

文章编号: 2095-2163(2020)04-0193-06

中图分类号: TP391

文献标志码: A

双向 GRU 和自注意力机制下微博情感倾向性分析

杨凡¹, 薛佳琦²

(1 西安建筑科技大学 理学院, 西安 710000 ;2 北京搜狗科技发展有限公司, 北京 100000)

摘要: 针对微博文本内容的即时性、随意性、碎片性, 本文提出将网络流行语词库、微博表情词库、网络流行词库加入到结巴分词词库中, 以确保 Word2vec 模型对文本准确词向量化; 针对上下文语义关联的微博文本特征以及双向长短期记忆网络(Bidirectional Long Short-Term Memory, BiLSTM) 难收敛、训练时间长的问题, 采用 BiGRU(Bidirectional, Gated Recurrent Unit) 神经网络学习微博文本的深层特征; 针对当下流行的注意力机制(AttentionMechanism), 调整情感词权重, 但同时忽略微博内部序列关系的问题。该文使用自注意力机制(Self-Attention) 综合微博文本内部关键特征对情感词进行定位, 调整情感词的概率权重; 然后用 softmax 函数进行分类。实验证明, 本文方法比现有方法的准确率提高了 5.34%。

关键词: 词库; 结巴分词; BiGRU; Self-Attention

Analysis of Weibo Emotional Tendency under BiGRU and Self-Attention Mechanism

YANG Fan¹, XUE Jiaqi²

(1 School of Xi'an University of Architecture and Technology, University Xian 71000, China;

2 Department of Beijing Sogou Technology Development Co., Ltd, Beijing 10000, China)

[Abstract] In view of the immediacy, randomness and fragmentation of the content of Weibo, this paper proposes to add the network popular words, the expression words and the network popular words into the Jieba segmentation lexicon, so as to ensure that the Word2vec model quantifies the text accurately. The characteristics of microblog texts for contextual semantic association and the shortcomings of Bidirectional Long Short-Term Memory (BiLSTM), which are difficult to converge and have long training time. BiGRU (Bidirectional, Gated Recurrent Unit) nerve is used in this paper. The deep features of the Weibo texts of online learning; the current attentional attention mechanism (AttentionMechanism) adjusts the weight of emotional words but ignores the internal sequence relationship of Weibo. This paper uses the Self-Attention mechanism to synthesize microblogs. The key features in the text locate the emotional words and adjust the probability weights of the emotional words. Finally, the softmax function is used to classify. Finally, the experimental method of this paper has a 5.34% improvement over the existing method.

[Key words] Thesaurus; Jieba segmentation lexicon; BiGRU; Self-Attention

0 引言

由于微博平台具有便捷性、传播性、原创性等特点, 已然成为大众交流信息、表达情感及观点的平台。目前, 对微博文本的分析主要有微博主题分析^[1]、舆情分析^[2]、情感分析^[3], 其中微博情感分析是微博文本分析的基础。当前微博文本情感分析采用支持向量机^[4]、卷积神经网络^[5]、长短时记忆神经网络^[6]等模型进行情感分析, 并取得了较好的结果。但在使用支持向量机模型进行情感分析时, 发现存在训练耗时、对大数据量样本难以实施、需谨慎选择核函数的缺点, 并不适用于大数据量的微博情感分析; 在采用卷积神经网络和长短时记忆神经网络模型进行情感分析时, 也存在训练时间长、参数多、内部结构复杂^[7], 以及不能准确定位情感词的

缺点。

近年来, 网络新词每天都在不断涌现, 微博作为大众网络交流信息的重要工具, 文本中也含有大量的网络热词, 如: “香菇蓝瘦”、“正能量”、“疯狂打call”、“雨女无瓜”、“大猪蹄子”等。在社交媒体中, 这些词对文本的情感表达、情感倾向性, 产生着重要的影响。而现有模型在处理微博文本分词时, 采用传统的分词方法, 极大地影响了微博情感分词的准确率。

为解决现有微博情感模型的不足, 以及网络新词影响情感极性的问题。本文提出在准确分词的基础上, 采用训练时间短、参数少、内部结构相对简单的 GRU 模型进行微博情感学习, 用 Self-Attention 模型进行情感词定位, 最后分类。通过在微博数据

