

# 融合时间序列与卷积神经网络的网络谣言检测

汪建梅, 彭云, 余晨钰

(江西师范大学 计算机信息工程学院, 南昌 330022)

E-mail: pengyun@jxnu.edu.cn

**摘要:** 针对卷积神经网络检测网络谣言没有考虑到谣言各生命周期之间的深层特征以及训练参数过于庞大问题, 提出了一种融合时间序列和卷积神经网络的谣言检测算法(CNN-TS). 首先将微博中的谣言事件向量化, 其次将这些向量根据时间顺序分组为潜伏期、滋生期、蔓延期、消退期, 然后通过卷积神经网络隐含层的学习训练来挖掘表示各阶段文本深层的特征, 最后通过修改分类函数进一步加强学习效果. 实验结果表明, 该检测算法在准确率、精确率及 F1 值上均优于支持向量机与卷积神经网络, 能够准确识别谣言事件.

**关键词:** 卷积神经网络; 谣言检测; 时间序列; 分类函数

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2022)05-1020-07

## Network Rumor Detection Combining Time Series and Convolutional Neural Network

WANG Jian-mei, PENG Yun, YU Chen-yu

(School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China)

**Abstract:** Aiming at the problem that the convolutional neural network detects network rumors without considering the deep features between the life cycles of the rumors and the excessive training parameters, a rumor detection algorithm combining time series and convolutional neural networks is proposed. First, the rumors in Weibo are vectorized, and then these vectors are grouped into incubation period, breeding period, spreading period, and fading period according to the time sequence, and then through the learning and training of the hidden layer of the convolutional neural network to mine the deep text of each stage. Finally, the learning effect is further enhanced by modifying the classification function. Experimental results show that the proposed method can accurately identify rumor events and is superior to comparison algorithms such as support vector machines and convolutional neural networks in terms of accuracy, precision and F1 index.

**Key words:** convolutional neural network; rumor detection; time series; classification function

## 1 引言

截至2020年6月,我国网民规模达到9.40亿,互联网普及率达67.0%<sup>[1]</sup>. 随着互联网的普及,大多数用户不断从社交媒体获取信息,同时通过社交媒体表达和传播信息. 随着在线社交媒体的迅速发展,大量不可靠的谣言信息可能会大量涌现. 社交媒体上的谣言信息泛滥,可能使人们难以将可信信息与众多复杂信息区分开,从而影响人们的正常生活秩序,尤其是在突发公共事件(自然灾害,意外事故)下,公共卫生事件,社会保障事件,经济危机等),谣言广泛传播可能极具破坏性. 因此,在谣言的早期阶段自动有效地识别网络中的谣言具有重要意义.

谣言检测的主要任务是检测微博帖子的相关事件是否为谣言,不是判断某一条具体微博是否为谣言. 现有的谣言检测模型可以大致分为两类:基于传统机器学习的谣言检测和基于深度学习的谣言检测.

基于传统机器学习的谣言识别模型主要是利用了微博文

本的符号特征、含有的链接特征、关键词分布特征和时间差等手工制作的特征,在综合用户特征和传播特征,利用SVM分类学习方法对微博进行分类<sup>[1-3]</sup>,然后在此基础上提出不同的系统识别框架. V. Qazvinian等<sup>[4]</sup>通过分析Twitter文本的浅层文本特征、行为特征和元素特征,构建多个贝叶斯分类研究来识别谣言. 早期Mendoza等<sup>[5]</sup>通过分析Twitter谣言传播的结构,得出在出现重大事件时, Twitter更容易传播谣言,并且通过图传播分析传播源的可信性来判断事件的可信度. 还有学者通过分析Twitter文本特征、转发率、传播时间点特征计算可信度分数<sup>[6,7]</sup>与基尼系数(Gini)<sup>[8]</sup>来识别谣言. Vasu等<sup>[9]</sup>分别讨论了词袋、n\_grams、计数矢量器、TF-IDF这四个基本分类器对谣言识别的效果. 也有学者另辟蹊径,提出基于话题<sup>[10,11]</sup>的谣言检测框架,可以降低谣言在没有转发或人群反应的情况下的互动缺失影响. 李明彩等<sup>[12]</sup>提出最大熵模型,将信息中的词汇作为特征,计算文本的最大熵,得出谣言与非谣言的概率,并将该模型与传统的SVM、BP、Bayes<sup>[13]</sup>方法比较,得到较好结果. 罗嗣卿等<sup>[14]</sup>、蔡国永等<sup>[15]</sup>、曾子明

<sup>1</sup> <http://www.cnnic.net.cn/hlwfzyj/hlwzxbg/hlwjtbg/202009/P020200929546215182514.pdf>

收稿日期:2020-11-26 收修改稿日期:2021-01-20 基金项目:国家自然科学基金项目(61966017,61662032)资助;江西省高校人文社科项目(JC19121,JC19117)资助. 作者简介:汪建梅,女,1997年生,硕士,研究方向为自然语言处理;彭云,男,1972年生,博士,副教授,CCF会员,研究方向为数据挖掘、自然语言处理;余晨钰,女,1997年生,硕士,研究方向为自然语言处理.

