

# 一种新的考虑注意力机制的微博谣言检测模型

潘德宇<sup>1</sup> 宋玉蓉<sup>1</sup> 宋波<sup>2</sup>

<sup>1</sup>(南京邮电大学 自动化学院、人工智能学院, 南京 210023)

<sup>2</sup>(南京邮电大学 现代邮政学院, 南京 210003)

E-mail: 291791609@qq.com

**摘要:** 随着社交媒体的迅速发展, 谣言通过社交媒体迅速传播, 识别社交媒体网络上的谣言是社交网络研究中一个至关重要的问题。本文提出了一种新的考虑注意力机制的微博谣言检测模型, 考虑到卷积神经网络(CNN)提取到的特征对输出结果影响力问题, 在经典的文本卷积神经网络(Text CNN)上引入了注意力机制, 通过CNN中的卷积层学习微博窗口的特征表示, 再根据每个特征表示对输出结果的影响力不同通过注意力机制赋予不同的权重来进行谣言事件的检测。研究结果表明, 本文提出的微博谣言检测模型准确率达到了96.8%, 并且在召回率和F1值上也有提升, 即本文提出的新的微博谣言检测模型具有更好的谣言识别能力。

**关键词:** 谣言检测; 卷积神经网络; 注意力机制; 准确率

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2021)02-0348-06

## New Microblog Rumor Detection Model Based on Attention Mechanism

PAN De-yu<sup>1</sup> SONG Yu-rong<sup>1</sup> SONG Bo<sup>2</sup>

<sup>1</sup>(School of Automation and Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

<sup>2</sup>(School of Modern Post, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** With the rapid development of social media, rumors spread rapidly through social media, and identifying rumors on social media networks is a crucial issue in social network research. This paper proposes a new microblog rumor detection model that considers the attention mechanism. Considering the influence of the features extracted by the convolutional neural network (CNN) on the output result, it is introduced in the classic text CNN. In order to understand the attention mechanism, the feature representation of the microblog window is learned through the convolutional layer in CNN, and then according to the different influence of each feature representation on the output result, different weights are assigned to the attention mechanism to detect rumor events. The research results show that the accuracy rate of the microblog rumor detection model proposed in this paper reaches 96.8%, and the recall rate and F1 value are also improved, that is, the new microblog rumor detection model proposed in this paper has better rumor recognition ability.

**Key words:** rumor detection; convolutional neural network; attention mechanism; accuracy

## 1 引言

谣言一般是指与某一事件相关未经核实的陈述或说明<sup>[1]</sup>。随着社交媒体的迅速发展, 谣言可以通过社交媒体迅速传播, 微博等社交平台已经逐渐成为滋养虚假消息和散布谣言的理想场所, 一旦虚假谣言被广泛传播, 就会引发诸多网络安全问题, 产生社会动荡。因此, 谣言检测<sup>[2]</sup>至关重要。微博作为时下最火热的社交平台之一拥有数以亿计的用户基数, 在居民生活中占据不可或缺的地位, 本文选用微博社交网络作为典型进行研究。

谣言检测问题实际上是一个二分类的问题, 目前关于社交媒体谣言检测的研究方法主要是基于深度学习方法。通过构造神经网络并搭配非线性函数学习文本中的潜在特征, 通过卷积神经网络(CNN)<sup>[3]</sup>、循环神经网络(RNN)<sup>[4]</sup>等神经

网络模型对文本序列进行特征表示学习, 解决句子分类<sup>[5]</sup>、情感分类<sup>[6]</sup>等问题。这些构造神经网络的方法在谣言检测中取得了显著效果。

Ma等人<sup>[7]</sup>在2015年将微博事件相关内容根据时间先后建模为变长时间序列, 提出基于循环神经网络模型学习文本抽象表示; 刘政等人<sup>[8]</sup>在2017年提出卷积神经网络的谣言检测模型, 将微博事件向量化, 通过隐含层的学习训练来挖掘文本深层特征, 但是这种方法提取出的文本特征并不全面; 李力钊等人<sup>[9]</sup>在2018年提出的结合卷积神经网络和门控循环单元(GRU)的谣言检测模型, 从微博事件中学习到窗口特征序列, 再通过GRU<sup>[10]</sup>学习序列特征来进行谣言检测, 解决了卷积神经网络提取特征不全面的问题, 但提取出的特征对输出结果影响力都一样, 无法区分明显特征。