作者简介: 杨凡(1995-), 女, 硕士研究生, 主要研究方向: 自然语言处理、深度学习、机器学习; 薛佳琦(1993-), 男, 硕士研究生, 主要研究方向: 自然语言处理、大数据。

收稿日期: 2020-02-26

集上的实验,与现有模型对比,证明本文所构建的模型更有利于微博情感倾向性分析。

1 相关工作

1.1 数据处理

本实验采用 python 编写程序代码,爬取微博数据集。该数据集随机获取了 2019 年几个时间段的数据集,总计 100 万条。实验所用微博内容包括以下几个方面:文字、表情、参与话题、链接、标点符号、特殊符号等。

(1)清洗数据。本文微博数据由 python 爬虫爬取,因此难免出现数据缺失、无效数据,对该类型数据进行排查和修改;因本文对微博文本情感倾向性进行研究,故对微博内容中特殊符号、标点符号、参与话题及微博话题,用正则表达式的方法进行去除。

(2)异常值与重复数据监测。只含有微博链接、标点符号、特殊字符的微博,经过数据清洗会出现异常值,将这些异常值去除;获取的数据有重复数据存在,将这些重复的微博进行去除。

(3)微博表情替换。微博表情是一种“携带意义感知”的符号,是一种视觉化语言,更加具体、生动和形象地表达微博情感,是人真实表情的延伸。获取的微博内容中含有大量的表情符,将这些表情符替换成对应的情感词,有利于后续研究的推进。微博表情库如图 1 所示。



图 1 微博表情库

Fig. 1 Weibo expression library

1.2 分词处理

结巴分词^[8]是国内程序员用 Python 开发的一种中文分词模块,该模块可以准确的识别出分词,将连续的字序列按照一定的规范重新组合成词序列。该模块支持 4 种分词模式:精确模式、全模式、搜索引擎模式、自定义字典模式。在分词时,针对含有大量网络新词的微博文本,如果仅仅依靠精确模式的结巴分词法,并不能正确分词。故使用自定义字典

模式将近年来微博表情词、搜狗网络新词、网络流行词、QQ 输入法网络流行语、搜狐网络新词加入结巴分词词库。分词结果见表 1。

表 1 分词结果

Tab. 1 Segmentation results

精确模式结巴分词	加入网络新词结巴分词
这是/正/能量/没错/啦	这是/正能量/没错/啦
休/了/3/天/之/后/的/第/一/个/工/作/日/有/些/许/不/适/应/上/班/路/上/还/交/通/拥/堵/严/重/蓝/瘦/香/菇	休/了/3/天/之/后/的/第/一/个/工/作/日/有/些/许/不/适/应/上/班/路/上/还/交/通/拥/堵/严/重/蓝/瘦/香/菇
看/小/欢/喜/看/得/我/好/压/抑/,/真/的/太/能/理/解/因/子/丁/一/了	看/小/欢/喜/看/得/我/好/压/抑/,/真/的/太/能/理/解/因/子/丁/一/了
香/港/废/亲/的/新/闻/了/简/直/辣/眼/睛/这/出/戏/啥/时/候/可/以/杀/青	香/港/废/亲/的/新/闻/了/简/直/辣/眼/睛/这/出/戏/啥/时/候/可/以/杀/青

分词后用哈工大停用词表,对微博文本集进行去停用词。

2 模型

2.1 word2vec 模型

Word2vec 是 Google 公司在 2013 开放的训练词向量软件工具,其可以根据给定的语料库训练合适的词模型。该词模型可以快速有效地将一个词语表达成向量形式,是自然语言处理领域的新工具。Word2vec 有 2 种模型:CBOW 模型和 Skip-gram 模型。CBOW 模型采用根据前后词预测中间词,如根据“这只____实在太萌了”预测“小猫”,适合较大的数据集。而 Skip-gram 模型则是根据中间词预测前后词,如:根据“小猫”预测“这只____实在太萌了”。Skip-gram 把每个词作为中心词对其它词进行预测,适合数据量适中的文本集。本文实验中有 100 万条微博数据,属于数据量适中的文本集,故采用 Skip-gram 模型训练词向量模型。