等<sup>[16]</sup>都加入了谣言微博文本的情感特征,并利用 LDA 主题模型深入挖掘微博文本的主题分布特征,再采用分别采用随机森林、决策树、组合决策树算法进行谣言识别。传统的机器学习方法从单纯的手工制作特征到利用器和集成分类器,虽然所有降低劳动强度,但是仍然不能形成重要特性之间的高层交互。

为了挖掘动态复杂的社交媒体场景中的关键特征,深度神经网络是一个很好的选择。采用循环神经网络 GRU<sup>[17,18]</sup>、LSTM<sup>[19]</sup>、RNN<sup>[20,21]</sup> 等以及这些基础网络的变体来深层挖掘谣言微博文本以及谣言传播的特征,结果证明基于神经网络的谣言检测方法优于人工构造特征的谣言检测模型。Ma J 等<sup>[22]</sup>同样利用 RNN 来捕捉帖子的上下文信息随时间的变化。Zhihong Wang 等<sup>[23]</sup>在 GRU 的基础上添加了情绪词典和动态时间序列特征来优化 GRU 算法。循环神经网络更加关注时间序列带来的特征影响,并且对卷积神经网络的内部特征关系没有深入的研究。刘政等<sup>[24]</sup>提出卷积神经网络模型,将微博中的谣言事件向量化,通过卷积神经网络隐含层的学习训练来挖掘表示文本深层的特征,Santhoshkumar 等<sup>[25]</sup>提出一种基于特定因素的卷积神经网络(CNN)方法。采用两个并行的 CNN 进行谣言事件分类,然后再利用决策树将这两个 CNN 的输出组合在一起,并提供分类输出。在卷积神经网络的基础上衍生出了基于图卷积神经网络模型<sup>[26]</sup>的谣言识别方法,将微博谣言数据转换为图数据,再利用卷积神经网络进行训练有标注数据,通过更新图中节点权重,将该信息传递给无标注数据,大大降低了谣言数据标注的工作量。

CNN 在谣言识别中取得了很好的结果,但没有考虑到谣言各个生命周期之间的时间序列特征,对现有的卷积神经网络模型进行了改进。

主要贡献如下:

- 1) 提出卷积神经网络(CNN)与时间序列算法进行微博谣言识别;
- 2) 改进卷积神经网络(CNN)与时间序列算法的全连接层分类函数,增强模型泛化能力。

## 2 相关工作

在自然语言处理方面,卷积神经网络在语句分类<sup>[27]</sup>、语义分析<sup>[28]</sup>、实体关系分类<sup>[29]</sup>、注意力机制<sup>[30]</sup>等中都取得了很好的效果。在谣言识别方面,刘政提出基于卷积神经网络的模型,利用 Doc2Vec<sup>[31]</sup>将微博帖子向量化,并且将卷积核的大小设为帖子向量的长度,可以有效地提取帖子与帖子之间、帖子内部之间的特征。刘政等的模型大大降低了人工构造特征的劳动强度,并且学习到微博帖子之间的深层特征,但是并没有考虑谣言不同生命周期之间帖子的时间序列特征。Adel 等<sup>[30]</sup>对卷积神经网络识别谣言进行改进,添加不确定检测因素的注意力机制,该模型利用文本的单词向量序列、外部网络环境变化来检测事件的真假。Feng Yu 等<sup>[32]</sup>提出一个用于错误信息和早期发现任务的 CAMI 模型,收集事件数据集的数据分布,观察假信息和真信息的长尾分布,将微博帖子分组,

再利用 Doc2Vec 得到文本向量表示,作为卷积神经网络的输入。Feng Yu 等该模型检测帖子中的真假信息,虽考虑到帖子之间的时间序列特征,但是参数过于庞大,占据太多资源。Dazhen 等<sup>[33]</sup>在考虑内容特征和社交特征的基础上,还考虑了文本单词序列上下文特征,采用 LSTM 获取内容的双向序列上下文信息,然后将深层序列上下文信息与社会特征结合,用 CNN 学习内容与社会特征之间的联系。Jing MA 等<sup>[34]</sup>提出 DSTS 算法,使用时间序列来考虑社会语境特征在谣言传播过程中随时间变化的特征,在利用 SVM 算法进行谣言识别,但是并没有考虑到微博文本上下文之间或微博文本之间的深层特征。

综合分析,现有的利用卷积神经网络检测网络谣言的模型存在以下问题:

- 1) 卷积神经网络训练模型,模型过于复杂,参数过于庞大,占据过大资源;
- 2) 卷积神经网络并未很好处理谣言生命周期之间的时间序列特征;

## 3 CNN-TS 谣言检测模型

### 3.1 基本定义

实验的研究对象是微博谣言事件,关注的是微博发布的相关事件是否是谣言,而不是关心与该事件相关的微博帖子是否为谣言。微博事件一旦被检测为谣言,那么传播与该事件相关的帖子也应该被视为谣言。例如,“中美不同的教育方式的结果是:教育进展国际评估组织的调查显示,在 21 个被调查国家中,中国孩子的计算能力排名第一,想象力排名倒数第一,创造力排名倒数第五。此外,在中国的中小学生中,认为自己有好奇心和想象力的只占 4.7%,而希望培养想象力和创造力的只占 14.9%。”<sup>2</sup>这条谣言以及相关微博帖子“这种发展趋势很可怕...”“[给力]”“家长、社会、教育专家和行政官员是推手,值得反思,但目前看不到希望,因为大家都跟着感觉走,没有彻底改革的勇气。”等就构成了一个谣言事件,而本文判断的仅是“中美不同的教育方式的结果是:教育进展国际评估组织的调查显示,在 21 个被调查国家中,中国孩子的计算能力排名第一,想象力排名倒数第一,创造力排名倒数第五。此外,在中国的中小学生中,认为自己有好奇心和想象力的只占 4.7%,而希望培养想象力和创造力的只占 14.9%。”<sup>3</sup>这个微博事件是否是谣言,而其相关微博帖子是否是谣言并不关心。因此,针对融合时间序列与卷积神经网络模型给出相关的符号解释和定义。

