

# 相互作用的情感与原因子句提取模型

陆丁天<sup>1</sup> 张志远<sup>1\*</sup> 刘佩佩<sup>2</sup> 于得水<sup>3</sup>

<sup>1</sup>(中国民航大学计算机科学与技术学院 天津 300300)

<sup>2</sup>(中国民航大学校报编辑部 天津 300300)

<sup>3</sup>(中国民航大学科技创新研究院 天津 300300)

**摘要** 无情感标注的原因提取获得较多关注。由于情感与原因是相互影响的,基于此,提出相互作用的情感与原因子句提取模型,加入相互作用的注意力矩阵提取情感与原因之间的关联性,进而提高情感子句与原因子句的提取效果。实验结果表明相互作用的注意力矩阵能显著提高模型效果。

**关键词** 情感子句 原因子句 情感分析 注意力 相互作用矩阵

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.06.030

## INTERACTIVE EMOTION AND CAUSE CLAUSE EXTRACTION MODEL

Lu Dingtian<sup>1</sup> Zhang Zhiyuan<sup>1\*</sup> Liu Peipei<sup>2</sup> Yu Deshui<sup>3</sup>

<sup>1</sup>(School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

<sup>2</sup>(Editorial Department of Journal, Civil Aviation University of China, Tianjin 300300, China)

<sup>3</sup>(Research Institute of Science and Technology Innovation, Civil Aviation University of China, Tianjin 300300, China)

**Abstract** More and more attention has been paid to the cause extraction without emotion tagging. Because emotion and cause interact with each other, this paper proposes an interactive emotion and cause clause extraction model. An interactive attention matrix was added to extract the correlation between emotion and cause, so as to improve the extraction effect of emotion clause and cause clause. The experimental results show that the interaction attention matrix can significantly improve the effect of the model.

**Keywords** Emotion clause Cause clause Emotional analysis Attention Interaction matrix

## 0 引言

情感原因提取最早是由 Lee 等<sup>[1]</sup>提出的以表述一个事件的词作为情感原因提取单元的任务。此后,在将该任务与机器学习相结合的研究中,Gao 等<sup>[2]</sup>提出了一个包含原因事件的情感模型,在规则算法基础上,提取情感词相对应的原因事件词。随着深度学习的兴起,以词作为情感原因提取单元的情感原因提取任务需要对语料库进行大量复杂的人工标注,语料库规模一般都比较小,无法满足目前深度训练模型所需要信息。Gui 等<sup>[3]</sup>使用新浪都市新闻建立了一个标注好的

中文情感原因发现语料库,将子句作为原因提取的基本单元,并在文档中预先标注好情感关键词。Gui 等<sup>[4]</sup>也首次将深度学习方法与情感原因提取任务相结合,利用卷积操作的记忆网络在先前构造的语料库上完成情感原因提取任务。该语料库能满足深度学习所需的信息,目前大部分情感原因提取任务的研究都是以该语料库作为数据源<sup>[5-10]</sup>。Li 等<sup>[11]</sup>提出了一种基于多注意力的神经网络模型(MANN 模型),将已标注的情感信息融入文档子句中进行原因子句提取,利用注意力来捕捉已标注的情感分句与原因候选分句之间的联系。

预先标注情感信息再进行原因提取的方式需要在

收稿日期:2021-03-12。空中交通管理系统与技术国家重点实验室开放基金资助项目(SKLATM201902);天津市教委科研项目(2019SK110);中央高校基本科研业务费中国民航大学专项(3122019009)。陆丁天,硕士生,主研领域:自然语言处理,文本情感分析。张志远,教授。刘佩佩,硕士。于得水,硕士。

数据集中人工标注情感信息,这增加了人工成本,也限制了原因提取任务在现实环境中的应用。为此, Xia 等<sup>[12]</sup>提出新的情感原因提取任务:给定一篇未标注情感信息的文档,同时提取其中的情感子句及对应的原因子句,并提出 Inter-EC 模型来进行该任务,为情感原因提取的研究提供了新思路。然而 Inter-EC 模型仅将情感子句预测结果嵌入到候选原因子句特征中进行原因提取,这只是单向地表现情感与原因之间的联系:情感→原因,而不能体现出情感与原因之间的相互关联性:情感↔原因。为了将情感与原因之间的相互关联性进一步体现,进而提高提取效果,本文提出一种相互作用的情感子句与原因子句提取模型 IECEM。该模型利用词水平 Bi-LSTM 捕获子句内词之间的上下文依赖特征,并利用注意力机制分别捕获情感表达和原因表达的关键语义表示;在此基础上,将子句特征输入句子水平 Bi-LSTM 来捕获文档中子句之间的上下文语义依赖关系;最后使用情感-原因相互作用矩阵对子句特征进行操作,使子句特征包含情感与原因相互联系的语义信息。此外,为增强情感与原因的联系,在原因子句提取部分,先将情感子句的预测结果嵌入到子句特征中,再与原因相互作用矩阵进行操作,最后进行原因子句预测。

# 1 模型介绍

## 1.1 任务描述

设计深度学习模型,使模型能提取一篇文档中的情感子句与原因子句。如图 1 所示,这篇文档有  $c_1$  到

$c_7$  共 7 个子句(以标点符号结尾的句子算一个子句),文档输入模型后,分别输出每个子句的情感子句标签和原因子句标签,其中: $y_i^e$  ( $i = 1, 2, \dots, 7$ ) 为情感子句标签, $y_i^c$  为原因子句标签;0 表示否,1 表示是。该文档的情感子句为  $c_4$ ,原因子句为  $c_3$ 。



