

基于情感膨胀门控 CNN 的情感-原因对提取^{*}

代建华^{1,2,3} 邓育彬^{1,3}

¹(湖南师范大学智能计算与语言信息处理湖南省重点实验室 长沙 410081)

²(湖南师范大学语言与文化研究院 长沙 410081)

³(湖南师范大学信息科学与工程学院 长沙 410081)

摘要:【目的】针对情感-原因对抽取这一情感分析任务,提出情感膨胀门控 CNN(EDGCNN)模型。【方法】首先使用情感判别模型 CNN 找出情感句,然后将情感句编码输入到融入情感特征的 EDGCNN 模型,找出相应的情感原因,得到情感-原因对,并在实验数据集进行情感原因关键字标注。【结果】召回率和 F1 值分别达到了 63.52% 和 60.45%,召回率优于已有方法最好结果,F1 值与已有方法最优性能相当,而且能从更细粒度实现情感-原因对抽取。【局限】情感-原因对语料规模较小,有待进一步扩充完善。【结论】EDGCNN 模型能够从文本中更好地抽取情感-原因对。

关键词: 情感-原因对抽取 EDGCNN 情感判别

分类号: TP391 G202

DOI: 10.11925/infotech.2096-3467.2019.1243

引用本文: 代建华, 邓育彬. 基于情感膨胀门控 CNN 的情感-原因对提取[J]. 数据分析与知识发现, 2020, 4(8): 98-106.(Dai Jianhua, Deng Yubin. Extracting Emotion-Cause Pairs Based on Emotional Dilation Gated CNN [J]. Data Analysis and Knowledge Discovery, 2020, 4(8): 98-106.)

1 引言

在信息爆炸的时代,社交网络为人们提供了一个共享各种新闻和信息源的平台,所以从互联网的海量文本(如商品的在线评论、舆论新闻等)中挖掘出情感倾向和情感原因是非常有意义的。文本中情感研究是自然语言处理领域的热门方向,情感研究又分为情感分类和情感原因抽取(Emotion Cause Extraction, ECE),目前大部分研究都集中在情感分类的研究上。如 Rout 等^[1]提出一种基于词权重的无监督的 Twitter 文本情感分类方法。Peng 等^[2]通过训练原域大量带有标签的数据及目标域少量带有标签的数据,对目标域大量无标签数据进行跨领域情感分类。为了解决特定方面的情感极性分类问题,刘全等^[3]提出结合区域卷积神经网络(Convolutional

Neural Networks, CNN)和分层长短期记忆网络(Long Short-Term Memory, LSTM)的深度分层模型。张庆庆等^[4]研究了深度信念网络在文本情感分类任务中的参数选择与性能。

一般来说,个人情感的产生、表达和感知受到许多因素的影响,情感原因被认为是能够引发相应情感的事件或条件,因此情感原因的分析对于民意挖掘和知识发现等至关重要。但由于 ECE 任务困难加上相关标注的数据集较少,所以相应的研究^[5-6]也较少。在深度学习之前,情感原因抽取的研究是基于规则的。例如, Lee 等^[7]提出一种基于文本驱动和规则的汉语情感原因关系检测方法; Yada 等^[8]使用自举技术,自动获取连接词作为文本提示模式以提取情感原因; Gao 等^[9]提出一种基于规则的中文微博

通讯作者:代建华,ORCID:0000-0003-1459-0833,E-mail:jhdai@hunnu.edu.cn。

^{*}本文系湖南省科技创新计划项目“湖湘高层次人才聚集工程-创新人才”(项目编号:2018RS3065)和“智能计算与语言信息处理湖南省重点实验室”(项目编号:2018TP1018)的研究成果之一。

3 实验模型

3.1 情感判别模型

CNN在图像领域有很多成功应用,后来也成功运用在自然语言处理领域中。Kim^[16]使用CNN进行文本分类并取得了较好的结果,表明CNN具有很强的语义特征捕捉能力。情感表达往往会带有诸如“担心”“感动”这种能够表示情感倾向的词,因此融合字词向量的CNN模型能够学习到这些关键词的特征,有利于情感语句的判别。