**定义 1.** 谣言事件<sup>[34]</sup>。定义所有微博事件集合  $E = \{E_i\}$ , 其中  $E_i = \{m_{i,j}\}$ ,  $E_i$  是指第  $i$  个事件,包含与其相关的所有微博帖子,  $m_{i,j}$  是指第  $i$  个事的第  $j$  条微博帖子。要达到的目标就是判断  $E_{i,j}$  是否是谣言事件。

### 3.2 理论基础

#### 3.2.1 Doc2Vec 段落向量的构建

一个谣言事件有  $E_{i,j}$  若干条微博帖子  $m_{i,j}$ , 将每一条微博

<sup>2</sup> [https://www.sohu.com/a/74593586\\_101403](https://www.sohu.com/a/74593586_101403)

<sup>3</sup> [https://www.sohu.com/a/74593586\\_101403](https://www.sohu.com/a/74593586_101403)

帖子的文本视为谣言事件的一个段落文本. 根据 Doc2Vec<sup>[31]</sup> 中对段落文本向量的处理如式(1)所示.

$$\arg \max_{D, W} \frac{1}{N} \sum_{n=k}^{N-k} \log p(w_n | w_{n-k}, \dots, w_{n+k}) \quad (1)$$

其中,  $N$  是段落文本的词个数, 一个单词表示成的单词向量为  $W_n$ , 单词向量组成的集合是  $W$ ,  $g_j$  是一个段落向量,  $D$  是所有段落向量组成的集合.

利用 Softmax 函数对段落中的单词进行概率预测的公式如公式(2)、公式(3)所示.

$$p(w_n | w_{n-k}, \dots, w_{n+k}) = \frac{\exp(\theta^T x_n)}{\sum_i \exp(\theta^T x_i)} \quad (2)$$

$$x_n = h(g_j, w_{n-k}, \dots, w_{n+k}; D, W) \quad (3)$$

其中,  $\theta$  是 Softmax 函数的参数,  $h$  代表着连接函数或者平均函数.

### 3.2.2 卷积神经网络结构原理

党建武等<sup>[35]</sup>中描述了卷积神经网络的结构, 如图 1 所示. 深度学习能够实现数据的逐层转换, 深入提取隐藏特征.

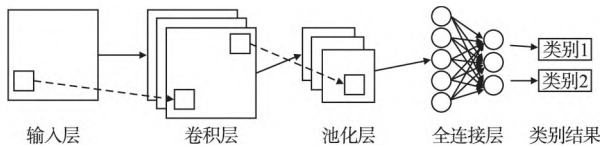


图 1 卷积神经网络结构图

Fig. 1 Structure of convolutional neural network

CNN 对输入数据进行一系列的卷积与池化操作. 其中, 卷积层是卷积神经网络的核心, 卷积层对输入层数据中的每个特征进行局部感知, 然后在更高层次上对局部特征进行综合运算, 从而得到全局信息. 卷积层的计算公式如(4)所示.

$$F = f(X \otimes W + b) \quad (4)$$

其中,  $F$  为卷积层中多个卷积核的输出特征矩阵,  $X$  为输入数据向量,  $f$  为非线性激活函数,  $\otimes$  为卷积操作,  $W$  为各个卷积核的权重矩阵,  $b$  为各个卷积核的偏置项. 非线性激活函数是对卷积层的输出进行非线性映射操作, 以增强函数的特定数值性质, ReLU 函数通常用于中间层的非线性激活, 因为其迭代速度快, Softmax 函数一般用于最后一层的激活功能.

池化层是池化卷积层的输出, 保留主要特征, 压缩数据和参数的数量, 减少过度拟合, 提高模型的容错性, 计算如公式(5)所示.

$$P = \text{pool}(F) \quad (5)$$

其中,  $P$  表示池化层的特征输出结果,  $\text{pool}$  表示池化函数, 一般分为 Max Pooling 和 Average Pooling.

全连接层是以平滑化后的特征矩阵作为全连通层的输入矩阵, 然后以激活函数作为分类函数, 输出每个分类标签的概率.

### 3.2.3 K 折交叉验证法

K 折交叉验证法是减少过拟合, 解决数据量不够多, 导致检测效果差的常用操作方法. K 折交叉验证就是将给定的数据集均分为  $K$  份, 经过若干次迭代, 第  $K$  次迭代取第  $K$  折数据作为测试集, 另外  $K-1$  份数据作为训练集, 直到  $K$  折中的每折数据都用作测试集. 然后将总迭代的测试结果与训练结果的平均值作为最后的测试与训练结果. K 折交叉验证法的

结构原理图如图 2 所示.

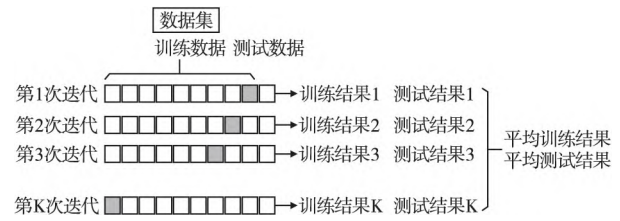


图 2 K 折交叉验证法原理图

Fig. 2 Schematic diagram of K folding cross verification method

### 3.3 CNN-TS 模型结构

#### 3.3.1 CNN-TS 模型流程图

提出的谣言检测的流程图如图 3 所示. 首先获取微博中的事件数据, 其中包含谣言数据与非谣言数据, 将其随机打乱, 保证训练集、验证集和测试集的数据分布都是相同的; 其次, 将每个事件数据作为一个整体, 并对与之相关的每个微博帖子进行文本向量化. 然后将该事件数据的文本向量按对应的时间序列进行排序, 均分为 20 组; 最后将这 20 组向量作为输入矩阵, 进行卷积神经网络模型的训练.