图 1 任务描述示例

## 1.2 IECEM 模型

IECEM 模型由情感子句提取和原因子句提取两部分组成。模型整体结构如图 2 所示,情感子句提取部分首先使用词水平的 Bi-LSTM 与注意力机制对子句的词向量进行编码,获取子句内词之间的上下文语义依赖关系及情感表达的关键语义信息;再将编码后的子句特征输入句子水平的 Bi-LSTM 以获取子句之间的上下文语义依赖;最后与情感-原因相互作用矩阵进行操作,获取情感与原因之间的关联特征,使用 Softmax 得到情感子句预测结果。原因子句提取部分与情感子句提取部分大体相同,只是在子句特征输入句子水平的 Bi-LSTM 前,将情感子句提取部分的预测结果嵌入到子句特征中,以增强情感与原因的联系。

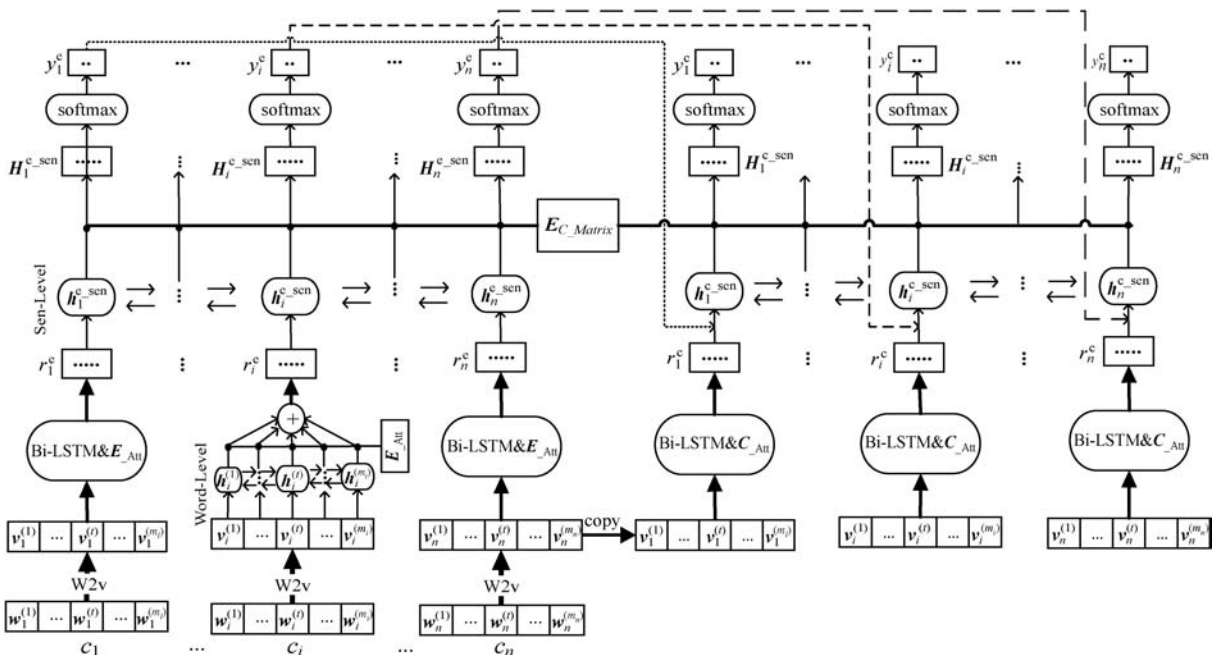


图 2 IECEM 模型

### 1.2.1 情感子句提取

给定一篇文档  $d_{oc} = \{c_1, c_2, \dots, c_i, \dots, c_n\}$ , 其中  $c_i$  表示文档中的第  $i$  个子句,  $n$  为这篇文档的长度。子句  $c_i$  由多个词组成, 表示为  $c_i = \{w_i^{(1)}, w_i^{(2)}, \dots, w_i^{(t)}, \dots, w_i^{m_i}\}$ ,  $w_i^t$  表示子句  $c_i$  中的第  $t$  个词,  $m_i$  表示子句的长度;  $w^{2v} = \{v_i^{(1)}, v_i^{(2)}, \dots, v_i^{(t)}, \dots, v_i^{(m_i)}\}$  为子句  $c_i$  中词对应的词向量表示,  $v_i^{(t)}$  表示第  $t$  个词对应的词向量。将  $w_{2v}$  输入词水平的 Bi-LSTM 中, 获取子句  $c_i$  词之间上下文依赖关系的子句特征  $h^e = \{h_i^{(1)}, h_i^{(2)}, \dots, h_i^{(t)}, \dots, h_i^{(m_i)}\}$ , 为捕获情感表达的关键部分, 使用注意力矩阵  $E_{Att}$  与  $h$  中每个中间状态  $h_i^{(t)}$  ( $t = 1, 2, \dots, m_i$ ) 进行如下操作, 得到子句  $c_i$  最新的特征表示  $r_i^e = \{a_{h_i^{(1)}}, a_{h_i^{(2)}}, \dots, a_{h_i^{(t)}}, \dots, a_{h_i^{(m_i)}}\}$ 。

$$u_i^{(t)} = \tanh(W_a h_i^{(t)} + b_a) \quad (1)$$

$$a_i^{(t)} = \frac{\exp(u_i^{(t)} \cdot E_{Att})}{\sum_{j=1}^{m_i} \exp(u_i^{(j)} \cdot E_{Att})} \quad (2)$$

$$a_{h_i^{(t)}} = a_i^{(t)} h_i^{(t)} \quad (3)$$

式中:  $W_a$  为权重矩阵;  $b_a$  为偏置项;  $E_{Att}$  为一个  $-0.01 \sim 0.01$  范围内随机初始化的正态分布矩阵表示上下文语义, 用以捕获子句中情感表达的关键语义信息;  $\cdot$  为矩阵相乘操作。  $W_a$ 、 $b_a$  与  $E_{Att}$  在模型训练过程中不断变化(参考文献[13])。  $E_{Att}$  的作用如图 3 所示,  $w_{2v}$  与  $E_{Att}$  进行操作后, 能捕获“感激”这个情感词的信息, 对应增加该词的权重, 图 3 中灰色表示权重增加。