2.2 GRU 模型

针对中文微博文本上下文语义关联的特征,本文选用 LSTM(Long Short-Term Memory)的一个变种模型 GRU(Gated Recurrent Unit),用此模型进行微博情感学习。该模型由重置门(r_t)、更新门(z_t)、隐藏层(h_t)3 部分组成。虽然每条中文微博的篇幅在 240 字之内,属于短文本,但决定文本的情感极性与该文本的上下文都有关。而 GRU 模型可以通过重置门(r_t)、更新门(z_t)、隐藏层(h_t)3 部分,记忆和更新每个词之前的情感信息;同时该模型比 LSTM 参数少、内部结构简单、训练时间短,更适用于本文数据集。GRU 模型的小单元结构如图 2 所示。

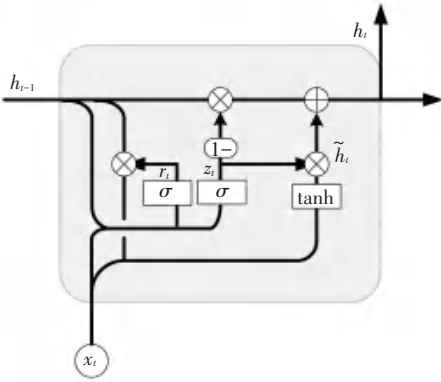


图 2 GRU 单元
Fig. 2 GRU unit

首先,GRU 的重置门 (\$r_t\$) 决定了如何将新的输入词信息与前面的情感记忆相结合。

$$r_t = \sigma(W^r x_t + U^r h_{t-1}), \quad (1)$$

更新门 (\$z_t\$) 决定当前时刻输入信息, 被附加到上一时刻状态中的量, 以及上一时刻的情感信息有多少保留在时间步 \$t\$ 中, 更新门的具体表达式为:

$$z_t = \sigma(W^z x_t + U^z h_{t-1}), \quad (2)$$

更新门将这 2 部分信息相加并作为 sigmoid 函数的输入。

最后, 计算隐藏层。隐藏层由 \$r_t\$ 控制保留之前的情感记忆 (\$\tilde{h}_t\$)。

$$\tilde{h}_t = \tanh(W^h x_t + r_t U^h h_{t-1}), \quad (3)$$

进而得到最后的隐藏层情感信息 (\$h_t\$)。

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

其中, \$W^r\$、\$W^z\$、\$W^h\$ 分别为重置门、更新门、隐藏层的权重矩阵, \$\tanh\$、\$\sigma\$ 为激活函数。

2.3 Self-Attention 机制

注意力机制 (Attention Mechanism) 源于对人类的视觉研究。在认知科学中, 由于信息处理的瓶颈, 人类会选择性地关注所有信息中的一部分, 同时忽略其它可见信息。这种机制通常被称为注意力机制。Bahdanau 等^[10]将注意力机制用于机器翻译模型中, 提出了 Soft Attention Model, 并将其应用到了机器翻译领域。2017 年 Google 机器翻译团队发现, 使用自注意力机制可以学习文本与训练集, 以及文本自身的关键特征表示^[11]。而后, Self-attention 机制因既可以学习到文本与总文本集的关系, 又可以学习到自身的关键特征, 成为近期在自然语言处理中研究的热点模型。微博文本的情感极性 with 微博文本集、微博文本自身都有极大关系, 故本文采用 Self

-Attention 模型对微博文本情感学习。

Self-Attention 的核心与 Attention 一样。设 \$X = (x_1, x_2, \dots, x_n)\$ 为 Self-Attention 的输入, 则 Self-Attention 可被描述为一个查询 \$q\$, 得到一系列 \$(k-v)\$ 对的映射, \$Q = X \times W^Q\$、\$K = X \times W^K\$、\$V = X \times W^V\$ \$Q = (q_1, q_2, \dots, q_i)\$、\$K = (k_1, k_2, \dots, k_j)\$、\$V = (v_1, v_2, \dots, v_j)\$。其中, \$W^Q\$、\$W^K\$、\$W^V\$ 是模型训练过程中学习到的合适参数。主要步骤如下:

(1) 计算 \$Q\$ 和 \$K\$ 的相似度。常用计算相似度的方法有计算余弦相似度、点积。

(2) 为防止 \$Q\$ 与 \$K\$ 点乘结果过大, 会除以一个尺度标度 \$\sqrt{d_k}\$, 其中 \$d_k\$ 为 \$Q\$ 和 \$K\$ 的行数。