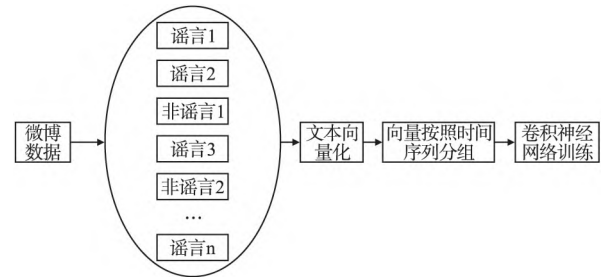


图 3 微博谣言检测流程图

Fig. 3 Flow chart of Microblog rumor detection

#### 3.3.2 微博帖子的时间序列分组

一个有影响力的微博事件的帖子数量至少有几百条, 甚至达到几万条, 不同事件的帖子数量差异大. 但是谣言事件的生存周期分为潜伏期、滋生期、蔓延期、消退期, 在这 4 个时间段内和 4 个时间段之间的微博帖子文本特征具有相似的特征及变化趋势. 因此将这些按照时间序列排好的相邻帖子视为一个组, 代表事件的一个特定生命周期. 这样做的考虑是: 更多关注各个时间段内的微博帖子内容的特征关系, 并且可以提取出各个时间段之间的微博文本的特征, 而不是单独一条微博帖子之间的特征关系. 一个微博谣言事件的发展在各个时间段内的帖子内容会有差异, 而且这样做能够大大减少模型输入数据的复杂度. 帖子的时间序列分割参考 MA 等<sup>[34]</sup>中的操作.

对于一个事件  $E_i$ , 其帖子的起始时间为  $\text{time\_Begin}_i$ , 帖子的结束时间为  $\text{time\_End}_i$ , 将每条微博  $m_{i,j}$  的时间  $t_{i,j}$  转换为 0 到  $N$  之间内的时间戳,  $N$  是时间间隔数, 在本文中  $N$  取 20. 那么对于事件  $E_i$  的帖子时间间隔  $\text{time\_Interval}_i$  和每一条帖子的时间戳数的  $\text{time\_Stamp}_{m_{i,j}}$  计算公式如公式(6)和公式(7)所示.

$$\text{time\_Interval}_i = \left\lfloor \frac{\text{time\_End}_i - \text{time\_Begin}_i}{N} \right\rfloor \quad (6)$$



$$time\_Stamp_{m_{i,j}} = \left\lfloor \frac{t_{i,j} - time\_Begin_i}{time\_Interval_i} \right\rfloor \quad (7)$$

然后收集所有相关微博帖子的时间戳,并从每个事件的所有时间戳中减去相应事件的开始时间戳.然后将这些时间戳规范化为 0-1 尺度.最后将整个时间戳按时间顺序平均分割为 20 个份额,每个时间窗口内的帖子表示为如式(8)所示.

$$T_i = [time\_Stamp_{m_{i,j-1}}, time\_Stamp_{m_{i,j}}], \quad j = 1, 2, 3, \dots, 20 \quad (8)$$

然后将分组后的帖子  $T_i$  利用 Doc2Vect 文本向量化,作为输入矩阵传入卷积神经网络训练.第  $i$  个微博谣言事件的相关帖子分组后,得到的向量矩阵的表示如式(9)所示.

$$V(E_i) = (F_{i,1}, F_{i,2}, \dots, F_{i,20}) \quad (9)$$

其中,  $E_i$  是第  $i$  个谣言事件,  $F_{i,20}$  是分组后的 20 个特征向量.

向量按时间序列分组的模型结构如图 4 所示.

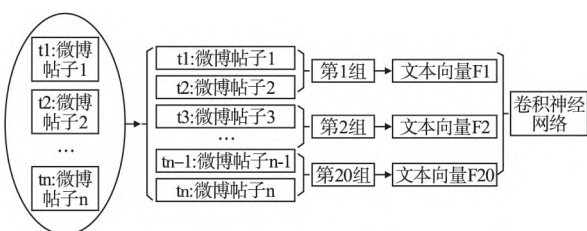


图 4 时间序列分组的模型结构图

Fig. 4 Model structure diagram of time series grouping

### 3.3.3 CNN-TS 的卷积神经网络模型

用于卷积神经网络的模型结构如图 5 所示.

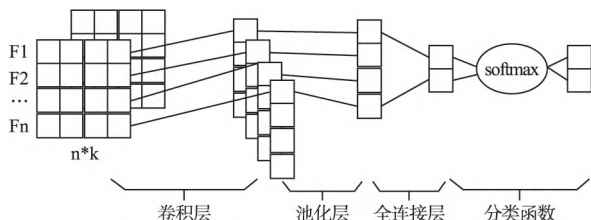


图 5 微博谣言的卷积神经网络模型图

Fig. 5 Convolutional neural network model diagram of Microblog rumors

其中  $F_{i,j} \in R^k$ , 表示谣言事件  $E_i$  的第  $j$  条微博的文本向量, 其维度为  $k$ . 一个包含  $n$  条相关微博的谣言事件的文本向量可以表示为如式(10)所示.

$$E_i = F_{i1} \oplus F_{i2} \oplus F_{i3} \oplus \dots \oplus F_{in} \quad (10)$$

其中  $\oplus$  表示串联操作.