Wang等人<sup>[11]</sup>在2019年运用了双头注意力机制来进行

收稿日期: 2020-02-28 收修改稿日期: 2020-04-23 基金项目: 国家自然科学基金项目(61672298)资助; 江苏高校哲学社会科学研究重点项目(2018SJDH42)资助; 教育部人文社会科学研究规划基金项目(17YJAZH071)资助。 作者简介: 潘德宇, 男, 1996年生, 硕士, 研究方向为复杂网络、谣言检测; 宋玉蓉, 女, 1971年生, 博士, 教授, 博士生导师, CCF会员, 研究方向为复杂网络中的信息传播及其控制策略, 自适应网络的建模、仿真和智能优化; 宋波, 男, 1987年生, 博士, 讲师, 研究方向为复杂网络传播动力学、自适应网络建模与分析、信息安全。

谣言检测,提出的模型可以自动捕捉上下文信息之间的依赖关系,并从上下文信息中学习全局表示;Fang 等人<sup>[12]</sup>在2019年建立了一个基于自我多头注意力的卷积神经网络模型,通过在预处理的词向量部分融合多头注意力机制挖掘词语之间的深层特征再通过卷积神经网络训练来进行谣言检测;Wu 等人<sup>[13]</sup>在2019年提出一种筛选的多任务学习方法,具有选定的共享层用于谣言检测,共享层选用门机制与注意力机制对任务之间的共享特征进行过滤选择。上述文献[11-13]都将注意力机制运用到了谣言检测研究中,解决了提取出的特征对输出结果影响力的问题。另外,考虑到了提取特征对输出结果的影响力问题的同时,本文还考虑到了提取特征不够全面的问题在预训练部分采用 BERT 模型<sup>[14]</sup>,进一步增强了谣言检测特征提取的全面性。

目前基于深度学习构造神经网络的微博谣言检测研究中,预训练模型大多采用的是 word2vec<sup>[15]</sup> 词向量或深度语境化词表征(ELMo)<sup>[16]</sup>。前者中得出的词向量无法解决多义词的问题使得训练出的每个词只能对应一个向量表示,后者可以根据上下文动态调整词嵌入,但是使用长短期记忆网络(LSTM)<sup>[17]</sup>进行特征抽取而不是 Transformer<sup>[18]</sup>,并且 ELMo 使用上下文向量拼接作为当前向量,这样融合出的向量特征较差。

本文针对目前的存在挑战提出了一种引入注意力机制的 Text CNN 模型用于微博谣言检测,通过给不同的特征根据其影响力赋予不同的权重,这样对输出结果影响较大的特征就会被赋予更多的权重,从而对结果产生更重要的影响,提高微博谣言检测准确性。本文方法通过实验结果表明,与目前最好的基线方法相比,该方法在新浪微博数据集上的准确率提高了约 1.1%,同时在召回率与 F1 值上也有好的表现。

## 2 模型方法

### 2.1 预训练模型

本文采用了 BERT 预训练模型,该模型相较于传统的 word2vec 和 ELMo 预训练模型在效果上有显著的提升。传统的预训练模型,主要限制是标准语言模型是单向的,使得模型预训练时可以使用的架构类型受到限制,而在 BERT 中提出了一种新的预训练目标:遮蔽语言模型(masked language model, MLM)。该模型克服传统的单向性局限,MLM 目标允许表征融合左右两侧的语境,从而可以预训练一个深度双向的 Transformer。除了遮蔽语言模型外, BERT 还引入了“下一句预测”任务,可以和 MLM 共同训练文本对的表示。BERT 模型运用了句子级负采样,对于句子级的连续性预测,就是预测输入 BERT 的两端文本是否连续。在训练过程中,输入模型的第 2 段将从所有文本中随机选择,概率为 50%,其余 50% 将选择第 1 段的后续文本。

BERT 预训练模型的结构如图 1 所示,它的每一层都对 Transformer 进行了双向的连接,通过多层的堆叠来构建出预训练模型。它的嵌入层由 Token 嵌入、分段嵌入和位置嵌入 3 部分组成。Token 嵌入就是最为常见的词向量形式,同时会在每个句子的首尾处位置分别嵌入 [CLS] 和 [SEP] 两个特殊的 Token,通过 Token 嵌入可以将预处理后的句子转化为固定维度的向量形式。分段嵌入用来判断属于文本是否存在句义相似。位置嵌入主要是用来解决编码无序的问题,该层会学习一

