

基于 BERT 模型的增强混合神经网络的谣言检测

梁兆君^{1 2} 但志平^{1 2*} 罗衍潮³ 李 奥^{1 2}

¹(三峡大学计算机与信息学院 湖北 宜昌 443002)

²(三峡大学水电工程智能视觉监测湖北省重点实验室 湖北 宜昌 443002)

³(西北大学信息科学与技术学院 陕西 西安 710127)

摘 要 网络社交平台中大量谣言的广泛传播严重影响社会稳定。传统谣言检测方法无法有效处理文本中多义词和突出重要关键词,造成检测效果不理想。针对该问题,提出一种基于 BERT 模型的增强混合神经网络的谣言检测方法。该方法使用 BERT 模型将推文向量化,通过 3 种不同尺寸的卷积核学习推文特征,将这些特征进行最大池化拼接得到特征序列,并输入到 BiLSTM 中学习序列特征。同时引入 Attention 机制计算注意力分值,实现谣言检测。在公开数据集 Twitter15 和 Twitter16 上的实验结果表明,该方法相较于其他方法在谣言检测性能方面提升明显,并在早期检测阶段表现出卓越的检测能力,检测准确率提高了 5 个百分点。

关键词 谣言检测 特征提取 卷积神经网络 双向长短期记忆网络 注意力机制

中图分类号 TP391 文献标志码 A DOI: 10.3969/j.issn.1000-386x.2021.03.022

RUMOR DETECTION OF IMPROVED HYBRID NEURAL NETWORK BASED ON BERT MODEL

Liang Zhaojun^{1 2} Dan Zhiping^{1 2*} Luo Yanchao³ Li Ao^{1 2}

¹(College of Computer and Information Technology, China Three Gorges University, Yichang 443002, Hubei, China)

²(Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, Hubei, China)

³(College of Information Science and Technology, Northwest University, Xi'an 710127, Shaanxi, China)

Abstract There are a large number of rumors existing in social networking platforms, while the widespread of rumors have affected social stability seriously. The traditional rumor detection approaches cannot solve the problem of polysemy in the text and highlight the significant keywords, which leads to unsatisfactory detection results. To solve the problems, a rumor detection method based on BERT model and improved hybrid neural network is proposed. In this method, BERT model was used to vectorize the tweets. The convolution kernel of three different sizes was used to learn the features of tweets, and these features were added to the max pooling layer and spliced into feature sequences, which were inputted into BiLSTM for learning sequence features. Moreover, the Attention mechanism was added to calculate attention value. The results on Twitter15 and Twitter16 datasets demonstrate that the proposed method achieves much better performance than state-of-the-art approaches, and demonstrates superior capacity on detecting rumors at very early stage. The accuracy increases by 5 percent.

Keywords Rumor detection Feature extraction CNN BiLSTM Attention mechanism

收稿日期: 2019-08-21。国家自然科学基金项目(61871258); NSFC-新疆联合基金项目(U1703261); 湖北省自然科学基金项目(2018CFC852)。梁兆君 硕士生, 主研领域: 自然语言处理。但志平 教授。罗衍潮 硕士生。李奥 硕士生。

0 引言

互联网时代网络社交平台发展迅速,网络信息数量庞大、资源共享、传播快,同时也滋生了大量虚假信息以及谣言。

网络谣言检测过程分为语言预训练模型和分类模型两个部分。目前,国内外网络谣言检测研究^[4]的语言预训练模型多采用 Word2vec^[5] 训练的词向量,相比简单的 One-hot 向量化方式,解决了数据稀疏的问题,但是 Word2vec 模型的缺点在于无法解决文本一词多义的问题。Devlin 等^[6] 提出一种新的语言表示模型 BERT,在训练某个单词词向量的过程中,根据上下文单词的语义动态地调整该单词的词向量表示,调整后的词向量能更好地表达在具体语境下的含义,可以解决文本中多义词的问题,成为目前自然语言处理领域最优的语言预训练模型。另一方面,分类模型方法对网络谣言检测影响非常大,目前网络谣言检测研究的分类检测方法主要包括传统机器学习方法和深度学习方法。其中,传统机器学习方法有朴素贝叶斯^[7] (Naive Bayesian Model, NBM)、决策树^[8]、支持向量机^[9] (Support Vector Machine, SVM)、随机森林^[10] 等。近年来,深度学习方法逐渐被用于自然语言处理领域, Kim^[11] 提出以 CNN 解决句子分类问题,首次将 CNN 应用到文本分类领域,实验表明其强大的特征学习和分类能力同样适用于谣言检测研究,但基础 CNN 模型的缺点在于单卷积层无法捕获远距离特征。Ma 等^[12] 提出利用循环神经网络(RNN)检测微博中的谣言,对于捕获远距离特征非常有效,在一定程度上提高了谣言检测准确率,但由于 RNN 网络的序列依赖结构使得其很难具备高效的并行运算能力,运行时间长,不利于网络谣言的早期检测。Zhou 等^[13] 提出一种基于 C-LSTM 的混合神经网络,结合 CNN 和 RNN 两大网络的优点,利用 CNN 提取短语的局部特征,利用 LSTM 提取句子的全局语义和时态语义,实验表明该模型在情感分类和问题分类上取得了很好的效果,但无法突出关键词对输出信息的影响。以上算法都无法体现关键词在文本中的重要程度,而对于谣言检测任务,由于不同的词在文本中的重要程度不同,因此需要计算权重,以达到将注意力^[14-15] 集中在重要词上的目的。