卷积层: 利用卷积核  $w \in R^{h \times k}$  对输入向量进行卷积操作得到新的特征向量. 例如, 第 1 层隐含层的 3 个卷积核的操作计算公式如式(11).

$$\begin{aligned} a_1^{(1)} &= f(W_1^{(1)} F + b_1^{(1)}) \\ a_2^{(1)} &= f(W_2^{(1)} F + b_2^{(1)}) \\ a_3^{(1)} &= f(W_3^{(1)} F + b_3^{(1)}) \end{aligned} \quad (11)$$

其中,  $W_1^{(1)}, W_2^{(1)}, W_3^{(1)}$  分别是每一个  $f$  卷积核的各个输入的权重,  $b_1^{(1)}, b_2^{(1)}, b_3^{(1)}$  分别是每一个卷积核的偏置值,  $f$  是非线性函数, 本文模型采用的是 ReLU 激活函数.  $a = [a_1^{(1)}, a_2^{(1)},$

$a_3^{(1)}]$  是新生成的特征向量, 作为第 2 层隐含层的输入矩阵, 以此类推, 执行之后的每一层操作.

池化层: 对卷积后的特征矩阵进行压缩, 一方面使特征矩阵变小, 另一方面提取主要特征. 本实验采用的是 maxpooling, 将特征矩阵切成几个区域, 取其最大值, 保持原有的矩阵特征得出池化后的特征值.

全连接层: 将经过多个卷积层与池化层后的特征矩阵作为全连接层的输入矩阵, 一般采用 Softmax 函数作为分类函数进行操作, 其输出为每一个分类标签的概率.

卷积核 filter 的设置是采用刘政等<sup>[24]</sup>文献中的操作: 将 filter 的宽度等于输入矩阵的宽度. 每一行是微博事件的按时间序列分组后的微博帖子向量, 这样设置, 能够更加关注微博帖子文本之间的特征关系, 而不是微博帖子文本内部的词语之间的特征关系.

### 3.4 模型结构的改进

卷积神经网络对标签的分类方法主要由两部分组成: 一个是分数函数, 它将原始数据映射到类别分数; 另一个是损失函数, 它量化了预测分数和实际分数之间的一致性, 主要用来表示一个标签的概率. 神经网络模型可以通过更新评分函数的参数使损失函数值最小化, 达到模型最优的目的. 分类函数 softmax 与 SVM 的损失函数是不同的, 它们的区别如下:

1) softmax 函数的损失函数是 cross-entropyloss, 如式(12)、式(13)所示.

$$L_i = -\log\left(\frac{e^{f_{yi}}}{\sum_j e^j}\right) \quad (12)$$

$$P_{yi} = \frac{e^{f_{yi}}}{\sum_j e^j} \quad (13)$$

其中,  $f_{yi}$  是把标签  $i$  预测为  $y_i$  的分数,  $j$  是分类结果中的一个,  $L_i$  是在  $[0, 1]$  之间的损失.  $P_{yi}$  是数据分类的 Softmax 值, 即分类正确的概率, Softmax 值越大, 该模型的损失也就越小.

2) SVM 函数的损失函数是 hinge-loss. 分类器输入样本  $x_i$  后, 第  $j$  个类别的评分是:

$$s_j = f(x_i, W)_j \quad (14)$$

则对第  $i$  个样本的损失为:

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + \Delta) \quad (15)$$

该计算公式只关注正确分数比不正确大  $\Delta$  的值, 在用 hinge-loss 训练时, 需考虑预期的损失函数该有多大.

因此, 先用 Softmax 函数对输入数据进行预训练, 得到预期的损失函数值大小, 然后再次输入模型, 将分类函数 Softmax 改为 SVM, 解决模型泛化能力不足的问题.

CNN-TS 模型采用的分类函数是 Softmax, 针对此问题, 修改模型中的分类函数 Softmax, 采用 SVM 对池化后的特征向量进行最终分类, 模型结构图如图 6 所示.

## 4 实验与分析

### 4.1 实验数据

采用的实验数据是 Ma 等<sup>[22]</sup>文献中的公开数据集. 该数据集集中的微博数据是从新浪社区管理中心获得的谣言与非谣言事件, 该中心报道了各种不实信息. 并通过爬虫捕获了这些事件的原始帖子以及所有相关的转发/回复消息. 该数据集的

统计情况如表 1 所示。

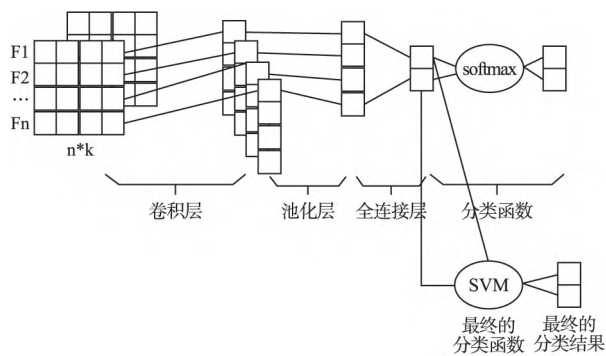


图 6 SVM 作为分类函数的 CNN-TS 改进模型结构图

Fig. 6 CNN-TS improved model structure diagram with SVM as a classification function

表 1 微博数据的详情统计情况表

Table 1 Detailed statistics table of microblog data

统计信息	微博
用户个数	2746818
帖子个数	3805656
事件个数	4664
谣言事件个数	2313
非谣言数据个数	2351
事件的平均时间跨度	2460.7 小时
事件帖子的平均个数	816
事件帖子的最多个数	59138
事件帖子的最少个数	10