输入文本: 她非常感激

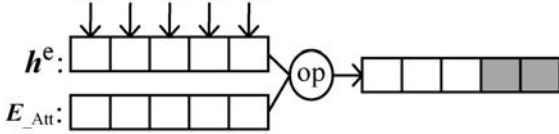


图 3  $E_{Att}$  作用示意图

为获取子句  $c_i$  与其他子句之间的上下文语义依赖关系, 将  $r_i^e$  输入到句子水平的 Bi-LSTM 中进行编码, 并使用隐藏层状态  $h_i^{e-sen}$  表示子句  $c_i$  与其他子句间的上下文依赖特征。然后使用情感-原因相互作用矩阵  $E_{C\_Matrix}$  与  $h_i^{e-sen}$  进行如下操作:

$$H_i^{e-sen} = (E_{C\_Matrix} \cdot h_i^{e-sen}) + b_{ec} \quad (4)$$

使得  $H_i^{e-sen}$  包含了情感与原因的相互联系特征。式中: 情感-原因相互作用矩阵  $E_{C\_Matrix}$  为一个随机初始化的特征矩阵, 在训练过程中学习情感与原因之间的相互关系并作用到情感子句提取和原因子句提取上;  $b_{ec}$  为偏置项。

情感-原因相互作用矩阵  $E_{C\_Matrix}$  的作用如图 4 所示, 图 4 以图 1 任务描述为例。  $h_1^{e-sen}$  到  $h_7^{e-sen}$  分别对应

情感子句提取部分子句  $c_1$  到子句  $c_7$  经过句子水平 Bi-LSTM 后的隐藏层状态;  $h_1^{e-sen}$  到  $h_7^{e-sen}$  分别对应原因子句提取部分子句  $c_1$  到子句  $c_7$  经过句子水平 Bi-LSTM 后的隐藏层状态。两者同时与情感-原因相互作用矩阵  $E_{C\_Matrix}$  进行操作, 使情感子句和原因子句对应的隐藏层状态权重增大。任务描述中,  $c_3$  为原因子句,  $c_4$  为情感子句, 故经过与  $E_{C\_Matrix}$  进行操作后对应的隐藏层状态  $H_3^{e-sen}$  与  $H_4^{e-sen}$  权重增加, 图 4 中灰色部分表示权重增加。

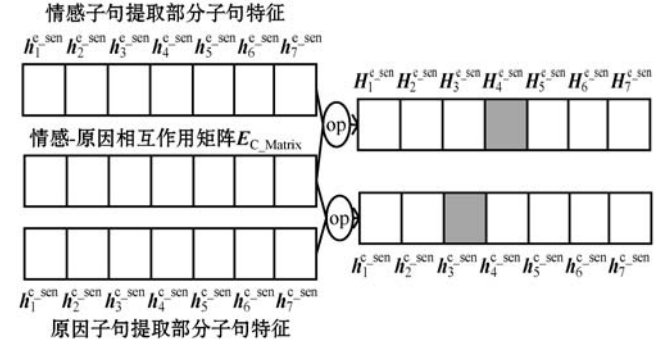


图 4  $E_{C\_Matrix}$  作用示意图

最后将每个  $H_i^{e-sen}$  进行 softmax 得到情感子句预测标签  $y_i^e$ :

$$y_i^e = \text{softmax}(W^e H_i^{e-sen} + b^e) \quad (5)$$

### 1.2.2 原因子句提取

在原因子句提取部分, 一些能表达原因的关联词对原因提取有着重要作用。由此利用经过词水平 Bi-LSTM 编码后的子句特征与  $C_{Att}$  注意力矩阵进行操作来提高对原因表达有作用的词的权重, 操作过程同式(1) - 式(3)。  $C_{Att}$  的作用如图 5 所示, 子句特征在与  $C_{Att}$  矩阵进行操作后, “因为”这个词的权重增加, 灰色表示权重增加。

输入文本: 因为得知整个事件经过后

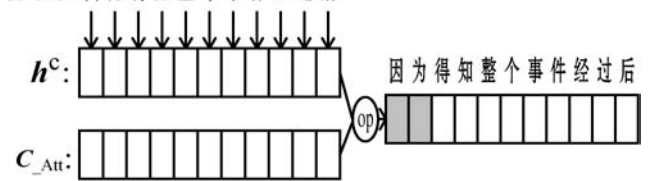


图 5  $C_{Att}$  的作用

子句  $c_i$  在通过词水平的 Bi-LSTM 和与  $C_{Att}$  矩阵进行操作后得到新的特征向量  $r_i^c$ , 为了增强情感与原因的关联性, 提高原因子句的提取效果, 先将情感子句提取部分的预测结果嵌入  $r_i^c$  中:

$$R_i^c = r_i^c \oplus y_i^e \quad (6)$$

式中:  $\oplus$  表示把  $y_i^e$  拼接在  $r_i^c$  后。再将  $R_i^c$  输入到句子水平的 Bi-LSTM 中, 获取其与其他子句之间的语义信息  $h_i^{e-sen}$ 。最后将  $h_i^{e-sen}$  与情感-原因相互作用矩阵  $E_{C\_Matrix}$