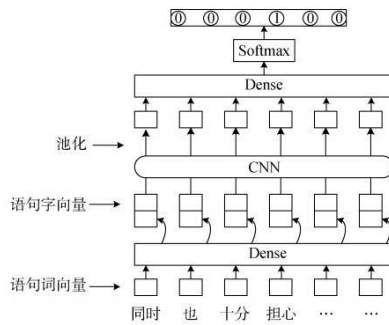


图3 情感语句判别模型

Fig.3 Emotional Discriminant Model

情感判别模型如图3所示,文本 d 中的每一条语句 c_i ,经过分词后,输入到预训练好的词向量Word2Vec模型得到语句的词向量序列 I_{c_i} ,如公式(2)所示。

$$I_{c_i} = \{x_{c_1}, \dots, x_{c_n}\} \quad (2)$$

其中, n 为语句 c_i 分词后词的个数, x_{c_n} 表示语句 c_i 的词向量。

语句 c_i 中的字经过Word2Vec模型得到字向量 z_i ,进而得到语句 c_i 的全部字向量 Z_i ,如公式(3)所示。其中, m 为语句中字的个数。

$$Z_i = \{z_1, \dots, z_m\} \quad (3)$$

得到句子的词向量 I_{c_i} 后,再经过变换矩阵得到 H_i ,如公式(4)所示。其中, h_m 为矩阵变换后的向量。

$$H_i = \{h_1, \dots, h_m\} \quad (4)$$

最后字词向量融合得到 ZC_i ,如公式(5)所示。

$$ZC_i = \{zc_1, \dots, zc_m\} \quad (5)$$

其中, m 为语句中字的个数, zc_m 为融合的字词向量。

字词融合向量进入CNN层,再进入池化层(Pooling)、全连接层(Dense),最后进入Softmax层,得到情感判别结果。最终得到文本带有情感的语句 $C = \{c_1, \dots, c_t\}$, t 为带有情感的语句个数。

3.2 情感-原因对提取

情感和原因往往是相互关联的,所以本文使用融合情感倾向特征的深度学习框架去找出情感-原因对。

(1) 基于EDGCNN的情感-原因对抽取模型

本文EDGCNN模型框架如图4所示。主要包含两个模块:情感语句编码模块和情感-原因对定位模块。

由图3情感判别模型得到带有情感倾向的语句 $C = \{c_1, \dots, c_t\}$,再经过Word2Vec模型分别得到语句的词向量序列 C_i ,依次作为情感语句编码模块的输入。本文使用情感判别模型Dense层的输出 O_{d_i} 作为EDGCNN的情感特征编码。情感编码模块的最终输入为 I_i ,如公式(6)所示。

$$I_i = \{C_i, O_{d_i}\} \quad (6)$$

随后进入CNN进行语义特征抽取得到语句上层语义特征 HC_i ,如公式(7)所示。

$$HC_i = \{hc_1, \dots, hc_v\} \quad (7)$$

其中, v 代表卷积核的个数, hc_v 表示在 I_i 上进行一维卷积^[16]所得到的上层特征。采用Laskar等^[17]提出的空间注意力(Spatial Attention)机制代替简单的池化,每一个卷积核的Attention矩阵参数共享,最后得到融合情感倾向的情感语句上层编码 HE ,如公式(8)所示。

$$HE = \{l_1, \dots, l_t\} \quad (8)$$

情感-原因对定位模块主要包括特征输入层、膨胀卷积层和模型输出层。卷积层负责捕获语句的特征,输出层负责最后的情感原因定位。模型的输入融合了语句的字词向量和情感语句编码。