## 4.2 数据预处理

收集到的实验数据包含大量噪声,会严重影响谣言识别效果,因此有必要在实验前进行数据噪声处理。

1) 去除标点符号。在本实验中,只考虑帖子的文本特征,但某些标点符号,如“!”也可作为一种特征,因此按照“!”“?”优先级保留,其余则需移除。

2) 去除表情符号。本实验中只考虑文本的特征,不考虑文本所展现出来的情感特征,因此需要移除表情符号,例如[开心]。

3) 去除特殊符号。微博帖子包含无关信息,例如超链接和@某某,可以通过正则表达式匹配将其从数据集中删除。

4) 分词。微博帖子为短文本,在进行文本处理时,其最小粒度应为词或词语,因此需要进行分词操作。采用的是 Python 中的 jieba 分词处理。

5) 去停用词。停用词是指不对文本特征有任何贡献的词。它不仅包含标点符号,还包含语气词,人称词和地点词之类的词。但语气词也代表了文本的一种隐含特征,因此本实验不去除语气词。本实验采用的是哈工大停用词表 hit\_stop-words。

微博文本数据进行噪声处理,可以大大降低文本无关信息的噪声影响,大大增强模型检测效果。

## 4.3 实验评价指标

准确率 (accuracy), 精确率 (precision), 召回率 (recall), F1 值常被用于评判实验效果,本文同样选取这 4 个指标。具

体的计算方式如下:

$$\text{准确率 (accuracy)} = \frac{\text{预测正确的样本数}}{\text{总样本数}}$$

$$\text{精确率 (precision)} = \frac{\text{预测正确的正样本数}}{\text{预测为正的样本数}}$$

$$\text{召回率 (recall)} = \frac{\text{预测正确的正样本数}}{\text{总样本中的正样本数}}$$

$$F1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (16)$$

预测正确的样本数是样本数中真谣言的样本预测为真和假谣言的样本预测为假的个数;预测正确的正样本数是预测正确的样本数中预测为真的个数;预测为正的样本数为不论真实标签为真还是为假,都将其预测为真的个数;总样本中的正样本数为总样本中真实标签为真的个数。准确率反映的是检测方法的准确性,召回率反映的是检测方法的覆盖率,这两个指标的数值越大,表明检测方法的整体效果越好。

## 4.4 实验设置

将提出的模型与刘政等<sup>[24]</sup>, Abhishek 等<sup>[36]</sup>的模型进行了对比。

刘政等<sup>[24]</sup>的模型:刘政等提出采用卷积神经网络模型进行微博谣言的检测,取得了良好的效果,本文的模型就是卷积神经网络模型的一种改进。

Abhishek 等<sup>[36]</sup>的模型:Abhishek 等的模型是利用朴素贝叶斯、随机森林和支持向量机这 3 种基本的机器学习模型对文本特征、用户特征以及文本与用户组合的特征进行训练、识别。本实验采用 Abhishek 中的基于 RBF 核模型的 SVM 方法,对组合特征进行训练。本文模型对分类函数的改进是利用了 RBF 核模型的 SVM 方法。

提出的模型 (CNN-TS): 时间序列和卷积神经网络模型的融合由卷积层、池化层和两个完全连接层组成。卷积层的宽度与输入数据矩阵的宽度相同。输入矩阵是使用 Doc2Vec 训练的文本向量矩阵。

超参数设置:超参数的设置是参考刘政等<sup>[24]</sup>的参数,filter 的高度为 3, dropout rate 为 0.5。每一条微博的维度 k 设为 50。

数据集的划分:K 折交叉验证方法用于评估模型的预测性能。经过训练的模型对新数据的性能可以在一定程度上减少过度拟合的情况,并可以从有限的数据中获得尽可能多的有效信息,将 K 设置为 10,并多次计算每个评估指标的 K 折交叉验证平均值。

## 4.5 实验结果

模型 CNN-TS 与 CNN, SVM-RBF 的实验结果比较如表 2 所示。

表 2 模型的实验结果比较图

Table 2 Comparison of experimental results of the models

模型	accuracy	precision	recall	F1
SVM-RBF	0.769	0.778	0.748	0.763
CNN	0.795	0.852	0.706	0.771
CNN-TS (Softmax)	0.851	0.839	0.866	0.852
CNN-TS (SVM)	0.866	0.868	0.860	0.863

从表 2 中可以看出,融合时间序列与卷积神经网络的模型优于基于支持向量机的 RBF 核模型和普通的卷积神经网络

络模型,比支持向量机模型准确率提高 0.082,精确率提高 0.061,F1 值提高 0.089;比普通的卷积神经网络模型准确率提高 0.056,recall 提高 0.16,F1 值提高 0.081.因此融合了时间序列的卷积神经网络模型是优于刘政等<sup>[24]</sup>提出的基本卷积神经网络模型以及 Abhishek 等<sup>[36]</sup>运用的支持向量机模型的。

同时对 CNN-TS 的模型进行改进,将最终的分类函数改为 SVM 得到改进后的模型,最终的实验结果比支持向量机模型准确率提高 0.097,精确率提高 0.09,recall 值提高 0.112,F1 值提高 0.1;比普通的卷积神经网络模型准确率提高 0.071,精确率提高 0.016,recall 提高 0.154,F1 值提高 0.092;比 CNN-TS 准确率提高 0.015,精确率提高 0.029,F1 值提高 0.013.CNN-TS-SVM 模型除了 recall 略低 0.006 外,其他准确率、精确率、F1 值都高于 CNN-TS 模型,因此在 CNN-TS 模型的基础上,修改分类函数,改进后的 CNN-TS (SVM)模型是优于 CNN-TS 模型的。