个位置向量来表征序列的顺序。BERT 预训练模型可以通过特殊的双向 Transformer 学习到上下文的词对当前词的影响

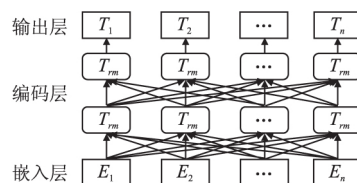


图 1 BERT 模型结构图

Fig. 1 BERT model structure diagram

程度,从而可以增强对语义深层特征的提取。本文通过 BERT 预训练模型可以得到微博与其对应评论的每一条句向量,句向量的维度固定在 768 维,这要比传统的只能转化出 100-200 维的 word2vec 方法来说可以提取出更多的语义特征。通过选取微博和对应的数条评论看作是一个事件,最后只需要判断该事件是否是谣言便可以完成谣言检测任务。

### 2.2 引入注意力机制的训练模型

#### 2.2.1 注意力机制

注意力模型<sup>[19]</sup>起初用于解码编码框架之中,编码就是将输入序列转化成一个固定长度的向量,而解码就是将之前生成的固定长度的向量再转化成输出序列,这种模型通常被应用于翻译、文本摘要的任务中。

但是解码编码模型具有很大的局限性,最明显的是作为编码与解码联系的语义向量长度已经固定,也就是说编码要将信息压缩到固定长度,这样便会出现两个缺点:1) 压缩出的向量无法表示整个序列的信息;2) 先输入的内容携带的信息会被后输入的信息稀释掉,整个序列越长,被稀释的信息就会越多。所以引入了注意力模型,通过注意力分配的不同,可以关注到每一个输入对于预测结果的影响。在编码的过程中,每一次预测都计算其对输入的影响,从而在解码时,输入序列的每一句都可以对预测的结果产生影响。

#### 2.2.2 训练模型

模型的整体流程如图 2 所示,首先对运用的数据集进行预处理,整理出微博新闻与对应评论组成事件,通过 BERT 预训练模型生成固定维度的事件句向量;接着通过 Text CNN 模型对生成的句向量进行卷积学习操作,得到基于过滤器的

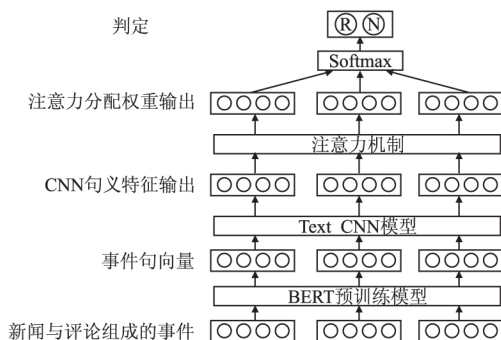


图 2 整体流程图

Fig. 2 Overall flow diagram

特征表示,将特征进行拼接得到完整的句义特征输出;再引入注意力机制,根据特征对输出结果的影响力不同分别赋予特

征不同的权重得到微博特征权重输出;最后通过 Softmax 分类器进行微博谣言事件判定.

详细的训练过程如下:

1) 输入预训练生成句向量. 对于数据集中所有的谣言与非谣言事件及其相应的评论,通过 BERT 预处理模型都训练成了句向量. 对于每一个微博事件,选取其事件下的相应数条评论和原微博一起作为输入. 输入层为一个  $m \times n$  的矩阵,  $m$  为选取的事件的总数量,  $n$  则为单条句向量的长度.

2) 使用过滤器卷积得到相应特征. 通过使用尺寸不同的三种过滤器进行卷积,分别得到对应不同过滤器的特征. 过滤器会在  $m \times n$  的输入矩阵中不停的滑动,为了方便提取特征,设定过滤器的长度为  $k$ ,宽度与输入矩阵宽度一样为  $n$ ,一个过滤器提取出的特征就可以表示为  $h \in \mathbb{R}^{k \times n}$ ,那么对应矩阵  $w$  中的任意一条  $u$  所获得的特征  $x_u$  为:

$$w_u = (x_u, x_{u+1}, \dots, x_{u-k+1}) \quad (1)$$

在对输入矩阵卷积完之后就会生成一个特征列表  $c$ ,每一次卷积生成的特征都会对应  $c$ :

$$c_u = f(w_u * h + b) \quad (2)$$

式中的  $f$  为  $ReLU$  函数,  $b$  为偏置项.