(2) EDGCNN卷积结构

① 门控卷积

模型采用Dauphin等^[18]提出的门控卷积结构,使用门控卷积可以有效降低梯度弥散的可能,并且还具有保留非线性变换的作用。所以,门控卷积在

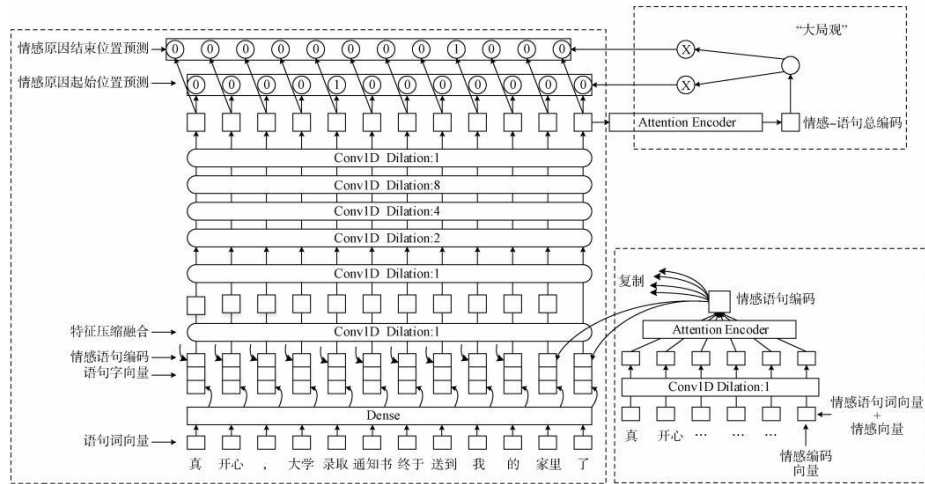


图4 情感-原因对提取模型
Fig.4 Emotion-Cause Pair Extraction Model

一些自然处理任务(如机器翻译)中取得不错的效果。

给定输入序列 $X = [x_1, x_2, \dots, x_n]$, 门控卷积神经网络与标准卷积不同的是, 前者在每层进行卷积操作后都加上一个输出门限, 用来控制信息的流动, 如公式(9)所示。

$$Y = \text{Conv1D}_1(X) \otimes \sigma(\text{Conv1D}_2(X)) \quad (9)$$

其中, $\sigma()$ 为 sigmoid 激活函数; Conv1D_1 和 Conv1D_2 表示卷积操作, 两个卷积的卷积核数一样多, 并且权值不共享。

② 膨胀卷积

传统卷积神经网络对序列建模受到卷积核大小的限制, 但是在自然语言处理中, 文字往往是具有依赖性的, 对于距离较远的依赖, 如果想要抓取这种依赖关系, 必须线性地堆叠模型的层数, 所以为了抓取文字更长的依赖关系。本文使用膨胀卷积神经网络代替传统的卷积神经网络, 膨胀卷积神经网络如图5所示。

相比于传统卷积神经网络, 膨胀卷积在每一层做卷积的时候允许间隔采样, 采样间隔受膨胀系数 d 的控制, 如图5所示, 当取 $d = 1$ 时, 表示对这一层每

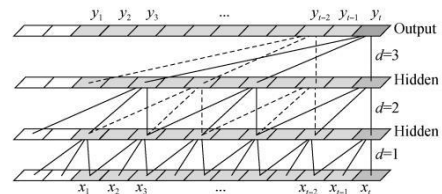


图5 膨胀卷积神经网络
Fig.5 Dilated Convolutions

个点都进行采样; 当 $d = 2$ 时, 表示每两个点采样一个作为卷积输入。因此, 层数越高, d 越大, 膨胀卷积的窗口随着层数的增加呈现指数增长, 从而很大程度上减少了卷积的层数。在图像领域, 可以极大地扩大视野; 在自然语言处理领域, 能够抓取传统卷积神经网络需要堆叠层数才能获取的长距离依赖关系, 所以使用膨胀卷积神经网络代替传统卷积神经网络进行特征抽取。图4模型中的 Dilation:4 表示膨胀卷积神经网络的膨胀系数为4。