模型的输入数据矩阵与刘政等<sup>[24]</sup>的卷积神经网络的输入矩阵的规模对比如表 3 所示。

表 3 模型的输入数据矩阵规模对比表

Table 3 Model input data matrix size comparison table

模型	微博
CNN	59138 × 50
CNN-TS	20 × 50
CNN-TS-SVM	20 × 50

模型融合了时间序列后,输入矩阵的规模大大减小,大大减少了占用系统内存,减少了训练时间,同时准确率,召回率,F1 值都优于刘政等<sup>[24]</sup>提出的 CNN 模型.而且迭代步数在 100 步内就可以收敛,同时不出现过拟合的情况。

## 5 结 语

提出了融合时间序列与卷积神经网络的谣言检测算法模型,可以大大降低模型的复杂度,减小数据集的规模,又考虑到时间序列上的特征变化.提出的方法是通过将与某一个事件相关的所有微博帖子按时间序列分成若干组,再通过 Doc2Vec 处理成文本向量,最后作为卷积神经网络的输入进行训练预测.结果显示,它们优于卷积神经网络模型.为了进一步改进,用 Softmax 分类函数的神经网络模型训练好后,再一次用 SVM 模型进行最后的分类,结果显示,这种改进优于未改进之前的状态.如何进一步优化神经网络模型,添加谣言文本和图片的特征进行更彻底的实验将是下一个进一步的研究工作。

## References:

- [1] He Gang, Lv Xue-qiang, Li Zhuo, et al. Automatic rumor identification on microblog[J]. Library and Information Service, 2013, 57(23):114-120.
- [2] Fan Yang, Yang Liu, Yu Xiao-hui, et, al. Automatic detection of rumor on sina weibo[C]//ACM SIGKDD Workshop on Mining Data Semantics, New York, 2012, 13:1-7.
- [3] Rosa Sicilia, Stella Lo Giudice, Yulong Pei. Twitter rumour detec-

- tion in the health domain[J]. Expert Systems with Applications, 2020, 110(11):33-40.
- [4] Vahed Qazvinian, Emily Rosengren, Dragomir R Radev, et al. Rumor has it:identifying misinformation in microblogs[C]//Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, 2011:1589-1599.
- [5] Mendoza M, Poblete B. Twitter under crisis: can we trust what we RT? [C]//1st Workshop on Social Media Analytics, 2011:71-79.
- [6] Takahashi T, Igata N. Rumor detection on twitter[C]//6th International Conference on Soft Computing and Intelligent Systems, and 13th International Symposium on Advanced Intelligence Systems, IEEE, 2012:452-457.
- [7] Gupta M, Zhao P, Han J. Evaluating event credibility on twitter [C]//12th SIAM International Conference on Data Mining, 2012:153-164.
- [8] Kumar, Geethakumari. Detecting misinformation in online social networks using cognitive psychology[J]. Kumar and Geethakumari Human-centric Computing and Information Sciences, 2014, 9(14):4-14.
- [9] Vasu Agarwal, H Parveen Sultana, Srijan Malhotra, et al. Analysis of classifiers for fake news detection[J]. Procedia Computer Science, 2019, 165(11):377-383.
- [10] Yang Wen-tai, Liang Gang, Xie Kai, et al. Rumor detection method based on burst topic detection and domain expert discovery[J]. Journal of Computer Applications, 2017, 37(10):2799-2805.
- [11] Fan Xu, Victor Sheng, Mingwen Wang. Near real-time topic-driven rumor detection in source microblogs[J]. Knowledge-Based Systems, 2020, 207(6380):106391-106400.
- [12] Li Ming-cai, Liu Feng-ming. Research on Maximum entropy model for rumor detection[J]. Journal of Chinese Computer Systems, 2017, 38(7):1475-1478.
- [13] Shuo Yang, Kai Shu, Suhang Wang, et, al. Unsupervised fake news detection on social media: a generative approach[C]//33rd AAAI Conference on Artificial Intelligence (AAAI), 2019:5644-5651.
- [14] Luo Si-qing, Wang Jia-yu, Li Bing-ke. Study on rumor discrimination based on improved combinatorial optimization decision tree [J]. Computer Simulation, 2018, 35(2):219-223.
- [15] Cai Guo-yong, Bi Meng-ying, Liu Jian-xing. A novel rumor detection method based on features of labeled cascade propagation tree [J]. Computer Science and Engineering, 2018, 40(8):1488-1495.
- [16] Zeng Zi-ming, Wang Jing. Research on Microblog rumor identification based on LDA and random forest[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(1):89-96.
- [17] Li Li-zhao, Cai Guo-yong, Pan Jiao. A microblog rumor events detection method based on C-GRU[J]. Journal of Shandong University (Engineering Science), 2019, 49(2):102-106 + 115.
- [18] Wu Zhi-yuan, Pi De-chang, Chen Jun-fu, et al. Rumor detection based on propagation graph neural network with attention mechanism[J]. Expert Systems with Applications, 2020, 158(11):113595-113611.
- [19] Yin Peng-bo, Pan Wei-min, Peng Cheng, et al. Research on early detection of weibo rumors based on user characteristics analysis [J]. Journal of Intelligence, 2020, 39(7):81-86.
- [20] Ruchansky N, Seo S, Liu Y. Csi: a hybrid deep model for fake news detection [C]//ACM on Conference on Information and