3) 特征拼接形成完整特征矩阵. 当过滤器在长度为  $m$  的输入上滑过时,特征列表的长度为  $(m-k+1)$ . 假设存在  $q$  个过滤器,则会产生  $q$  个特征列表. 将  $q$  通过拼接得到矩阵  $W_1$

$$W_1 = [c_1, c_2, \dots, c_q] \quad (3)$$

$c_q$  代表第  $q$  个过滤器产生的特征列表. 而本文一共运用了三种不同尺寸的过滤器,最后产生的总的矩阵  $W$  即为

$$W = [W_1, W_2, W_3] \\ = [c_1, c_2, \dots, c_q, c_{q+1}, \dots, c_{2q}, c_{2q+1}, \dots, c_{3q}] \quad (4)$$

4) 池化,全连接. 对每种过滤器获得的特征采取最大池化操作得到输出特征. 将不同过滤器输出特征进行全连接得到 CNN 输出矩阵  $W'$

$$W' = [c_{11}, c_{22}, \dots, c_{kk}] \quad (5)$$

上述步骤 1)~步骤 4) 介绍了完整的 CNN 卷积操作,操作过程如图 3 所示.

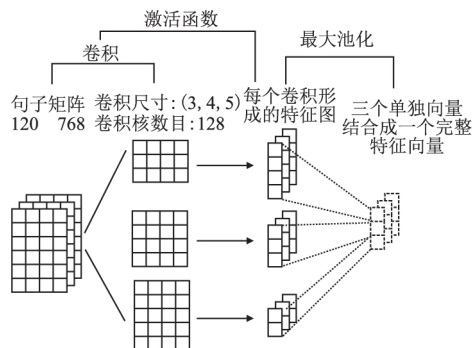


图3 CNN 卷积操作过程

Fig. 3 CNN convolution operation process

5) 引入注意力机制进行加权操作. 采用注意力层对 CNN 层的输出进行加权求和,以获取微博序列的隐层表示. 对 CNN 网络引入注意力机制能给 CNN 网络输出的隐状态序列  $W$  赋予不同的权重. 这样在学习微博序列的表示时模型能够有侧重的利用微博序列信息. 该注意力层将 CNN 网络的输出

$c_{kk}$  作为输入,输出微博序列对应的表示  $v_{kk}$

$$h_i = \tanh(W_A * c_{kk} + b_A) \quad (6)$$

$$\alpha_i = \frac{\exp(h_i^T h_A)}{\sum_{j=1}^n h_j^T h_A} \quad (7)$$

$$v_i = \sum_{kk} \alpha_i c_{kk} \quad (8)$$

组成矩阵  $V = [v_{11}, v_{22}, \dots, v_{kk}]$ ,  $W_A$  为权重矩阵,  $b_A$  为偏置值,  $h_i$  为  $c_{kk}$  的隐层表示,  $\alpha_i$  为  $h_i$  与上下文  $h_A$  的相似度,  $v_i$  为输出向量.

6) 分类器进行微博谣言判别. 将输出向量送入全连接层,通过 Softmax 分类器对微博事件进行谣言与非谣言的二分类,得到相关概率输出  $P_c$ :

$$P_c = \text{Softmax}(W_s * v_i + b_s) \quad (9)$$

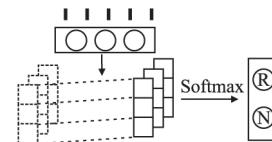


图4 谣言判别示意图

Fig. 4 Rumor discrimination diagram

$W_s$  为权重矩阵,  $b_s$  为分类层偏置向量. 步骤 5)~步骤 6) 介绍了引入注意力机制并通过分类器进行判别操作,详细过程如图 4 所示.

### 3 实验分析

#### 3.1 数据集

本文选用的数据集一个是来自 Ma 等人<sup>[7]</sup>整理并用于论文里的一系列的基于微博平台的事件数据,该微博数据集是由 Ma 通过微博 API 捕获的原始信息以及给定事件的所有转发和回复信息,包含 2746818 名用户,3805656 篇帖子. 另一个是 CED\_Data set<sup>[20]</sup>,包含微博原文与其相应转发/评论内容. 数据集详细统计情况如表 1 所列.

表1 数据集详情统计结果

Table 1 Data set statistics

统计	Ma	CED_Data set
事件数量	4664	3387
谣言事件数量	2313	1538
非谣言事件数量	2351	1849

研究将所有的数据按照 4:1 的比例并进行五折交叉验证,分别用作模型的训练集与测试集,划分情况如表 2 所列.

表2 数据集划分

Table 2 Data set partition

名称	数据集	谣言	非谣言	总数
Ma	训练集	1813	1851	3664
	测试集	500	500	1000
CED_Data	训练集	1188	1499	2687
	测试集	350	350	700

句向量维度设置为 768,参数优化器采用 Adam Optimizer<sup>[21]</sup>更新,dropout 设置为 0.8,输入矩阵长度设置 120,学习率设置 0.001.