(3) 使用 Softmax 函数, 进行归一化为概率分布函数。

$$\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right),$$

然后, 再乘以矩阵 \$V\$, 得到权重求和的表示如下:

$$\text{Self-Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V.$$

3 基于 Self-attention 的 BiGRU 模型

通过清洗数据、训练 Word2vec 词模型后, 在 Word2vec 模型中寻找微博的词向量, 作为该模型的输入层, 并将这些词向量输入到隐藏层中。用 BiGRU 学习微博的上下文关系, 将学习到的上下文关系信息输入到 Self-attention 模型中, 用权重确定该微博的情感词。最后用 Softmax 函数进行分类作为该模型的输出层。其中隐藏层包括: BiGRU 层、Self-Attention 层、Dense 层。模型分层如图 3 所示。

3.1 输入层

为了对微博情感进行准确分析, 本文使用结巴分词的自定义字典模式, 将微博表情词、搜狗网络流行、加入结巴分词的词库中, 对微博文本进行准确分词, 然后训练出 Word2vec 模型。

在训练词向量时, 分别采用精确模式结巴分词和加入网络新词的结巴分词 2 种模式, 训练得到 Model1 和 Model2 这 2 个词向量模型。如: 表 2 中“这是正能量没错啦”的词向量。精确模式下, 结巴分词并去停用词后, 训练词向量模型和加入网络新词的结巴分词并去停用词后, 训练的词向量模型在微博句向量的表示中相差甚多。加入网络新词并去停用词后训练的词向量模型比精确模式下训练的模型更准确。可以得到即时性、含有网络新词的词

模型。

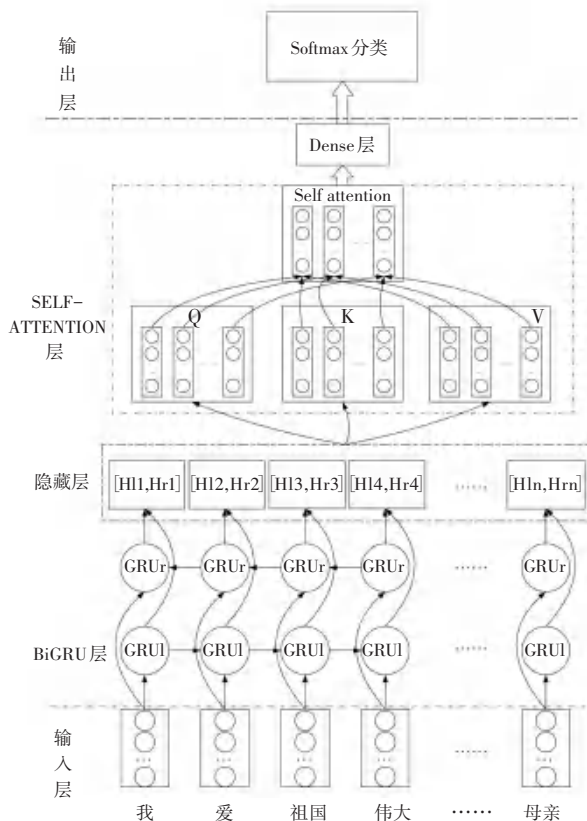


图 3 模型分层图

Fig. 3 Model layering diagram

表 2 词向量

Tab. 2 Word vectors

Model1	这是	正	能量	没错
1	-0.121 91	0.109 48	-0.568 44	-0.158 64
2	-0.122 57	0.169 86	-0.032 30	-0.148
3	0.041 772	0.436 84	0.081 071	-0.151 96
4	-0.082 69	-0.174 03	0.093 198	-0.078 54
5	-0.249 89	0.062 44	-0.595 49	-0.103 23
6	-0.381 49	0.031 176	-0.693 11	-0.236 79
7	-0.171 86	0.191 39	0.163 152	0.163 847
8	-0.187 18	-0.104 13	0.163 83	0.236 18
...
100	-0.447 4	-0.208 9	0.175 987	-0.202 94

Model2	这是	正能量	没错
1	-0.121 91	-1.048 2	-0.158 64
2	-0.122 57	0.029 940	-0.148
3	0.041 772	0.303 661	-0.151 96
4	-0.082 69	-0.171 57	-0.078 54
5	-0.249 89	-0.537 31	-0.103 23
6	-0.381 49	-0.099 10	-0.236 79
7	-0.171 86	-0.131 21	0.163 847
8	-0.187 18	0.143 86	0.236 18
...
100	-0.447 4	0.159 969	-0.202 94