- Knowledge Management, 2017; 797-806.
- [21] Xu Y, Wang C, Dan Z, et al. Deep recurrent neural network and data filtering for rumor detection on sina weibo[J]. Symmetry, 2019, 11(11): 1408, doi:10.3390/sym11111408.
- [22] Jing Ma, Wei Gao, Prasenjit Mitra, et al. Detecting rumors from microblogs with recurrent neural networks[C]//25th International Joint Conference on Artificial Intelligence (IJCAI' 16), 2016; 3818-3824.
- [23] Wang Zhi-hong, Guo Yi, Wang Jia-hui, et al. Rumor events detection from Chinese microblogs via sentiments enhancement[J]. IEEE Access, 2019, 7(10): 103000-103018.
- [24] Liu Zheng, Wei Zhi-hua, Zhang Ren-xian. Rumor detection based on convolutional neural network[J]. Journal of Computer Applications, 2017, 37(11): 3053-3056 + 3100.
- [25] Santhoshkumar S, Babu L. Earlier detection of rumors in online social networks using certainty-factor-based convolutional neural networks[J]. Social Network Analysis and Mining, 2020, 10(11): 10-20.
- [26] Mi Yuan, Tang Heng-liang. Rumor identification research based on graph convolutional network[J/OL]. Computer Engineering and Applications; 1-9, <http://kns.cnki.net/kcms/detail/11.2127.tp.20201020.1727.018.html>, 2020-11-18.
- [27] Guo Chong. Modeling and simulation of online customer loyalty based on analysis of big data[J]. Computer Simulation, 2015, 32(10): 239-242 + 304.
- [28] Yih Wen-tau, He Xiaodong, Meek, et al. Semantic parsing for single-relation question answering[C]//52nd Annual Meeting of the Association for Computational Linguistics, 2014, 2: 643-648.
- [29] Vu NT, Adel H, Gupta P, et al. Combining recurrent and convolutional neural networks for relation classification[C]//The North American Chapter of the Association for Computational Linguistics; Human Language Technologies (NAACL-HLT), 2016; 534-539.
- [30] Adel H, Schutze H. Exploring different dimensions of attention for uncertainty detection[C]//15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2017, 1: 22-34.
- [31] Quoc Le, Tomas Mikolov. Distributed representations of sentences and documents[C]//31st International Conference on International Conference on Machine Learning (ICML' 14), 2014, 32: 1188-1196.
- [32] Feng Yu, Qiang Liu, Shu Wu, et al. A convolutional approach for misinformation identification[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence, 2017; 3901-3907.
- [33] Dazhen L, Ben M, Donglin C, et al. Chinese microblog rumor detection based on deep sequence context[J]. Concurrency and Computation: Practice and Experience, 2019, 31(23): e4508. 1-e4508. 12.
- [34] Jing Ma, Wei Gao, Zhongyu Wei, et al. Detect Rumors using time series of social context information on microblogging websites[C]//24th ACM International on Conference on Information and Knowledge Management (CIKM' 15), 2015; 751-1754.
- [35] Dang Jian-wu, Cong Xiao-qing. Research on Hybrid stock index forecasting model based on CNN and GRU[J/OL]. Computer Engineering and Applications; 1-11, <http://kns.cnki.net/kcms/detail/11.2127.TP.20201027.1415.016.html>, 2020-11-11.
- [36] Abhishek Vijeev, Anushreya Mahapatra, Arundhati Shyamkrishna, et al. A Hybrid approach to rumour detection in microblogging platforms[C]//International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018; 337-342.

#### 附中文参考文献:

- [1] 贺刚, 吕学强, 李卓, 等. 微博谣言识别研究[J]. 图书情报工作, 2013, 57(23): 114-120.
- [10] 杨文太, 梁刚, 谢凯, 等. 基于突发话题和领域专家的微博谣言检测方法[J]. 计算机应用, 2017, 37(10): 2799-2805.
- [12] 李明彩, 刘凤鸣. 谣言信息识别的最大熵模型研究[J]. 小型微型计算机系统, 2017, 38(7): 1475-1478.
- [14] 罗嗣卿, 王佳玉, 李冰珂. 改进的组合优化决策树谣言判别方法研究[J]. 计算机仿真, 2018, 35(2): 219-223.
- [15] 蔡国永, 毕梦莹, 刘建兴. 基于标记信息级联传播树特征的谣言检测新方法[J]. 计算机工程与科学, 2018, 40(8): 1488-1495.
- [16] 曾子明, 王婧. 基于LDA和随机森林的微博谣言识别研究——以2016年雾霾谣言为例[J]. 情报学报, 2019, 38(1): 89-96.
- [17] 李力钊, 蔡国永, 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), 2019, 49(2): 102-106 + 115.
- [19] 尹鹏博, 潘伟民, 彭成, 等. 基于用户特征分析的微博谣言早期检测研究[J]. 情报杂志, 2020, 39(7): 81-86.
- [24] 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 + 3100.
- [26] 米源, 唐恒亮. 基于图卷积网络的谣言鉴别研究[J/OL]. 计算机工程与应用; 1-9, <http://kns.cnki.net/kcms/detail/11.2127.tp.20201020.1727.018.html>, 2020-11-18.
- [27] 郭崇. 基于大数据分析的网购客户忠诚度建模仿真[J]. 计算机仿真, 2015, 32(10): 239-242 + 304.
- [35] 党建武, 从筱卿. 基于CNN和GRU的混合股指预测模型研究[J/OL]. 计算机工程与应用; 1-11, <http://kns.cnki.net/kcms/detail/11.2127.TP.20201027.1415.016.html>, 2020-11-11.