### 3.2 实验评价指标

本文采用的用来评估模型有效性的评价指标与之前文献中所运用的指标一致,分别是精确率、准确率、召回率和 F1 值 4 个。对于二分类问题,预测结果和实际结果两两结合会出现如表 3 所列 4 种情况:

表 3 二分类问题的结果表示

Table 3 Result representation of binary classification problem

预测结果	实际结果		
		1	0
	1	TP	FP
	0	FN	TN

1) 准确率(Accuracy): 预测正确的结果占总样本的百分比。

2) 精确率(precision): 在所有预测为正的样本中实际为正的比率。

3) 召回率(recall): 针对原样本而言,在实际是正的样本中被预测为正样本的概率。

4) F1 值(F1-score): 同时考虑精确率与召回率,让两者同时达到最高,再取得平均。

计算方式分如式(10)~(13)所示:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (10)$$

$$precision = \frac{TP}{TP + FP} \quad (11)$$

$$recall = \frac{TP}{TP + FN} \quad (12)$$

$$F_1 = \frac{2 \times precision \times recall}{precision + recall} \quad (13)$$

### 3.3 实验对比模型

将本文方法与其他基线方法在同样的微博数据集上进行实验,本文选取了如下几个基线方法:

SVM-TS 模型<sup>[22]</sup>: 该模型通过时间先后序列来捕捉微博事件随时间先后的特征变化,使用 SVM 分类器进行分类。

GRU-1 模型<sup>[7]</sup>: 该模型将传统循环神经网络中的循环单元改为门控单元,不仅可以捕捉到当前时间步长中的特征,还能融入前一时刻的特征,同时添加了一个嵌入层有效降低了模型复杂度,最后通过 Softmax 进行分类。

GRU-2 模型<sup>[7]</sup>: 在 GRU-1 的模型上又添加了一层 GRU 层,这一层可以捕获不同时间步长之间更高级别的特征交互,最后也通过 Softmax 进行分类。

CNN-GRU 模型<sup>[9]</sup>: 该模型在传统的 CNN 上又添加了 GRU,既解决了 CNN 模型无法在时间上学习微博时间先后序列相关特征的问题,也关注到了 RNN 模型虽然可以学习事件序列表示,但是无法学习到微博事件的语句的句义的难题。

CNN-attention 模型: 该模型是本文所提出的模型,在传统 CNN 模型上融入了注意力机制,在对每一个过滤器经过卷积池化操作后所形成的特征,再经过全连接层形成完整的特征序列,在这之后加入注意力机制可以根据特征对输出产生的不同影响程度赋予不同的权重特征。

### 3.4 实验结果分析

本文分别在两个不同微博数据集上进行实验对比。通过使用 Ma 等人的数据集,本文方法与基线方法在谣言检测中

效果比较实验结果如表 4 所列,不同模型准确率的对比图如图 5 所示,从中可以发现传统的 SVM-TS 方法使用分类器进行谣言检测最后的准确率仅仅只有 85.7%,效果并不是特别优秀,对比 GRU-1、GRU-2、CNN-GRU 三种模型的最后结果可以发现训练模型中加入卷积神经网络后因为可以通过过滤器提取到输入中不同的潜在特征,所以在准确率上有更好的表现最高达到了 95.7%,而我们的模型在引入注意力机制后,考虑 CNN 输出特征表示对最终输出结果影响力问题,对不同输出特征赋予不同权重这样对输出结果影响较大的特征就会被赋予更多的权重,从而对结果产生更重要的影响,有助于进行谣言检测,结果表明我们的模型准确率达到到了 96.8%,并且在召回率和 F1 值上也有不错的提升。

表 4 不同检测方法下的实验结果对比(Ma 数据)

Table 4 Comparison of experimental results under different detection methods(Ma data)

模型	类别	准确率	精确率	召回率	F1
SVM-TS	R	0.857	0.839	0.885	0.861
	N		0.878	0.830	0.857
GRU-1	R	0.896	0.871	0.958	0.913
	N		0.952	0.858	0.903
GRU-2	R	0.910	0.876	0.956	0.914
	N		0.952	0.864	0.906
CNN-GRU	R	0.957	0.912	0.958	0.934
	N		0.950	0.934	0.942
CNN-Attention	R	0.968	0.932	0.962	0.946
	N		0.960	0.931	0.945

通过使用 CED\_Data set 数据集,使用相同预训练模型,比较不同训练模型谣言检测中效果。实验结果如表 5 所列,不同模型准确率的对比图如图 6 所示,实验结果表明,通过 BERT 预训练模型获得的句向量在不同的训练模型上进行训练在准确率方面仍然会有偏差,但是偏差幅度对比之前使用不同预训练模型要小。通过实验可以得出,SVM-TS 的准确率大概为 86.7%,其次依次是 GRU-1、CNN-GRU、GRU-2 模型,效果最好的是我们提出的 CNN-Attention 模型,准确率达到到了 95.3%,并且在召回率和 F1 值上体现出的效果也是众多模型中最好的。