3.2 隐藏层

3.2.1 BiGRU 模型

在单向的神经网络 GRU 模型中,是从往后输出状态,这样 GRU 模型记忆单元中只记得该词之前的微博的特征。如果仅仅从前往后输入就忽略了下文。基于此问题,本文采用 BiGRU 神经网络模型, BiGRU 含前向 GRU 模型、后向 GRU 模型。前向 GRU 层学习该条微博的上文信息 (Hl), 后向 GRU 层学习该条微博的下文信息 (Hr)。将每条微博中每个词的上文信息 (Hl) 和下文信息 (Hr) 拼接,得到每个词的上下文信息 [Hr, Hl]。

3.2.2 Self-Attention 机制层

为了定位情感词,引入 Self-Attention 模型。该机制可以计算每条微博中每个词的情感概率权重和自身情感关系。在训练模型过程中,将 BiGRU 层学习到微博上下文信息输入到 Self-Attention 模型中,学习得到适合微博情感文本集的 W^Q, W^K, W^V 3 个权重矩阵。将测试集中 BiGRU 模型学习到的微博深层特征 ($X = [Hr, Hl]$), 输入 Self-Attention 机制。即可得到 $Q = X \times W^Q, K = X \times W^K, V = X \times W^V$ 。那么微博句子中每个词的权重值为:

$$Self - Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V.$$

其中, d_k 为 Q 和 K 的行数。本文对微博数据进行情感分类,通过 Self-Attention 机制学习到概率权重中,情感词的概率权重较大,地点词、名词、量词等与情感无关的词概率权重较小。

3.2.3 Dense 层

Dense 层^[12],也称全连接层。该层是将每一个单元和 Self-Attention 层中每一个单元相连接,得到每个词属于哪个分类的概率权重。

3.3 输出层

最后用 Softmax 函数对每条微博的概率权重进行归一分类得到情感类别。

4 实验

4.1 实验设置

为了证明该模型的有效性,使用获取的 100 万微博情感集训练 Word2vec 词向量,将数据中 10 000 条进行手工标记。其中,正向情感 5 320 条,负向情感 4 680 条。训练集和测试集按照 8:2 进行划分。

本实验采用 keras 深度学习框架^[13],后端环境为 TensorFlow^[14],使用 Python 语言编程实现;实验运行环境为 Anaconda3 软件、Win7 系统、内存 8 GB。实验参数设置见表 3。

表3 实验参数设置

Tab. 3 Experimental parameter settings

参数	值
BiGRU 隐藏节点数	100
Loss 函数	categorical_crossentropy
Optimizer	Adam
Batch_size	128
词向量维度	100

4.2 模型性能

对微博数据集分类时会出现以下4种情况:

(1)真正类(True Positive, TP):被判定为正向情感,事实上也是正向情感;

(2)假正类(False Positive, FP):被判定为正向情感,但事实上是负向情感;

(3)假负类(False Negative, FN):被判定为负向情感,但事实上是正向情感;

(4)真负类(True Negative, TN):被判定为负向情感,事实上也是负向情感。

本实验采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、 F 值^[15]以及运行时间对试验结果进行评价。准确率是指分类正确的样本数占总样本的比例;精确率是指该模型预测的正向情感微博数据集中正向情感的比例;召回率体现该模型对正向情感微博数据集的识别能力; F 值为精确率和召回率的调和值,更接近于两个数较小的那个,所以精确率和召回率越接近 F 值越大,说明该模型越稳健。

$$Accuracy = \frac{TP + TN}{(TP + FN + FP + TN)}, \quad (5)$$

$$Precision = \frac{TP}{(TP + FP)}, \quad (6)$$

$$Recall = \frac{TP}{(TP + FN)}, \quad (7)$$

$$F = \frac{2 * Recall * Precision}{Recall + Precision}. \quad (8)$$

4.3 对比试验设置

为了证明该模型对微博文本分类的有效性,本文实验与当下较为流行的微博文本分类模型进行了比对。

实验一 Word2vec+BiLSTM 微博情感分析

通过 Skip-gram 模式对结巴分词后的微博文本进行训练,得到词向量模型,将词向量输入 BiLSTM 神经网络模型中,捕捉微博情感后进行分类。

实验二 Word2vec+BiGRU 微博情感分析

在实验一的实现过程中,可以发现 LSTM 神经网络的训练参数较多、内部结构复杂、训练时间过长

且收敛速度较慢。而 GRU 神经网络模型是 LSTM 的简化模型,具有训练参数少、内部结构简单、训练时间短且收敛速度快的特点。因此,实验二中将 BiLSTM 模型替换为 BiGRU 模型。