进行操作得到  $H_i^{c-sen}$ ,操作过程同式(4)。将  $H_i^{c-sen}$  进行 Softmax 便得到原因子句预测标签:

$$y_i^c = \text{softmax}(W^c H_i^{c-sen} + b^c) \quad (7)$$

模型总的损失函数等于情感子句提取部分的交叉熵损失加上原因子句提取部分的交叉熵损失:

$$L_{\text{oss}} = \lambda L_{\text{oss}}^e + (1 - \lambda) L_{\text{oss}}^c + \lambda_{\text{reg}} \|\Theta\|^2 \quad (8)$$

式中: $L_{\text{oss}}^e$ 为情感子句提取部分的交叉熵损失; $L_{\text{oss}}^c$ 为原因子句提取部分的交叉熵损失; $\lambda$ 为损失系数; $\Theta$ 为 L2 正则项。

## 2 实验与结果分析

### 2.1 实验数据集

IECEM 模型实验使用 Xia 等提出的新情感原因提取任务数据集,数据集构成如表 1 所示,数据集共由 1 945 篇文档构成,每篇文档又由多个子句组成。模型训练集与测试集占比为 9:1。

表 1 数据集统计信息

一篇文档中原因子句个数	文档数量	占整篇文档比例/%
1 个	1 746	89.77
2 个	177	9.10
多于 2 个	22	1.13
总计	1 945	100

### 2.2 模型参数

模型使用预训练好的微博语料库词向量,词向量维度为 300 维,对于未登录词,采用均匀分布  $U(-0.01, 0.01)$  来随机初始化词向量。一个子句的截断长度为 45 个词,一篇文档的截断长度为 75 个子句。所有 Bi-LSTM 的隐藏层单元个数为 100。模型学习率为 0.000 5,并使用 Adam 优化算法。训练数据的批次大小为 32,迭代次数为 15 次。模型学习率为 0.005,L2 正则项为  $10^{-5}$ 。

### 2.3 实验结果分析

#### 2.3.1 加入情感-原因相互作用矩阵

为验证情感-原因相互作用矩阵的有效性,模型在没有将情感子句预测结果嵌入原因子句的条件下,仅仅进行了使用情感-原因相互作用矩阵和不使用情感-原因相互作用矩阵的实验。连续进行 5 次实验,以减少实验的偶然性。实验结果如图 6 所示,图 6(a)为情感子句提取结果,图 6(b)为原因子句提取结果。

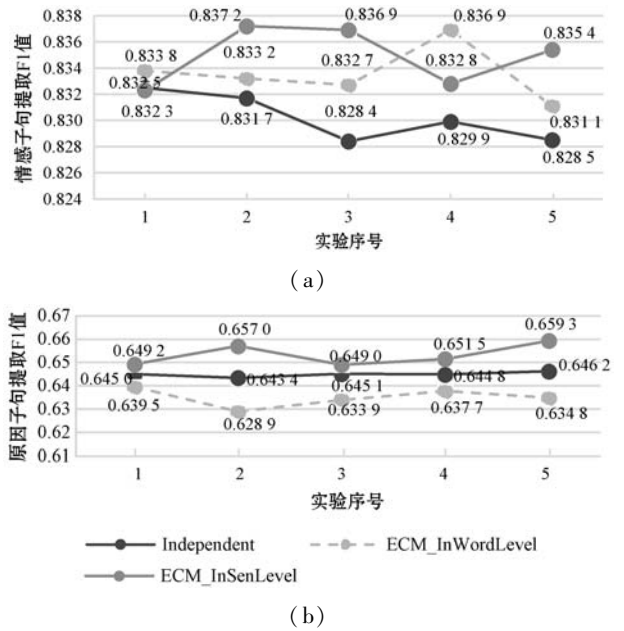


图 6 情感-原因相互作用矩阵效果

图 6 中 Independent 表示情感子句提取与原因子句提取独立进行,两者之间没有相互联系;ECM\_InwordLevel 表示将情感-原因相互作用矩阵作用于经过词水平 Bi-LSTM 编码后的子句特征;ECM\_InSenLevel 表示将情感-原因相互作用矩阵作用于经过句子水平 Bi-LSTM 编码后的子句特征。从图 6 可以看出在进行情感子句提取时,加入情感-原因相互作用矩阵的提取效果要比 Independent(情感子句与原因子句独立进行提取)条件下的效果好,这是由于情感-原因相互作用矩阵能捕捉情感与原因之间的关系,对情感子句的提取效果有促进作用。此外,情感-原因相互作用矩阵在 ECM\_InwordLevel 条件下的情感子句提取效果与在 ECM\_InSenLevel 条件下的提取效果相近,这说明在进行情感子句提取时,在有情感-原因相互作用矩阵的情况下,子句间的上下文语义依赖关系并不是太重要,模型在判断子句是否为情感子句时,会更多地依赖于子句中某个情感关键词,而非子句之间的上下文语义依赖特征。

图 6(b)为原因子句提取的 F1 值,可以看出,情感-原因相互作用矩阵在 ECM\_InwordLevel 条件下的原因子句提取效果比 Independent 条件下的提取效果还要差,说明情感-原因相互作用矩阵捕获到的信息作用于词水平 Bi-LSTM 编码后的原因子句特征时,给特征增加了噪声,进而影响了原因子句的提取效果。而情感-原因相互作用矩阵在 ECM\_InSenLevel 条件下的原因子句提取效果比 Independent 条件下的提取效果还要好,这说明情感-原因相互作用矩阵在一个独立的子句特征中捕获情感-原因信息有限,并不能作用在原因子句的提取上;而在拥有子句间上下文依赖关系的子

句特征中就能很好地捕捉到情感-原因信息。这也比较符合事实,因为仅仅依靠独立的子句是无法判断该子句是否为原因子句,还需要获取该子句与文档中其他子句间的上下文语义信息。