表 5 不同检测方法下的实验结果对比(CED\_Data)

Table 5 Comparison of experimental results under different detection methods(CED\_Data)

模型	类别	准确率	精确率	召回率	F1
SVM-TS	R	0.867	0.846	0.894	0.868
	N		0.888	0.840	0.863
GRU-1	R	0.938	0.905	0.951	0.927
	N		0.948	0.901	0.924
GRU-2	R	0.948	0.914	0.955	0.934
	N		0.952	0.911	0.931
CNN-GRU	R	0.942	0.902	0.945	0.923
	N		0.938	0.922	0.930
CNN-Attention	R	0.953	0.919	0.960	0.939
	N		0.961	0.916	0.938

综上所述,我们的模型在两个不同的数据集上都表现出了良好的效果,通过使用 BERT 预训练模型可以提取更为全



面的特征表示,引入注意力机制的 Text CNN 模型考虑到提取特征对输出结果的影响力问题,通过给提取特征赋予权重可以有效提高谣言检测效率。

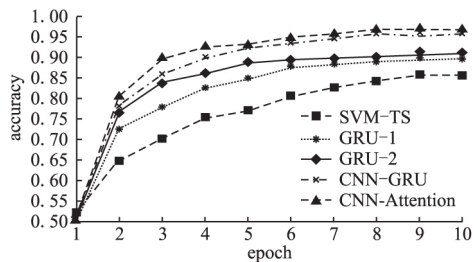


图5 模型准确率对比图(Ma数据)

Fig. 5 Model accuracy comparison chart(Ma data)

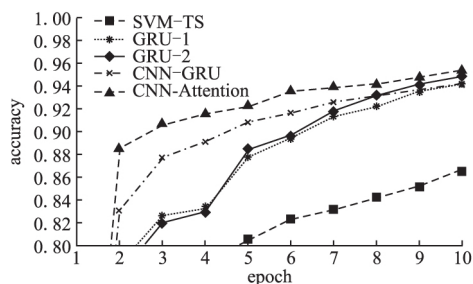


图6 模型准确率对比图(CED\_Data)

Fig. 6 Model accuracy comparison chart(CED\_Data)

CNN 模型中使用不同尺寸的过滤器可以提取出不同程度的语义特征,但是选择的过滤器尺寸过小,就会无法完整的提取句义表征;选择的过滤器尺寸过大,就会引入不必要的噪声。因此本环节选用了不同尺寸的过滤器进行实验。

表6 不同卷积尺寸结果表

Table 6 Results table for different convolution sizes

尺寸	准确率
3,4	0.955
3,5	0.956
4,5	0.958
3,4,5	0.968

通过表6与图7实验结果说明:卷积尺寸的不同会对准确率产生影响,但是总的来说影响并不是很大,选用卷积尺寸

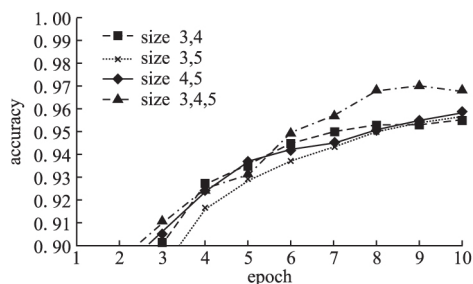


图7 不同卷积尺寸对比图

Fig. 7 Comparison chart of accuracy of different convolution sizes

为4,5的过滤器的准确率要比尺寸为3,4和3,5的过滤器的效果好,而选用尺寸为3,4和5的过滤器效果则优于前三个,

准确率高了大约1%,达到了96.8%。

## 4 结论

本研究提出了一种新的考虑注意力机制的谣言检测模型,能够有效解决微博谣言事件判断问题。采用 BERT 模型将部分下游自然语言处理任务转移到预训练模型,有效解决了提取特征不够全面的问题;采用引入注意力机制的 Text CNN 训练模型,根据所提取特征对输出结果影响力不同赋予不同权重,可以更有效的提取更为显著的文本深层特征,有效解决了提取特征对输出结果的影响力问题。本方法在真实微博数据集上经过实验验证具有较好的谣言检测效果。