(3) 注意力机制

EDGCNN 模型采用注意力机制^[9], 用来代替替

单的池化对序列信息进行整合。图4模型中的情感语句编码为一个总的情感语句向量,也将语句的序列编码为一个总的编码向量。注意力机制具体描述如公式(10)和公式(11)所示。

$$z = \text{Encoder}(z_1, z_2, \dots, z_n) = \sum_{i=1}^n \lambda_i z_i \quad (10)$$

$$\lambda_i = \text{Softmax}(\alpha^T \text{Act}(Wx_i + b) + \beta) \quad (11)$$

其中, α 和 W 为可训练参数; $\text{Act}()$ 为激活函数, 本模型取 \tanh 函数; β 为偏置项。

(4) EDGCNN模型输出设计

模型的输出参考文献[15]中的双标注输出, 使用两次标注的方式分别标注出情感原因的起始位置和结束位置, 如公式(12)和公式(13)所示。

$$P_i^{start} = \sigma(\alpha_1^T \text{Act}(W_1 x_i + b_1) + \beta_1) \quad (12)$$

$$P_i^{end} = \sigma(\alpha_2^T \text{Act}(W_2 x_i + b_2) + \beta_2) \quad (13)$$

其中, P_i^{start} 和 P_i^{end} 分别表示模型的起始位置和结束位置; x_i 为卷积提取的特征; $\text{Act}()$ 为 \tanh 激活函数; $\sigma()$ 为 sigmoid 激活函数; 其他为可训练参数。

为了增加图4模型的“大局观”, 本文先将语句的序列编码成一个整体向量, 同样使用注意力机制代替简单的池化, 然后接全连接层得到一个总的打分 p^{global} , 如公式(14)-公式(17)所示。

$$o = \text{Encoder}(x_1, x_2, \dots, x_n) \quad (14)$$

$$p^{global} = \sigma(Wo + b) \quad (15)$$

$$P_i^{start} = p^{global} \sigma(\alpha_1^T \text{Act}(W_1 x_i + b_1) + \beta_1) \quad (16)$$

$$P_i^{end} = p^{global} \sigma(\alpha_2^T \text{Act}(W_2 x_i + b_2) + \beta_2) \quad (17)$$

全局打分对模型的收敛和情感原因定位效果有重要的作用, 如果不存在情感原因, “大局观” p^{global} 直接为0。

损失函数根据 Lin 等^[20]提出的改进的交叉熵损失函数(Cross-Entropy, CE), 即焦点损失函数(Focal Loss, FL)设计, 它能够解决序列标注不平衡的问题, 在情感原因定位模型中, 情感原因关键字显然要比非情感原因少, 损失函数计算如公式(18)-公式(20)所示。

$$loss_{start} = \sum_{j=1}^m \sum_{i=1}^n -\alpha y_{ji} \log^{z_j} (1 - \hat{y}_{ji})^z - (1 - \alpha) (1 - y_{ji}) \log^{1-z_j} \hat{y}_{ji} \quad (18)$$

$$loss_{end} = \sum_{j=1}^m \sum_{i=1}^n -\alpha z_{ji} \log^{z_j} (1 - \hat{z}_{ji})^z - (1 - \alpha) (1 - z_{ji}) \log^{1-z_j} \hat{z}_{ji} \quad (19)$$

$$loss = loss_{start} + loss_{end} \quad (20)$$

其中, α 和 λ 为可调参数; 因为是双标注序列输出, 所以有 $loss_{start}$ 和 $loss_{end}$; y_{ji} 和 z_{ji} 为真实序列; \hat{y}_{ji} 和 \hat{z}_{ji} 为模型预测输出; n 为模型序列长度; m 为批处理(batch)大小。

(5) 模型解码部分

最后输出两列浮点数序列, 一列预测情感原因起始位置, 另外一列预测结束位置, 如公式(21)和公式(22)所示。

$$P_1 = \{f_1, \dots, f_n\} \quad (21)$$

$$P_2 = \{l_1, \dots, l_n\} \quad (22)$$

P_1 、 P_2 分别表示起始位置和终止位置的打分。为了得到最终的情感原因定位, 设置最大情感原因区间 $maxlen$, 遍历输出序列长度不超过 $maxlen$ 的区间, 分别计算 P_1 、 P_2 的起始位置和终点位置的打分, 实验使用和、乘积或积的平方根进行计算, 发现使用乘积效果最好, 最后选择得到的最大值区间作为模型预测的情感原因结果。