基于上述研究存在的问题,本文提出一种基于 BERT 模型^[16-17] 的增强混合神经网络的谣言检测方法(BERT + CNN-3-BiLSTM-Attention, BC3BLA)。该方

法不仅可以提取局部短语特征和全局上下文特征,还可以提取出文本关键词,减少无关信息对检测结果的不合理影响,实验表明融合多网络结构的模型可以有效提高谣言检测的准确率。同时,使用 BERT 模型预训练相比 Word2vec 模型的收敛时间短,有助于在网络谣言传播早期识别并抑制谣言传播。

1 网络谣言检测模型

本文提出的 BC3BLA 谣言检测方法在 Twitter 数据上进行实验,图 1 为 BC3BLA 模型结构。

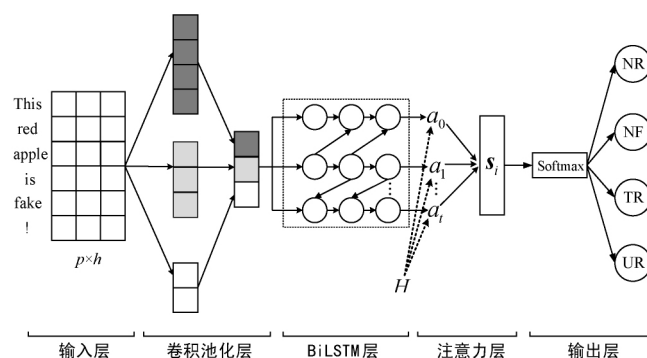


图1 BC3BLA 模型结构

该模型网络层次结构包含以下 5 部分:

- (1) 输入层: 使用 BERT 模型将推文映射为低维向量;
- (2) CNN-3 层: 使用 3 种不同尺寸的卷积核对输入矩阵进行卷积池化操作,并将推文特征横向拼接成特征序列;
- (3) BiLSTM 层: 将特征序列输入到 BiLSTM 层,利用双向 LSTM 从前后两个方向全面捕捉推文更深层次的语义特征;
- (4) 注意力层: 计算注意力分布概率,生成含有注意力概率分布的语义编码,优化特征向量;
- (5) 输出层: 将最终提取到的特征向量通过 Softmax 激活函数完成分类检测。

2 语言预训练模型

传统的语言预训练模型无法解决文本中多义词的问题。多义词在文本中经常出现,同一个词在不同的语境中的含义往往不同,如单词“bill”在一些语境中的含义是“账单”,在另一些语境中的含义是“钞票”。语义不同,词的向量表示应不同,Word2vec 模型对一个词语只有一个向量表示,而 BERT 会根据语料中的上下文语境,动态调整词语对应的词向量表示。

本文采用 BERT 模型对推文进行预训练,图 2 为 BERT 模型的结构图。该模型包含以下两个阶段:

- 1) 编码阶段,首先通过查询词典将推文中的每个词进行编码,转化成编码向量,由于文本中的词所在的位置不同表达的语义信息也有所差异,因此该模型针对不同位置的词分别附加一个不同的位置向量作为区分。
- 2) 生成向量阶段,将以上编码向量和位置向量送入 BERT 模型进行训练学习,输出融合了全文语义信息后的向量表示,并将其送入下一层 CNN-3 神经网络进行训练。