实验三 网络新词+Word2vec+BiGRU

在处理微博数据集时发现,用传统的分词模式对其分词,会导致大量微博数据集失去情感特征,甚至会歪曲情感倾向。故本文将网络新词、微博表情词加入结巴新词库中,利用结巴分词对微博文本集进行准确分词,将词向量输入到 BiGRU 神经网络中捕捉微博深层特征,并对其进行分类。

实验四 网络新词+Word2vec+双向GRU+Attention

为能准确定位情感词,以达到更好的分类效果,实验三在采用 BiGRU 学习微博深层特征的基础上,用 Attention 机制来确定微博中每个词的权重。情感词的权重较大。

实验五 网络新词+word2vec+BiGRU+Self-Attention

Self-Attention 机制比 Attention 机制的性能更好,更注重微博中词与词之间的位置关系、结构关系、情感关系。本文使用基于网络新词、网络流行语的 Word2vec 模型,结合 BiGRU 和 Self-Attention 的神经网络模型,对微博情感倾向性进行分析。

4.4 实验结果比较

在微博文本集上运用以上模型,结果对比见表4。

表4 结果对比

Tab. 4 Comparison of results

对比模型	准确率/%	召回率/%	精确率/%	F 值/%	时间/s
实验一	85.94	87.10	86.37	86.73	750
实验二	86.30	87.61	86.46	87.03	550
实验三	89.92	90.96	89.94	90.44	537
实验四	90.50	91.77	90.22	90.98	834
实验五	91.64	92.50	91.63	92.06	982

由实验一与实验二对比发现:BiLSTM 和 BiGRU 在性能上不分伯仲,BiGRU 比 BiLSTM 在本文的微博情感数据集上分类性能略好。BiGRU 比 BiLSTM 参数少、容易收敛,分类时间短 200 s。

对比实验二和实验三发现:由于微博的时效性、随意性、碎片性,加入网络新词、网络流行语及微博表情词,对微博数据集进行训练的 Word2vec 模型比传统分词后训练的 Word2vec 模型,对微博的词向量化更加准确。实验三比实验二模型的准确率、召回率、精确率、 F 值分别有 3.62%、3.35%、3.48%、3.41%的提升。

对比实验三和实验四发现:加入 Attention 层对微博情感词定位后分类,可以提高模型性能。因为,

Attention 模型通过模型训练可以计算每个词的概率权重,减少处理高维数据的计算复杂度,从而降低了数据的维度,使得模型更快、更容易学习到输入微博数据与输出信息的关联信息,进而达到定位情感词的效果。

对比实验四和实验五发现:加入 Self-Attention 机制比 Attention 机制模型准确率性能有 1.14% 提升。Self-Attention 机制是 Attention 机制的一种改进。从参数角度上,Self-Attention 机制对外部参数依赖较少;从性能角度来看,Self-Attention 可以学习每条微博不同位置的词之间特征,以捕捉长距离的依赖关系。结合加入微博表情词库、网络新词库、网络流行词库训练的 Word2vec 模型,对微博准确的词向量化,比实验一、实验二、实验三、实验四有非常显著地提高,故本模型可以更加准确地关联情感词,更适合研究微博情感倾向性。

5 结束语

针对微博文本的特征,该文提出将网络流行词、微博表情词、网络流行语加入到结巴分词的词库中,以至于 Word2vec 模型对微博文本进行准确的词向量化。进而采用 BiGRU 和 Self-Attention 机制对微博文本进行情感特征学习、微博内部关键特征学习和情感词定位,最终使用 Softmax 函数分类。通过在同一数据集上的对比实验,以及模型性能分析,发现本模型更适合微博倾向性分析。

参考文献

- [1] 熊回香,叶佳鑫. 基于 LDA 主题模型的微博标签生成研究[J]. 情报科学,2018,36(10):7-12.
- [2] 连芷萱,兰月新,夏一雪,等. 基于首发信息的微博舆情热度预

测模型[J]. 情报科学,2018,36(9):107-114.

