 分别对“菜品”和“交通”两个方面进行褒贬不一的评论。若判断整个句子的情感极性为积极、消极或中立显然是不精准的, 因此关注方面情感更为重要。基于方面的情感分析前期主要采用基于机器学习的利用情感词典的特征构建分类器^[1], 其主要依赖人工规则和特

征工程, 且耗时成本高。近年来神经网络在方面情感分析中得到广泛使用, 文献[2]提出一种用于方面相关的自适应递归神经网络, 但是其依赖语境和句法结构预测方面的情感极性。长短期记忆模型^[3]能解决长距离依赖的问题, 文献[4]改进了该模型, 引入树结构长短期记忆模型, 准确度有所提升, 但模型训练时间长。文献[5]提出应用注意力机制方法实现对方面词周围情感词的关注, 但注意力机制不能捕捉序列的顺序, 且注意力层涉及指数

操作和句子中所有单词的所有对齐分数的标准化,所需计算时间长。文献[6]将注意力机制、长短期记忆网络和位置信息相结合,建立情感分析模型,但当句子中含有多个方面以及多个方面的情感词时,其判断可靠性一般,训练时间相比长短期记忆网络更长。文献[7]提出 GCAE 模型,其改善了系统的性能,提高了情感极性的预测准确度,但忽略了位置信息的重要性。

为了解决以上问题,本文提出了融合位置信息的卷积门控网络方法,采用卷积神经网络(Convolution Neural Network, CNN)模型进行并行处理,融合位置信息实现对文本中方面词顺序地捕捉。本文提出的模型相比较于长短期记忆网络模型和注意力机制模型运行的时间更短,性能有较大的改善。相对 GCAE 模型,对多方面多情感的文本的情感预测准确度更高。

1 卷积门控网络模型

1.1 卷积神经网络模型

卷积神经网络(CNN)主要由输入层、卷积层、池化层和全连接层组成,如图1所示。输入层为文本的词向量;卷积层采用不同尺寸的过滤器对输入的词向量有效地提取特征;池化层对提取的特征执行下采样操作;全连接层连接池化层操作并输出。

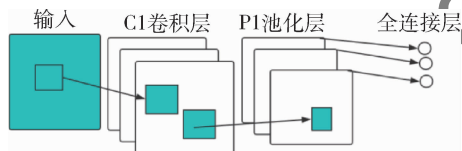


图1 CNN模型

1.2 GCAE 模型

卷积门控网络模型采用卷积神经网络与门控机制相结合。卷积神经网络模型输入层为一个文本句的嵌入向量,通常使用预先训练好的 Glove 词向量初始化每一个单词,输出为嵌入向量 $v_i \in \mathbf{R}^D$, D 是嵌入向量的维度。当输入句子含有 L 个单词,通过嵌入层可以用一个矩阵 $\mathbf{X} = [v_1, v_2, \dots, v_L]$ 表示。因为每一个评论文本的长度不一致,为了统一处理,在每一批次文本输入时取本批次中句子长度最长的 L ,对于长度短的句子在句尾补充 0 向量。

卷积层是通过卷积过滤器 $W_c \in \mathbf{R}^D \times K$ 将 k 个单词映射为一个特征,过滤器遍历整个句子时,获得了一个新的特征 C 为 $[C_1, C_2, \dots, C_L]$ 。

$$c_i = f(x_{i:i+k} * w_c + b_c) \quad (1)$$

式中 b_c 是偏差, f 是一个非线性的激活函数,“*”表明是卷积操作。

门控机制是由两个非线性激活函数组成的,根据方面嵌入向量连接到两个卷积神经层分别获得方面特征和情感特征。最终输出特征 C_i 为:

$$c_i = s_i \times a_i \quad (2)$$

其中,特征 a_i 是通过 ReLU 激活函数控制输出的方面特征, s_i 是通过 tanh 激活函数控制输出的情感特征。门控机制根据给定的方面信息与方面特征 a_i 之间的相关性输出相似度分数,如果分数小于 0,这个方面的情感特征将在门口被阻止,否则被输出。