### 2.3.2 嵌入情感子句预测结果

为增加原因子句的提取效果,将情感子句预测结果嵌入到原因候选子句特征中。实验结果如图 7 所示。

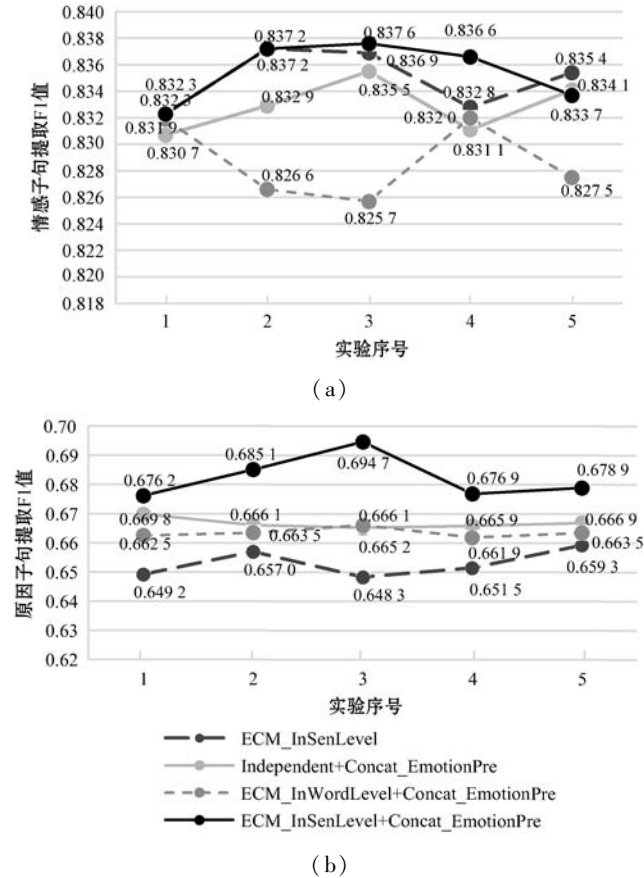


图 7 嵌入情感子句预测结果

图 7 中 CEM\_InSenLevel 为上一个实验中的 CEM\_InSenLevel(表示没有嵌入情感子句预测结果并将情感-原因相互作用矩阵作用于经过句子水平 Bi-LSTM 编码后的子句特征);Independent + Concat\_EmotionPre 表示没有加入情感-原因相互作用矩阵,仅将情感子句预测结果嵌入原因候选子句特征中;ECM\_InWordLevel + Concat\_EmotionPre 表示在上一个实验中 ECM\_InWordLevel 的基础上嵌入情感子句预测结果;ECM\_InSenLevel + Concat\_EmotionPre 表示在上一个实验中 ECM\_InSenLevel 的基础上嵌入情感子句预测结果。

从图 7(a)可以看出,实验结果与上一个实验相比,嵌入情感子句的预测结果对情感子句的提取并无明显的影响。但对于原因子句的提取却有明显的提高,从上一个实验的最好结果 0.6593 提高到了 0.6947,提高了 5.37%左右,说明嵌入情感预测结果对原因提取有效,表明在模型中考虑情感与原因间的联系是有

用的。此外 ECM\_InSenLevel + Concat\_EmotionPre 的原因子句提取 F1 值的最好结果 0.6947 比 Independent + Concat\_EmotionPre 的最好结果 0.6698 高出了 3.72%左右,这说明在嵌入情感子句预测结果的同时加入情感-原因相互作用矩阵(即情感→原因基础上再加上情感↔原因)对原因子句的提取有进一步促进作用。

### 2.3.3 嵌入原因子句预测结果

为增加情感子句提取效果,将原因子句预测结果嵌入到情感候选子句特征中。实验结果如图 8 所示。

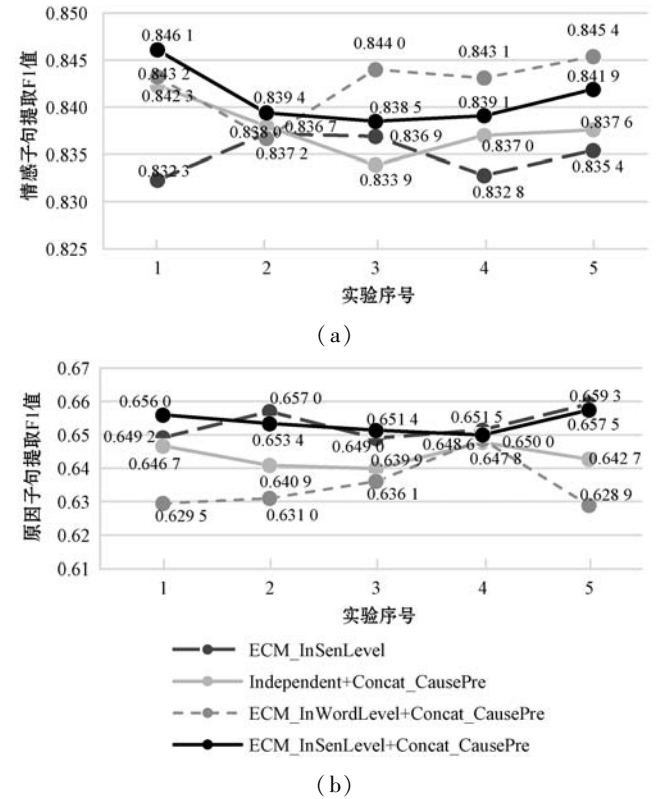


图 8 嵌入原因子句测试结果

其中 CEM\_InSenLevel 为实验 2.3.1 中的 CEM\_InSenLevel 部分;Independent + Concat\_CausePre 表示没加入情感-原因相互作用矩阵,仅将原因子句预测结果嵌入情感候选子句特征中;ECM\_InWordLevel + Concat\_CausePre 表示在实验 2.3.1 中 ECM\_InWordLevel 的基础上嵌入原因子句预测结果;ECM\_InSenLevel + Concat\_CausePre 表示在实验 2.3.1 中 ECM\_InSenLevel 的基础上嵌入原因子句预测结果。