4 实验部分

4.1 实验数据集

本文使用文献[14]所提供的新闻数据集, 每个文档标记了一条或多条情感句以及相应的情感原因句, 数据集中不同数量的情感原因对的文档所占的比例如表1所示。

表1 具有不同情感-原因对数量的文档的比例
Table 1 The Proportion of Documents with Different Number of Emotion-Cause Pairs

数据集	数量	比例
只有一对情感-原因对的文本	1 746	89.77%
有两对情感-原因对的文本	177	9.10%
超过两对情感-原因对的文本	22	1.13%
总计	1 945	100%

表1在此基础上, 为了抽取情感原因的关键因素, 在实验数据集做了如图6所示的情感原因关键字标注。图6左边 happiness 表示第2句为带有喜悦

的情感语句,并且喜悦的原因也在第 2 句,引起喜悦的关键原因是由于女儿的出生;图 6 右边 sadness 表示第 5 句为带有难过倾向的情感语句,关键原因是由于他们(指父母)还在为我操心。*中间所标注的词表示情感原因的关键字,EDGCNN 模型会抽取这些情感原因关键字。

随机选择数据集的 10% 作为实验测试集,其余作为训练集。为了获得更加可信的结果,进行 5 次实验,取平均值。实验采用正确率(P)、召回率(R)、F1 值(F_1)作为评判,如公式(23)-公式(25)所示。

$$P = \frac{\sum correct_pairs}{\sum proposed_pairs} \quad (23)$$

$$R = \frac{\sum correct_pairs}{\sum annotated_pairs} \quad (24)$$

$$F_1 = \frac{2PR}{P + R} \quad (25)$$

其中, *correct_pairs* 表示模型预测和标记的情感-原因对一致的个数; *proposed_pairs* 表示模型预测的情感-原因对个数; *annotated_pairs* 表示测试集标记好的情感-原因对个数。

1.null,null,49年前 2.happiness,喜悦.*女儿的出生*为黄魁吾夫妇带来了无与伦比的喜悦 3.null,null,可没多久 4.null,null,两人渐渐发现女儿似乎与其他孩子不同 5.null,null,一天天长大的她 6.null,null,不会坐 7.null,null,不会爬 8.null,null,甚至不会自己翻身 9.null,null,奔波了多家医院后 10.null,null,夫妇俩不得不接受女儿患有先天性脑瘫的事实	1.null,null,真的不想治了 2.null,null,我的家庭承受不起 3.null,null,一直都没有能力让父母过上好日子 4.null,null,现在还要让*他们为我操心* 5.sadness,惭愧,很惭愧 6.null,null,我知道即使手术后风险还是很大 7.null,null,所以就不想治了
---	---

图 6 情感原因关键字标注示例
Fig.6 Example of Emotion Cause Keyword Tagging

4.2 实验参数设置

词向量为北京师范大学与北京神州泰岳软件股份有限公司人工智能联合实验室^[21]提供的搜狗新闻训练的 300 维词向量。情感判别模型 CNN 卷积窗口设置为 2, 3, 4, 卷积核个数为 100 个。情感-原因对定位模型膨胀卷积网络卷积窗口设置为 3, 膨胀率一般为指数上升, 选择 1, 2, 4, 8。Adam 具有性能优秀、效率高的优点, 适用于梯度稀疏或梯度存在很大噪声的情况, 所以模型选择它作为优化器。为了防止过拟合, 加入了 dropout 优化策略, 通常设为 0.3, 批处理大小都为 30, EDGCNN 情感原因关键字最大长度 $maxlen=8, 10, 12$ 。