## References:

- [1] Huo En-bi. Oxford advanced English dictionary [M]. Beijing: Commercial Press 2018.
- [2] Liang Gang, He Wen-bo, Xu Chun, et al. Rumor identification in microblogging systems based on users' behavior [J]. IEEE Transactions on Computational Social Systems 2015, 2(3): 99-108.
- [3] Yann Lecun, Bernhard Boser, Denker J. S., et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1(4): 541-551.
- [4] Jeffrey L. Elman. Finding structure in time [J]. Cognitive Science, 1990, 14(2): 179-211.
- [5] Yoon Kim. Convolutional neural networks for sentence classification [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2014: 1746-1751.
- [6] Chen Hui-min, Sun Mao-song, Tu Cun-chao, et al. Neural sentiment classification with user and product attention [C]//Proceedings of Conference on Empirical Methods in Natural Language Processing, Austin 2016: 1650-1659.
- [7] Ma Jing, Gao Wei, Mitra Prasenjit, et al. Detecting rumors from microblogs with recurrent neural networks [C]//Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, New York, USA: AAAI Press 2016: 3818-3824.
- [8] Liu Zheng, Wei Zhi-hua, Zhang Ren-xian. Rumor detection based on convolutional neural network [J]. Computer Application, 2017, 37(11): 3053-3056.
- [9] Li Li-zhao, Cai Guo-yong, Pan Jiao. C-GRU-based microblog rumor event detection method [J]. Journal of Shandong University (Engineering Edition), 2018, 49(2): 102-106.
- [10] Rahul Dey, Fathi M. Salemt. Gate-variants of gated recurrent unit (GRU) neural networks [C]//Proceedings of 2017 IEEE 60th International Midwest Symposium on Circuits and Systems, Boston, USA 2017: 1597-1600.
- [11] Wang Yang-qian, Han Hao, Ding Ye, et al. Learning contextual features with multi-head self-attention for fake news detection [C]//Proceedings of the ICCV 2019: Cognitive Computing-ICCC 2019: 132-142.
- [12] Fang Yong, Gao Jian, Huang Cheng, et al. Self multi-head attention-based convolutional neural networks for fake news detection [J]. PLOS ONE 2019, 14(9): 1-13.
- [13] Wu Lian-wei, Rao Yuan, Jin Hao-lin, et al. Different absorption from the same sharing: sifted multi-task learning for fake news detection [C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing 2019: 4644-4653.
- [14] Jacob Devlin, Chang Ming-wei, Kenton Lee, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//arXiv preprint arXiv: 1810.04805 2018.
- [15] Bai Xue, Chen Fu, Zhan Shao-bin. A study on sentiment computing and classification of Sina Weibo with Word2vec [C]//Proceedings of

- the 2014 IEEE International Congress on Big Data 2014:358-363.
- [16] Matthew E Peters, Mark Neumann, Mohit Iyyer et al. Deep contextualized word representations [C]//arXiv preprint arXiv: 1802.05365 2018.
- [17] Martin Sundermeyer, Ralf Schlüter, Hermann Ney. LSTM neural networks for language modeling [C]//Proceedings of the 13<sup>th</sup> Annual Conference of the International Speech Communication Association 2012: 194-197.
- [18] Ashish Vaswani, Noam Shazeer, Niki Parmar et al. Attention is all you need [C]//Proceedings of the Advances in Neural Information Processing Systems 2017: 6000-6010.
- [19] Song Chang-he, Yang Cheng, Chen Hui-min et al. CED: credible early detection of social media rumors [C]//Proceedings of the 2019 IEEE Transactions on Knowledge and Data Engineering, 2019: 1-13.
- [20] Zhang Yang-sen, Zheng Jia, Huang Gai-juan et al. Sentiment analysis method of Weibo based on dual attention model [J]. Journal of Tsinghua University ( Natural Science Edition) 2018, 58(2): 122-130.
- [21] Diederik P. Kingma, Jimmy Ba. ADAM: a method for stochastic optimization [C]//arXiv preprint arXiv: 1412.6980 2014.
- [22] Ma Jing, Gao Wei, Wei Zhong-yu et al. Detect rumors using time series of social context information on microblogging websites [C]//Proceedings of the 24th ACM International Conference on Information and Knowledge Management, Melbourne 2015: 1751-1754.

#### 附中文参考文献:

- [1] 霍恩比. 牛津高阶英语词典 [M]. 北京: 商务印书馆, 2018.
- [8] 刘政, 卫志华, 张韧弦. 基于卷积神经网络的谣言检测 [J]. 计算机应用, 2017, 37(11): 3053-3056.
- [9] 李力钊, 蔡国勇, 潘角. 基于 C-GRU 的微博谣言事件检测方法 [J]. 山东大学学报(工学版), 2018, 49(2): 102-106.
- [20] 张仰森, 郑佳, 黄改娟, 等. 基于双重注意力模型的微博情感分析 [J]. 清华大学学报(自然科学版), 2018, 58(2): 122-130.