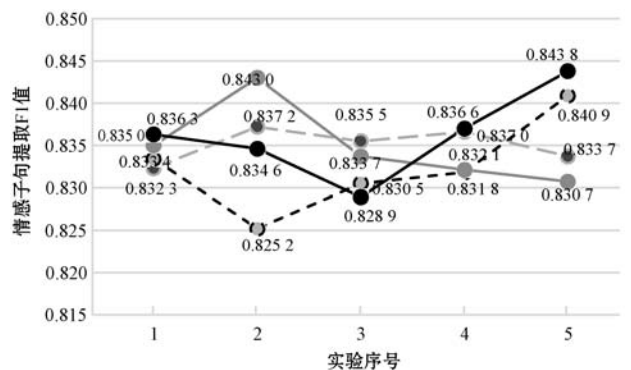
从图 8 中可以看出,嵌入原因子句预测结果,原因子句的提取结果与实验 2.3.1 并没有明显差距,但能提高情感子句的提取效果,而且 ECM\_InSenLevel + Concat\_CausePre 的提取效果比 Independent + Concat\_CausePre 更好,这说明在嵌入原因子句预测结果的同时加入情感-原因相互作用矩阵(即原因→情感基础上再加上情感↔原因)对情感子句的提取有进一步促进作用。

总体而言,把情感子句预测结果嵌入原因候选子句特征中对原因子句提取的效果比把原因子句预测结果嵌入情感候选子句特征中对情感子句提取的效果要好。这证实了文献[12]中的一个结论:在知道情感的情况下,更容易找出原因;但仅知道原因却不能很容易地确定情感。因此本文模型采用情感子句预测结果嵌入原因候选子句特征的方式来原因子句的提取效果。

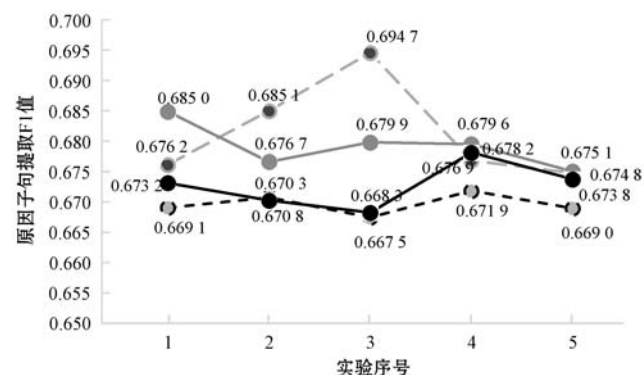
2.3.2节与2.3.3节的实验结果表明情感-原因相互作用矩阵与句子水平 Bi-LSTM 编码后的子句特征进行操作的模型效果更好,故模型采用此方式进行后续实验。

### 2.3.4 子句特征与情感-原因相互作用矩阵操作方式

将子句特征与情感-原因相互作用矩阵通过不同的运算方式进行实验,以获取更好的提取效果。实验结果如图9所示。



(a)



(b)

图9 不同运算方式实验结果

图9中 ECM\_InSenLevel + Concat 为实验2.3.2中的 ECM\_InSenLevel + Concat\_EmotionPre, 其子句特征与情感-原因相互作用矩阵的运算方式为式(4);图9中的“+”运算表示子句特征与情感-原因相互作用矩阵的运算方式由式(4)变为:

$$H_i^{e\_sen} = E_{C\_Matrix} + h_i^{e\_sen} \quad (9)$$

图9中的“\*”与“+”运算表示运算方式由式(4)变为:

$$H_i^{e\_s} = (E_{C\_Matrix} \cdot h_i^{e\_sen}) + b_{ec} \quad (10)$$

$$H_i^{e\_sen} = H_i^{e\_s} + h_i^{e\_sen} \quad (11)$$

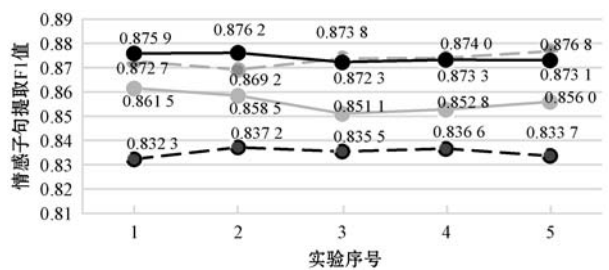
图9中的“\*”和“softmax”和“+”表示运算方式由式(4)变为:

$$H_i^{e\_sen} = \text{softmax}(E_{C\_Matrix} \cdot h_i^{e\_sen}) + h_i^{e\_sen} \quad (12)$$