4.3 实验结果及分析

(1) 情感句判别

首先进行情感句抽取, 由于 LSTM^[22]和 CNN 都能较好地对本进行情感分类, 所以选择 LSTM 和 CNN 进行对比, 对句子进行二分类, 找出文本中的情感句。为了平衡训练数据集, 随机选择训练集文

本中非情感句作为负样本, 使得训练集中情感句和非情感句数量相当, 进行 5 次实验求取平均结果, 如表 2 所示。

表 2 情感句判别结果

Table 2 Result of Emotional Sentences Discrimination

实验模型	情感句判别结果		
	P	R	F_1
CNN	0.734 0	0.876 9	0.799 1
LSTM	0.674 4	0.824 6	0.741 9

实验对测试集中句子长度小于 3 的句子不进行情感判别。从表 2 可以发现, 由于是对文本中的每一个句子都进行情感判别, 且 CNN 更容易抓取到短文本上的情感关键词, 池化层也会过滤掉模型认为不重要的特征, 所以效果优于 LSTM, 之后将得到的情感句作为下一步情感原因抽取模型的输入。

(2) 情感原因抽取

ECPE 任务最近才被提出, 相关研究还不多, 目前有三种代表性方法^[14], 包括独立多任务学习

Indep、交互式原因情感多任务学习 Inter-CE 和交互式情感原因多任务学习 Inter-EC, 而 Inter-EC 在 ECPE 任务上性能最好。Indep、Inter-CE 和 Inter-EC 实际上都是句子级别的学习模型。

Indep 模型具有两层 Bi-LSTM 网络, 第一层的 Bi-LSTM 提取语句级别特征; 第二层中一个任务是情感提取, 另一个任务是原因提取, 每个 Bi-LSTM 为句子级别, 接收 d 个句子表示, 然后把 d 个句子特征作为输入接一个 Softmax 对该子句进行分类。

Inter-CE 是在 Indep 基础上, 把预测得到的情感原因标签嵌入到第一层的语句级别特征, 然后共同作为第二层的输入来预测情感句分类。

Inter-EC 与 Inter-CE 相反, 用预测得到的情感标签融入到第一层抽取的语句特征来预测情感原因句分类。

为验证本文模型的有效性, 将最后抽取结果与上述三种代表性方法进行对比, 实验结果如表 3 所示。

表 3 实验结果对比
Table 3 Experimental Results

实验模型	情感原因对提取结果		
	P	R	F_1
Indep	0.683 2	0.508 2	0.581 8
Inter-CE	0.690 2	0.513 5	0.590 1
Inter-EC	0.672 1	0.570 5	0.612 8
EDGCNN	0.575 8	0.635 2	0.604 5

从表 3 可以看出, EDGCNN 模型拥有最高的召回率, 而 F_1 值和现有最好模型在性能上相当。Indep 将情感句识别和原因句识别当作两个独立的子任务, 上层的两个 Bi-LSTM 分量没有特征交互, 而 Inter-CE 和 Inter-EC 在上层 Bi-LSTM 进行特征融合, 性能更好, 结果也表明情感特征更容易影响原因句的抽取。但是这三种模型最后的过滤器没有考虑到上下文的信息, 会过滤掉很多正确的情感-原因对, 所以在召回率上不如 EDGCNN。F1 值上, EDGCNN 与现有最好模型 Inter-EC 性能上相当, 也说明了本文提出的方法的有效性。此外, 特别需要指出的是 EDGCNN 可以进行更加细粒度的情感原因抽取, 部分情感原因关键字抽取结果如表 4 所示。

ECPE 任务是从文本中找出情感-原因对, 从表

表 4 EDGCNN 模型部分实验结果

Table 4 Experimental Results of EDGCNN Model

文本中部分语句	情感句	EDGCNN 输出	情感-原因对
8.null, 白金跃陷入回忆 9.happiness, 并略显激动	9	跃陷入回忆	9-8
3.happiness, 她为自己再过几天就可基本康复出院而感到高兴	3	康复出院	3-3
4.happiness, 让郭惠英感动的是 5.null, 她醒来时 6.null, 还有一位民警守在身边	4	民警守在身边	4-6
1.null, 每当在新闻上看到血库告急的消息 2.fear, 刘淑梅就很焦虑	2	血库告急的消息	1-2
6.null, 当唐某向张某提出离婚后 7.disgust, 怨恨转化成杀意	7	提出离婚	7-6
19.null, 汤某也完全站在女儿一边 20.anger, 这让陈晓毅很是气愤	20	数落女婿	20-19