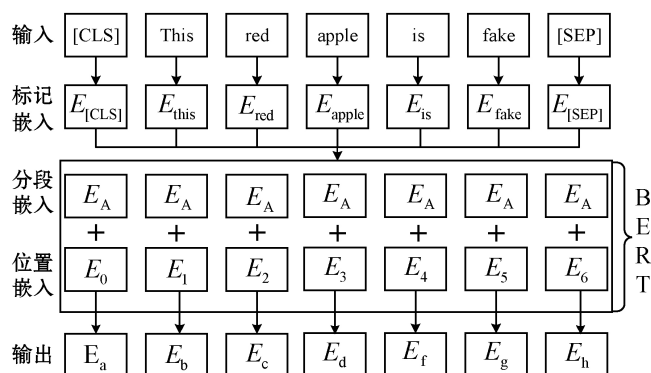


图2 BERT模型结构

3 分类模型

3.1 CNN-3 层

本文方法中 CNN-3 层使用了尺寸为 3、4、5 的卷积核和一个最大池化层进行特征提取。设置卷积核的宽度等于词向量的维度有助于更好地提取 Twitter 句子中单词的语义和语序信息。图 3 是 CNN-3 模型进行卷积、最大池化以及对生成的特征图列向量拼接的过程。

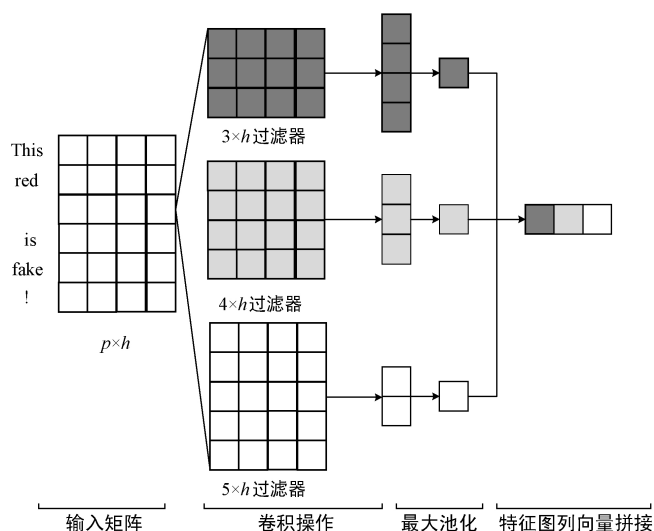


图3 CNN-3模型图

将每一条推文作为卷积层的一个输入矩阵,设 $t_i \in \mathbf{R}^h$ 表示一条推文中第 i 个词的 h 维向量,长度为 p 的推文表示为:

$$t_{1:p} = t_1 \oplus t_2 \oplus \cdots \oplus t_p \quad (1)$$

式中: \oplus 表示串联操作,令 $t_{i:i+j}$ 表示词 $t_i, t_{i+1}, \dots, t_{i+j}$ 的串联。设卷积核的长度为 l ,卷积核的宽度等于输入矩阵的宽度,则向量 $f \in \mathbf{R}^{l \times h}$ 表示卷积操作的卷积核。卷积核 f 对每条推文不同位置的单词窗口向量进行卷积操作,产生新的特征,例如某条推文中第 i 个单词 t_i 到第 $i+l-1$ 个单词 t_{i+l-1} 的窗口向量所产生的特征 m_i 的计算公式如下:

$$m_i = f(w \cdot t_{i:i+l-1} + b) \quad (2)$$

式中: $b \in \mathbf{R}$ 为偏置项; f 为非线性函数,如双曲正切函数。此卷积核应用于推文 $\{t_{1:l}, t_{2:l+1}, \dots, t_{p-l+1:p}\}$ 中每个可能的单词窗口,生成一个特征映射:

$$m = [m_1, m_2, \dots, m_{p-l+1}] \quad (3)$$

式中: $m \in \mathbf{R}^{p-l+1}$ 。然后池化层对此特征映射进行最大池化操作,将最大值 $\hat{m} \in \max\{m\}$ 作为该特定卷积核对应的特征,本文选择采用 max pooling 目的是为每个特征映射捕捉最重要的特征。最后将不同尺寸卷积核对应的特征进行横向拼接得到矩阵:

$$W = [\hat{m}; \hat{n}; \hat{e}] \quad (4)$$

式中: $\hat{m}, \hat{n}, \hat{e}$ 分别为三种不同尺寸卷积核经过卷积和最大池化后生成的新特征。

3.2 BiLSTM 层

本文方法的 BiLSTM 层使用了双向长短期记忆网络捕获双向的语义特征。单向 LSTM 网络将信息通过隐层状态从前向后传播,无法编码从后向前的信息,但推文的上文影响力和下文同样重要,因此使用双向长短期记忆网络能更好地捕获双向的语义特征。图 4 是 BiLSTM 网络的模型图,其输入 $w_0, w_1, w_2, \dots, w_v$ 是 CNN-3 层的输出特征向量 v 为卷积核的个数。