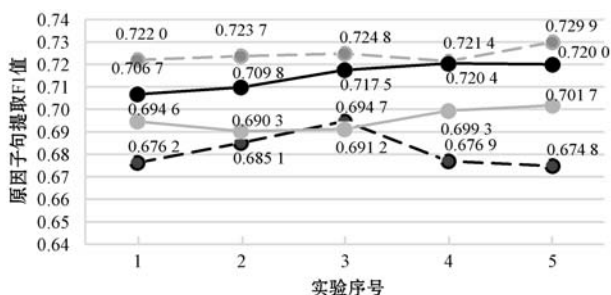
图9中的结果说明子句特征与情感-原因相互作用矩阵以不同方式进行运算,对情感子句与原因子句的提取都有不同的影响,且同一运算方式对两部分提取的作用效果也不成正比,如“\*”和“softmax”和“+”运算在情感子句提取上取得的最好结果比其他运算方式的高,但在原因子句的提取上,其取得的结果不是最好的,反而是 ECM\_InSenLevel + Concat 的运算方式取得的结果最好。这是由于情感子句提取注重的特征与原因子句提取注重的特征不一样,在同一种运算方式下,可能情感子句特征得到放大,进而促进情感子句提取效果;但原因子句特征放大得不够,进而不能很好地促进原因子句的提取效果。

### 2.3.5 不同词向量对模型的影响

不同语料库预训练生成的词向量不同,为验证不同语料库生成的词向量对模型提取效果的影响。为此,模型使用不同词向量进行实验,实验结果如图10所示。图10中,ECM\_InSenLevel + Concat 为使用300维的中文微博语料库预训练生成的词向量;W2Vec\_By\_wiki 为使用300维的维基百科语料库预先生成的词向量;W2Vec\_By\_renmin 为使用300维的人民日报语料库预训练生成的词向量;W2Vec\_By\_SogouNews 为使用300维的搜狗新闻语料库生成的词向量。词向量的生成参见文献[14-15]。



(a)



(b)

图10 不同词向量的提取效果

从实验结果可以看出,在情感子句提取上,模型使用搜狗新闻词向量的最好 F1 值 0.876 2,比使用微博词向量的最好 F1 值 0.837 2 高了 4.66% 左右;在原因子句提取上,使用搜狗新闻词向量最好 F1 值 0.720 4 比使用微博词向量最好 F1 值 0.694 7 高了 3.7% 左右。这说明不同词向量对模型提取效果的影响很明显。从整体上看,使用人民日报词向量与搜狗新闻词向量的提取效果更相近也更好,这是由于模型使用的数据集由新浪都市新闻构成,而人民日报和搜狗新闻也属于新闻,在词性词义上会更贴合数据集,且用这两种新闻训练生成的词向量在很大程度上能更好地表示数据集的信息。由此,使用更贴近模型数据集的语料库生成的词向量能提高模型完成任务的效果。

### 2.3.6 与其他模型对比

与其他模型对比主要分为两部分,第一部分是与传统的情感原因提取任务模型比,这些模型都是在数据集中标注好情感信息的条件下进行原因提取;第二部分与新的情感原因提取任务定义下的模型作比较,该模型是在数据集未标注情感信息的前提下去提取情感子句与原因子句。对比结果如表 2 所示,其中对比的模型具体如下:

(1) RB: rule-based method, 基于规则匹配的情感原因发现方法<sup>[1]</sup>。

(2) CB: commonsense-based method, 基于情感常识库匹配方法<sup>[16]</sup>。

(3) RB + CB + ML: machine learning, 结合规则匹配、常识库方法和机器学习方法<sup>[3]</sup>。

(4) Multi-kernel: 基于支持向量机分类器的方式<sup>[3]</sup>。

(5) ConvMS-M: 基于卷积的层级记忆网络的情感原因提取方法<sup>[4]</sup>。

(6) PAE-DGL: 结合相对位置信息和全局标签信息的情感原因提取方法<sup>[7]</sup>。

(7) MANN: 基于上下文多注意力的情感原因分析方法<sup>[11]</sup>。

(8) Inter-EC: 新的情感原因提取任务方法,在未给出情感信息的情况下,同时提取情感子句与原因子句<sup>[12]</sup>。

表 2 与其他模型对比

模型	情感提取			原因提取		
	准确率	召回率	F1	准确率	召回率	F1
RB	—	—	—	0.671 2	0.524 7	0.589 0
CB	—	—	—	0.267 2	0.713 0	0.388 7
RB + CB + ML	—	—	—	0.592 1	0.530 7	0.559 7

续表 2

模型	情感提取			原因提取		
	准确率	召回率	F1	准确率	召回率	F1
Multi-kernel	—	—	—	0.658 8	0.692 7	0.675 2
ConvMS-Mem	—	—	—	0.707 6	0.683 8	0.695 5
PAE-DGL	—	—	—	0.761 9	0.690 8	0.724 2
MANN	—	—	—	0.784 3	0.758 7	0.770 6
Inter-EC	0.836 4	0.810 7	0.823 0	0.704 1	0.608 3	0.650 7
<b>IECEM</b>	<b>0.830 1</b>	<b>0.842 4</b>	<b>0.835 5</b>	<b>0.733 6</b>	<b>0.661 7</b>	<b>0.694 7</b>

表 2 中,从 RB 到 MANN 方法都为传统的情感原因提取任务方法,这些方法均在已给出情感信息的前提下提取原因,故对应的情感提取部分为“—”。可以看出,IECEM 模型在没有给出情感信息的前提下提取情感子句与原因子句所得原因提取效果仍优于大部分传统情感原因提取模型的原因提取效果,说明在今后的原因提取任务研究中不需要事先在数据集中标注情感信息,减少人工标注数据集的成本,进而拓展这项任务在现实中的应用范围。

在与新情感原因提取任务定义模型 Inter-EC 的对比中,Inter-EC 模型使用微博词向量进行实验,故在进行对比时,IECEM 模型也使用微博词向量进行实验,并提取原因子句提取 F1 值最好的结果。从表 2 中看出,IECEM 模型在情感子句提取的 F1 值上比 Inter-EC 的提高了 1.52% 左右;在原因子句提取的 F1 值上比 Inter-EC 的提高了 6.76% 左右。这是由于 IECEM 模型增加了情感-原因相互作用矩阵,在进行提取任务时融入了情感原因相互作用的信息(情感↔原因),比 Inter-EC 模型仅融入情感到原因(情感→原因)的信息获取的特征更全面,表明 IECEM 模型的提取效果比 Inter-EC 的提取效果好。