GCAE 模型如图2所示。最大池化层进一步移除句子中不重要的特征。全连接层将最大池化层的输出全连接输出。由于对餐厅评论进行基于方面的情感分析是一个多分类的问题,选用多分类的交叉熵函数作为损失函数,最终通过最小化所有数据样本的真实输出值 y 与预测 $\hat{y}$ 的交叉熵来训练模型,如式(3)所示:

$$L = -y'_i \log y_i^j \quad (3)$$

其中, i 是数据样本的索引, j 是一个情感类的索引。目前的 GCAE^[7] 网络模型忽略了序列顺序对情感分析的重要性。

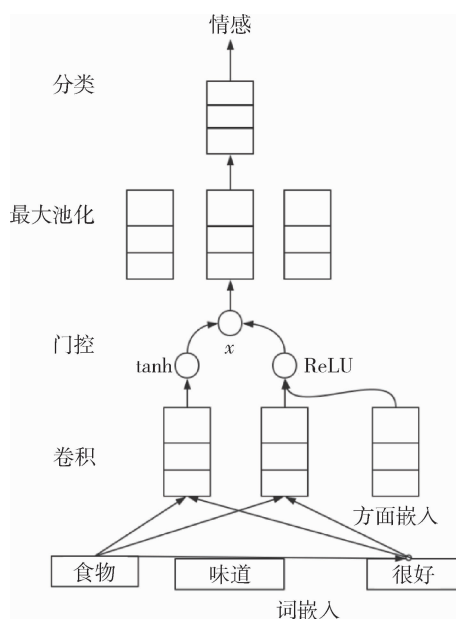


图2 GCAE模型

2 融合位置信息的卷积门控网络模型

本文提出的模型融合了位置信息,对文本中词

语出现的位置进行编码,实现对顺序信息的捕捉。采用正余弦函数,如式(4)和(5)所示。在偶数位置使用正弦编码;在奇数位置使用余弦编码。将 id 为 p 的位置映射为一个 d_{pos} 维的位置向量,第 i 个元素的数值为 $PE_i(p)$,其中 pos 为词语在序列中的位置。

$$PE_{2i}(p) = \sin(p/10\,000^{2i/d_{pos}}) \quad (4)$$

$$PE_{2i+1}(p) = \cos(p/10\,000^{2i/d_{pos}}) \quad (5)$$

向量本身是一个绝对位置信息且词向量本身没有局部结构,为了构造局部结构将词向量与位置向量相加作为输入层。式(6)更加精准地输出每一个方面的情感极性。

$$a_i = \text{relu}(x_{i:i+k} * w_a + V_a v_a + P_a p_a + b_a) \quad (6)$$

融合位置信息的卷积门控网络模型如图3所示。

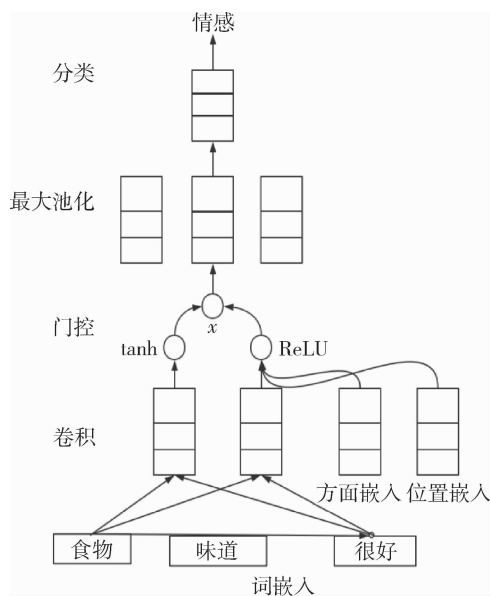


图3 本文模型

3 实验结果及分析

3.1 数据与实验准备

为了验证本文提出的结合位置信息的卷积门控网络模型的效果,本文选择了来自 SemEval 任务4中餐厅公开数据集。数据集分为训练数据、验证数据和测试数据。内容为顾客对餐厅的评价,方面有5个:食物(food)、价格(price)、服务(service)、环境(ambience)和设备其他(misc);数据集的情感类别有4类:积极的(positive)、消极的(negative)、中性的(netural)和冲突的(conflict),通过超参数的设置来实现情感极性四分类。因为不同方面具有不

同情感标签的句子在评论数据中很常见,而 SemEval 数据集中文本长度较短,一个文本评论大多包含两个方面两个情感。因此为了判断本模型对一个复杂的包含多个方面多个情感词的文本情感极性预测的准确性,依据原数据创建 Hard 数据集。

Hard 数据是由两个以上方面和相关联的方面情感标签的句子构成。

3.2 实验参数

在所有实验中,使用 Glove 向量初始化每个单词的嵌入向量,这些向量在未标记的 8 400 亿个分词数据上预先训练,每个词向量的维度都是 300, Glove 词汇表中的单词随机初始化为符合均匀分布 $U(-0.25, 0.25)$ 。融合位置信息和方面信息的 CNN 卷积核的宽度分别设为 3, 4, 5, 迭代次数为 30。每次的批量大小为 30, 由于文本句子的个数不一定是 batch 的整数,因此以最后的数据为准。Adagrad 中初始的学习率设置为 0.01, 所有神经模型都在 PyTorch 中实现。