### 征 稿 简 则

一、征稿范围 《小型微型计算机系统》杂志刊登文章的内容涵盖计算技术的各个领域(计算数学除外), 包括计算机科学理论、体系结构、计算机软件、数据库、网络与通讯、人工智能、信息安全、多媒体、计算机图形与图像、算法理论研究等各方面的学术论文。

二、来稿要求: 本刊主要刊登下述各类原始文稿:

1. 学术论文: 科研成果的有创新、有见解的完整论述。对该领域的研究与发展有促进意义。论文字数最好在 10000 字左右(含图表)。
2. 综述: 对新兴的或活跃的学术领域或技术开发的现状及发展趋势的全面、客观的综合评述(各类综述稿件一经录用三个月见刊)。
3. 技术报告: 在国内具有影响的重大科研项目的完整的技术总结。

三、注意事项

1. 来稿务求做到论点明确、条理清晰、数据可靠、叙述简练, 词义通达。
2. 来稿必须是作者自己的科研成果, 无署名和版权争议。引用他人成果必须注明出处。
3. 本刊采用在线投稿方式, 可登陆 <http://xwxt.sict.ac.cn/> 进行在线投稿。
4. 格式要求: 题目(中、英文)、摘要(中、英文)、作者的真实姓名(中、英文)、作者的单位、城市(中、英文)、邮政编码、E-mail(便于联系的)、关键词(中、英文 4~7 个)、中图分类号、作者简介、基金项目。

(1) 英文部分的作者姓名使用汉语拼音, 单位英文名称须给出英文全称, 不要使用缩略语;

(2) 作者简介包含作者姓名、性别、出生年、最高学历、技术职称、研究方向(若作者中有中国计算机学会(CCF)会员, 请注明, 并给出会员号)。凡第一作者为 CCF 会员/高级会员/学生会员者, 并在初次上传稿件中注明的, 将享受版面费八五折优惠;

(3) 基金项目的类别与项目编号。

5. 中、英摘要: 文章摘要具有独立性和自明性, 含正文等量的主要信息, 一般为 250~300 字, 采用第三人称表述。

6. 参考文献: 未公开发表的文献不得列入。文后所列参考文献统一排序, 且必须在正文中引用。中文参考文献应给出对应的英文译文。其具体书写格式为:

(1) 图书 [编号] 作者姓名(姓在前, 名在后), 书名, 出版社地址, 出版社, 出版年。

(2) 期刊 [编号] 作者姓名、文章题目、刊物名称, 出版年, 卷号(期号): 起止页码。

(3) 会议论文 [编号] 作者姓名, 论文题目。见: 编者、论文集名称、出版地: 出版年, 起止页码。

(4) 网络文献: 请给出文献作者或单位名, 文章题目、网址、发布日期。

7. 插图和表: 插图必须精绘并用计算机激光打印, 一般不超过 7 幅。图应结构紧凑, 不加底纹, 不要做成彩色的, 图宽最好不要超过 8 厘米。图内字号统一使用 6 号宋体, 字迹、曲线清晰。必要时给出坐标名称和单位。每个图、表均给出中英文图注(如“图 1: \* \* \* 图;” “Fig. 1: \* \* \*”)和表注(如“表 1: \* \* \* 表”, “Table 1: \* \* \*”)。

8. 计量单位: 稿件中一律使用《中华人民共和国法定计量单位》。外文和公式中应分清大、小写和正、斜体, 上、下角的字母、数码位置准确。易混淆的字母或符号, 请在第一次出现时标注清楚。

9. 本刊在收到作者稿件经初审后立即给作者电子邮箱发“稿件收到通知”。除作者另有明确要求外, 本刊原则上只与第一作者联系, 作者投稿后若 4 个月无消息, 可自行改投它刊。通过初审的稿件将收到本刊给予的编号, 并需邮寄审稿费。

10. 本刊对不拟录用的稿件只发给“退稿通知”, 恕不退回原稿, 请自留底稿。

11. 稿件一经发表, 将酌致稿酬, 并寄送样刊。

本刊文章现被国内外多家数据库收录, 作者著作权使用费与本刊稿酬一并给付, 作者若不同意将文章收录, 请在投稿时说明。

编辑部地址: 沈阳市浑南区南屏东路 16 号《小型微型计算机系统》编辑部 邮政编码: 110168

电 话: 024-24696120 E-mail: xwjt@sict.ac.cn 网 址: <http://xwxt.sict.ac.cn>