## 3 结 语

本文提出一种相互作用的情感与原因子句提取模型,利用情感与原因之间是有联系的这一性质,在进行情感子句与原因子句提取时,使用情感-原因相互作用矩阵来捕捉情感与原因之间的关联性来提高提取效果。实验结果表明,本文模型能提高情感子句与原因子句的提取效果且优于新情感原因提取任务定义下的 Inter-EC 模型。

在实验结果中,情感子句提取 F1 值提升并不明显,原因子句提取的 F1 值跟传统任务定义下的模型最

好结果还有一定的差距;且子句特征与情感-原因相互作用矩阵在同一运算方式下,对情感子句特征的放大与原因子句特征的放大效果也不成正相关。后续研究将聚焦这些问题,改进提取方式以进一步提高情感子句与原因子句的提取效果。

## 参 考 文 献

- [ 1 ] Lee S Y M, Chen Y, Huang C R. A text-driven rule-based system for emotion cause detection[C]//Workshop on Computational Approaches to Analysis and Generation of Emotion in Text,2010:45 - 53.
- [ 2 ] Gao K, Xu H, Wang J S. Emotion cause detection for Chinese micro-blogs based on ECOCC model[C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining,2015: 3 - 14.
- [ 3 ] Gui L, Wu D Y, Xu R F, et al. Event-driven emotion cause extraction with corpus construction[C]//Conference on Empirical Methods in Natural Language Processing,2016:1639 - 1649.
- [ 4 ] Gui L, Hu J N, He Y L, et al. A question answering approach for emotion cause extraction[C]//Conference on Empirical Methods in Natural Language Processing,2017:1593 - 1602.
- [ 5 ] 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.
- [ 6 ] Xia R, Zhang M R, Ding Z X. RTHN: A RNN-transformer hierarchical network for emotion cause extraction[C]//28th International Joint Conference on Artificial Intelligence, 2019:5285 - 5291.
- [ 7 ] Ding Z X, He H, Zhang M R, et al. From independent prediction to reordered prediction: Integrating relative position and global label information to emotion cause identification [J]. Proceedings of the AAAI Conference on Artificial Intelligence,2019,33:6343 - 6350.
- [ 8 ] 巫继鹏,鲍建竹,蓝恭强,等. 结合规则蒸馏的情感原因发现[J]. 清华大学学报(自然科学版),2020,60(5):422 - 429.
- [ 9 ] Tang H, Ji D H, Zhou Q J. Joint multi-level attentional model for emotion detection and emotion-cause pair extraction[J]. Neurocomputing,2020,409:329 - 340.
- [ 10 ] Wu S X, Chen F, Hunag Y F, et al. A multi-task learning neural network for emotion-cause pair extraction[J]. European Conference on Artificial Intelligence,2020,325:2212 - 2219.
- [ 11 ] 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.
- [ 12 ] Xia R, Ding Z X. Emotion-cause pair extraction: A new task to emotion analysis in texts[C]//57th Annual Meeting of the Association for Computational Linguistics,2019:1003 - 1012.
- [ 13 ] Zhang Y S, Zheng J, Jiang Y R, et al. A text sentiment classification modeling method based on coordinated CNN-LSTM-attention model[J]. Chinese Journal of Electronics, 2019,28(1):120 - 126.
- [ 14 ] Li S, Zhao Z, Hu R F, et al. Analogical reasoning on Chinese morphological and semantic relations[C]//56th Annual Meeting of the Association for Computational Linguistics, 2018:138 - 143.
- [ 15 ] Qiu Y, Li H Z, Li S, et al. Revisiting correlations between intrinsic and extrinsic evaluations of word embeddings[C]// Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data,2018:209 - 221.
- [ 16 ] Russo I, Caselli T, Rubino F, et al. EMOCause: An easy-adaptable approach to emotion cause contexts [ C ]//2nd Workshop on Computational Approaches to Subjectivity & Sentiment Analysis,2011:153 - 160.
- ~~~~~
- (上接第155页)
- [ 7 ] Yang L Q, Wang B Z, Li D Z, et al. Four slot dynamic binary search anti-collision algorithm[C]//International Conference on Computation, Communication and Engineering, 2020:138 - 141.
- [ 8 ] Su J, Sheng Z G, Hong D F, et al. An effective frame breaking policy for dynamic framed slotted ALOHA in RFID [J]. IEEE Communications Letters, 2016, 20(4):692 - 695.
- [ 9 ] 周艳聪,董永峰,张晶,等. 基于哈希分组的动态帧时隙 ALOHA 防碰撞算法[J]. 计算机工程与设计,2016,37(2):540 - 544,551.
- [ 10 ] 袁莉芬,杜余庆,何怡刚,等. 可并行识别的分组动态帧时隙 ALOHA 标签防碰撞算法[J]. 电子与信息学报,2018,40(4):944 - 950.
- [ 11 ] 史长琼,肖瑞强,吴丹. 改进的动态帧时隙 ALOHA 防碰撞算法[J]. 计算机工程与设计,2014,35(6):1897 - 1900, 1910.
- [ 12 ] 杨帆,任守纲,徐焕良,等. 基于比特估计的 RFID 标签数量估计算法[J]. 计算机工程与科学,2020,42(3):449 - 455.
- [ 13 ] Zheng Y, Wang X M, Yang D Y, et al. An efficient RFID tag cardinality estimation protocol based on bit detection [ C ]//17th International Conference on Communication Technology,2017:602 - 606.