3.3 实验结果与分析

选用 SemEval 任务4中餐厅评论测试数据集和 Hard 测试数据集进行实验,分别对 GCAE 模型和本文模型进行验证。为了避免不稳定的数据显示,每个模型进行 5 次实验取其均值。

GCAE 模型结合了 CNN 两个卷积层,根据上下文向量的每个维度上给定的方面信息控制情感的信息流,产生方面和情感特征,输出粒度精细到方面而不是对齐关注层中整个句子这一维度进行评分。这对于基于方面的情感分类提供了更加详细的预测依据。本文模型在 Hard 数据集准确率为 55%, 优于 GCAE 的 53%, 表明本文模型对多个方面和多个情感的文本预测,加入的位置信息提高了准确度。根据文本长度以及方面的位置信息实现所有方面特征的输出,为识别并输出每一个方面特征的情感特征提供精确的数据,最终实现多个方面多个情感准确度的提高。

本文模型采用结构简单的卷积神经网络,在相同的底层框架、CPU、GPU 下分析不同网络模型在餐厅数据集上完成一次迭代的训练时间。参考文献[8]数据得知,广泛使用的长短期记忆网络模型的训练时间为 81 s, 几乎是本文使用卷积神经网络模型训练时间 5.96 s 的 13 倍, 本文模型实现了性能上的优化。

GCAE 模型和本文模型分别在测试集和 Hard 数据集上对基于方面的情感极性预测的准确率如表 1 所示。

表 1 GCAE 模型和本文模型分别在测试集和 Hard 数据集上准确率对比 (%)

模型	餐厅数据集	
	测试集	Hard 数据集
GCAE	78	53
本文模型	77	55

4 结论

本文提出了一种融合位置信息的卷积门控网络模型实现基于方面的情感分析,模型通过两个卷积层融合位置信息将方面特征和情感特征更加准确和快速地输出并预测情感的极性。近几年基于评论的推荐系统应运而生,由于评论^[9]包含了用户对于产品的描述和情感倾向,能直接反映用户特征和产品特征,可以帮助人们更好地建模用户和商品,从而实现更准确的推荐。研究发现评分的高低与评论者真实的情感不一,因此如何将评论文本情感趋向与推荐系统有效地结合^[10]以提供更精准的推荐将是未来的研究方向。

参考文献

[1] 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,2017,41(12):2637-2652.

[2] 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 Association for Computational Linguistics,2014:49-54.

[3] TANG D Y, QIN B, LIU T. Document modeling with gated recurrent neural network for sentiment classification[C]. Proceedings of the 2015 Conference on Empirical Methods

in Natural Language Processing,2015:1422-1432.

[4] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long-term memory network[C]. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7rd International Joint Conference on Natural Language Processing, Stroudsburg, PA: Association for Computational Linguistics, 2015:1556-1566.

[5] WANG Y Q, HUANG M L, ZHAO L. Attention-based LSTM for aspect-level sentiment classification[C]. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing,2016:606-615.

[6] 武婷,曹春萍. 融合位置权重的基于注意力交叉注意力的长短期记忆方面情感分析模型[J]. 计算机应用, 2019,39(8):2198-2203.

[7] 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, 2018.

[8] 袁丁,章剑林,吴广建. 基于方面级的餐厅用户评论细粒度情感分析[J]. 软件,2019,40(8):181-189.

[9] CHEN C, ZHANG M, LIU Y, et al. Neural attentional rating regression with review-level expression[C]. Proceedings of the 2018 WorldWide Web Conference on World Wide Web, 2018:1583-1592.

[10] HE X N, CHEN T, KAN M Y, et al. Trirank: review-aware explainable recommendation by modeling aspects [C]. Proceedings of the 24th ACM International Conference on Information and Knowledge Mangement,2015:1661-1670.

(收稿日期:2019-11-07)

作者简介:

时昭丽(1995-),女,硕士研究生,主要研究方向:自然语言处理。

范红(1976-),通信作者,女,博士,副教授,主要研究方向:多媒体通信与图像处理。E-mail:dhufanhong@dhu.edu.

(上接第 58 页)

作者简介:

张志(1983-),男,学士,助理工程师,主要研究方向:国防工程与防护。

邓忠凯(1991-),通信作者,男,硕士,助理工程师,主要研究方向:烟气扩散,人员疏散仿真。E-mail:dzhongkai@163.com。

杨水亲(1986-),男,学士,助理工程师,主要研究方向:通信工程。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所