- [3] XIANG Z Q, ZOU Y X, WANG X. Sentiment analysis of Chinese micro-blog using vector space model [C]// Signal and Information Processing Association Annual Summit and Conference (APSIPA). Siem Reap: IEEE, 2014: 1-5.
- [4] TONG S, KOLLER D. Support vector machine active learning with applications to text classification [J]. Journal of Machine Learning Research, 2002, 2(1): 45-66.
- [5] 张海涛,王丹,徐海玲等. 基于卷积神经网络的微博舆情情感分类研究[J]. 情报学报,2018,37(7):695-702.
- [6] 陈思. 微博文本情感分类研究[D]. 长春:吉林大学,2016.
- [7] 张玉环,钱江. 基于两种 LSTM 结构的文本情感分析[J]. 软件,2018, 39(1): 116-120.
- [8] Sun J. Jieba chinese word segmentation tool [J]. 2018-01-21 [2018-06-25]. <https://github.com/fxsjy/jieba>, 2012.
- [9] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [J]. arXiv preprint arXiv: 1406.1078, 2014.
- [10] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [J]. arXiv preprint arXiv:1409.0473, 2014.
- [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in neural information processing systems 31:5998-6008.
- [12] 黄婕,张丰,杜震洪,等. 基于 RNN-CNN 集成深度学习模型的 PM_{2.5} 小时浓度预测[J]. 浙江大学学报(理学版), 2019, 46(3): 370-379.
- [13] GULLI A, PAL S. Deep Learning with Keras [M]. Packt Publishing Ltd, 2017.
- [14] ABADI M, BARHAM P, CHEN J, et al. Tensorflow: A system for large-scale machine learning [C]//12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). 2016: 265-283.
- [15] KIM Y, DENTON C, HOANG L, et al. Structured attention networks [J]. arXiv preprint arXiv:1702.00887, 2017.

(上接第 192 页)

组建医联体。在民营发展初期给予更多扶持,消除区域卫生规划上存在的障碍,在医保定点、职称评定、科研立项、资金支撑、税收减免等方面加大对民营医院的政策支持,通过市场调节力量提高医联体运作成效,以此倒逼公立医院从实际出发真正推进改革。

参考文献

- [1] 郑芸,农圣. 部分发达国家分级诊疗设计的政策启示[J]. 卫生经济研究,2017(12):27-30.
- [2] 李亚男,雷涵,吴海波. 国外分级诊疗及其对我国的启示[J]. 国外医学·卫生经济分册,2017,34(2):49-53.
- [3] 孙晓凡,陈曼洁,闻大翔,等. 英、美、荷、澳、日分级诊疗实践的启示[J]. 中国卫生质量管理,2016,23(5):105-108.
- [4] 张兴祥,庄雅娟. 西方发达国家分级诊疗体系比较及经验启示[J]. 经济资料译丛,2018(3):14-23.
- [5] 郑蕾. 英国分级诊疗对我国的启示[J]. 中国卫生质量管理,2017,24(3):103-106

- [6] 王德永,王莉. 分级诊疗制度国际比较[J]. 劳动保障世界,2018(29):30-32.
- [7] 顾亚明. 日本分级诊疗制度及其对我国的启示[J]. 卫生经济研究,2015,(3):8-11,12.
- [8] 邢梦琪. 环顾他国分级诊疗制度[J]. 劳动保障世界,2018,(22):71.
- [9] 邢梦琪. 环顾他国分级诊疗制度[J]. 劳动保障世界,2018(22):71.
- [10] 张皓. 基于系统动力学模型的分级诊疗体系及政策仿真研究[D]. 杭州:浙江大学,2017.
- [11] 郝磊俊. 居民的分级诊疗认知及社区就医影响因素研究[D]. 杭州:浙江大学,2018.
- [12] 黄丞,张文信. 上海市分级诊疗成效、难点和问题分析[J]. 中国医院,2018,22(11):14-18.
- [13] 蒲宽. 大连市分级诊疗制度实施现状调查研究[D]. 大连:大连理工大学,2018.
- [14] 沈晓初. 上海市构建分级诊疗制度的改革与探索[J]. 中国卫生资源,2016,19(01):1-3
- [15] 殷一宁. 上海市社区卫生服务中心在社区首诊制度运行中的角色研究[D]. 上海:上海师范大学,2015.