4 可以看出, EDGCNN 模型会根据上下文语义关系, 利用情感判别模型识别出的情感句, 融入情感特征, 最后解码输出的情感原因关键字定位在文本中哪条语句, 就表明这条语句为相应的情感原因句。例如表 4 中第三行, 模型解码输出“康复出院”, 所以第 3 句就视为情感原因句, 得到情感-原因对 3-3。

实验中发现最大区间范围的选择会影响模型解码输出, 因此尝试使用区间最大长度分别为 8、10 和 12 进行实验, 实验中发现最大长度为 10 时效果最好, 部分情感原因关键字抽取结果如表 4 所示。另外, 第一步情感判别 CNN 模型会得到非情感句, 这些句子也会经过 EDGCNN 模型, 最后输出两列浮点数序列。为了过滤掉这些非情感句经过 EDGCNN 产生的结果, 模型的解码设定了阈值, 本文实验选择浮点数序列中的值大于 0.15 时, 才进行相应位置打分的计算。

总的来说, 本文 EDGCNN 模型具有较高的召回率, 可以更全面地挖掘文本中的情感-原因对。此外, EDGCNN 模型能抽取更细粒度的情感-原因对, 从而精细完成情感-原因对抽取任务。

5 结语

情感-原因对提取是情感原因挖掘的新任务, 为

了有效解决这一任务,并且从更细粒度的角度分析情感原因,本文首先通过情感句判别模型CNN找出带有情感的语句,然后将情感句编码,输入到本文融入情感特征的情感原因定位模型EDGCNN,找出相应的情感原因关键字,得到情感-原因对。与已有三种模型相比,本文模型召回率最高,并且在F1值性能上与现有最好模型结果相当,说明本文模型的有效性。在未来的工作,笔者将尝试构造一个端对端的模型,直接一步获取情感-原因对。

参考文献:

- [1] Rout J K, Choo K K R, Dash A K, et al. A Model for Sentiment and Emotion Analysis of Unstructured Social Media Text[J]. Electronic Commerce Research, 2018, 18: 181-199.
- [2] Peng M L, Zhang Q, Jiang Y G, et al. Cross-Domain Sentiment Classification with Target Domain Specific Information[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 2505-2513.
- [3] 刘全, 梁斌, 徐进, 等. 一种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. (Liu Quan, Liang Bin, Xu Jin, et al. A Deep Layered Network Model for Aspect-Based Sentiment Analysis[J]. Chinese Journal of Computers, 2018, 41(12): 2637-2652.)
- [4] 张庆庆, 贺兴时, 王慧敏, 等. 基于深度信念网络的文本情感分类研究[J]. 数据分析与知识发现, 2019, 3(4): 71-79. (Zhang Qingqing, He Xingshi, Wang Huimin, et al. Text Sentiment Classification Based on Deep Belief Network[J]. Data Analysis and Knowledge Discovery, 2019, 3(4): 71-79.)
- [5] Chen Y, Hou W J, Cheng X Y, et al. Joint Learning for Emotion Classification and Emotion Cause Detection[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 646-651.
- [6] Li X J, Feng S, Wang D L, et al. Context-Aware Emotion Cause Analysis with Multi-Attention-Based Neural Network[J]. Knowledge-Based Systems, 2019, 174: 205-218.
- [7] Lee S Y M, Chen Y, Huang C R, et al. Detecting Emotion Causes with a Linguistic Rule-Based Approach[J]. Computational Intelligence, 2013, 29(3): 390-416.
- [8] Yada S, Ikeda K, Hoashi K, et al. A Bootstrap Method for Automatic Rule Acquisition on Emotion Cause Extraction[C]// Proceedings of 2017 IEEE International Conference on Data Mining Workshops. IEEE, 2017: 414-421.
- [9] Gao K, Xu H, Wang J S. A Rule-Based Approach to Emotion Cause Detection for Chinese Micro-Blogs[J]. Expert Systems with Applications, 2015, 42(9): 4517-4528.
- [10] Chen Y, Hou W J, Cheng X Y. Hierarchical Convolution Neural Network for Emotion Cause Detection on Microblogs[C]// Proceedings of the 27th International Conference on Artificial Neural Networks. 2018: 115-122.
- [11] Gui L, Hu J N, He Y L, et al. A Question Answering Approach for Emotion Cause Extraction[C]// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1593-1602.
- [12] Weston J, Chopra S, Bordes A. Memory Networks[OL]. arXiv Preprint, arXiv: 1410.3916.
- [13] Yu X Y, Rong W G, Zhang Z, et al. Multiple Level Hierarchical Network-Based Clause Selection for Emotion Cause Extraction [J]. IEEE Access, 2019, 7: 9071-9079.
- [14] Xia R, Ding Z X. Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019: 1003-1012.
- [15] 苏剑林. 基于CNN的阅读理解式问答模型[EB/OL]. [2018-04-15]. <https://spaces.ac.cn/archives/5409>. (Su Jianlin. CNN Based Reading Comprehension Question and Answer Model [EB/OL]. [2018-04-15]. <https://spaces.ac.cn/archives/5409>.)
- [16] Kim Y. Convolutional Neural Networks for Sentence Classification[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014: 1746-1751.
- [17] Laskar Z, Kannala J. Context Aware Query Image Representation for Particular Object Retrieval[C]// Proceedings of Scandinavian Conference on Image Analysis. 2017: 88-99.
- [18] Dauphin Y N, Fan A, Auli M, et al. Language Modeling with Gated Convolutional Networks[C]// Proceedings of the 34th International Conference on Machine Learning. 2017: 933-941.
- [19] Wang W H, Yang N, Wei F R, et al. Gated Self-Matching Networks for Reading Comprehension and Question Answering [C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017: 189-198.
- [20] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980-2988.
- [21] Li S, Zhao Z, Hu R F, et al. Analogical Reasoning on Chinese Morphological and Semantic Relations[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018: 138-143.
- [22] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

作者贡献声明:

代建华:提出研究思路,设计研究方案,起草论文,最终版本修订;
邓育彬:采集、清洗和分析数据,进行实验,起草论文,最终版本修订。

支撑数据:

支撑数据由作者自存储,E-mail:jhdai@hunnu.edu.cn。
[1] 邓育彬. EDGCNN.txt. 情感-原因对话料数据集。

利益冲突声明:

所有作者声明不存在利益冲突关系。

收稿日期:2019-11-14
收修改稿日期:2020-05-29

Extracting Emotion-Cause Pairs Based on Emotional Dilation Gated CNN

Dai Jianhua^{1,2,3} Deng Yubin^{1,3}

¹(Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China)

²(Research Institute of Languages and Cultures, Hunan Normal University, Changsha 410081, China)

³(College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China)

Abstract: **[Objective]** This paper proposes an Emotional Dilation Gated CNN (EDGCNN) model, aiming to extract emotion-cause pairs for sentiment analysis. **[Methods]** First, we used the emotional discriminant model to identify sentiment sentences. Then, we input coding for these sentences to the EDGCNN model and located corresponding reasons. Finally, we tagged keywords of reasons generated from the experimental dataset. **[Results]** The new model's recall and F1 values reached 63.52% and 60.45% respectively on the training dataset, which were better or very similar to the existing ones. The proposed model also extracted emotion-cause pairs at finergranularity level. **[Limitations]** The experimental corpus size was small. **[Conclusions]** The proposed model can extract emotion-cause pairs effectively.

Keywords: Emotion-Cause Pair Extraction EDGCNN Emotion Discrimination