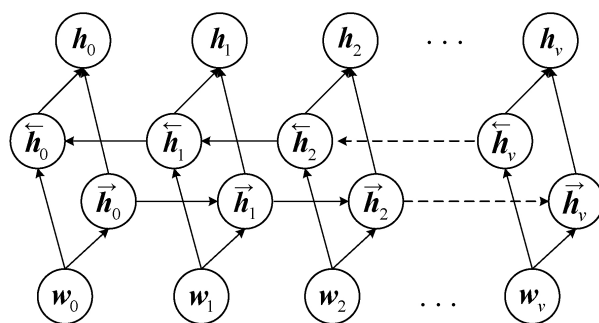


图4 BiLSTM模型图

BiLSTM 模型使用两个门来控制单元状态 c 的内容: 遗忘门 f_i 和输入门 i_i 。遗忘门决定了上一时刻的单元状态 c_{i-1} 有多少保存到当前时刻 c_i :

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (5)$$

输入门决定了当前时刻网络的输入 x_t 有多少保存到单元状态 c_t :

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (6)$$

然后根据上一次输出的隐层状态信息和本次的输入计算当前输入的单元状态 c'_t :

$$c'_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + W_{cc}c_{t-1} + b_c) \quad (7)$$

接下来计算当前时刻的单元状态 c_t 。用上一次的单元状态 c_{t-1} 按元素乘以遗忘门 f_t , 再用当前输入的单元状态 c'_t 按元素乘以输入门 i_t , 再将两个积加和, 这样新的单元状态 c_t 就由当前记忆 c'_t 和长期记忆 c_{t-1} 组合而成:

$$c_t = f_t \circ c_{t-1} + i_t \circ c'_t \quad (8)$$

BiLSTM 的输出门 o_t 控制了长期记忆对当前输出的影响:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (9)$$

输出 h_t 由输出门和单元状态共同决定:

$$h_t = o_t \circ \tanh(c_t) \quad (10)$$

BiLSTM 在原有的正向 LSTM 网络层上增加一层反向的 LSTM 层, 所以第 i 个单词向量的输出可以表示为:

$$h_i = [\vec{h}_i \oplus \overleftarrow{h}_i] \quad (11)$$

式(5) - 式(11)中: W_{xf} 、 W_{hf} 、 W_{cf} 、 b_f 表示遗忘门 f_t 相应的权重矩阵; W_{xi} 、 W_{hi} 、 W_{ci} 、 b_i 表示输入门 i_t 相应的权重矩阵; W_{xc} 、 W_{hc} 、 W_{cc} 、 b_c 表示输出门 o_t 相应的权重矩阵。

3.3 注意力层

本文引入注意力机制, 解决了 BiLSTM 层的输出信息无法突出重要关键词的问题。单词在 Twitter 文本中所占的权重越大, 该词受到的注意力值越大, 说明该词在谣言检测任务中发挥的作用越大。注意力机制模型图如图 5 所示, 其中输入的 $h_0, h_1, h_2, \dots, h_v$ 是 BiLSTM 层的输出特征向量。

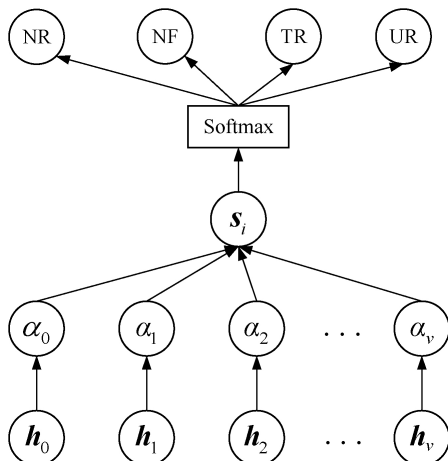


图5 注意力模型图

首先计算 BiLSTM 不同时刻隐层的输出 $h_0, h_1, h_2, \dots, h_v$ 在整个文本中对应分配的注意力概率分布值 $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_v$, 该值越大, 在 Twitter 文本中的注意力越大, 然后用 Softmax^[18] 进行归一化, 最后得到的聚焦化的向量 s_i 是所有向量的加权和, 计算式为:

$$u_i = w^T \tanh(W h_i + b_h) \quad (12)$$

$$a_i = \frac{\exp(u_i \sigma_w)}{\sum_j \exp(u_j \sigma_w)} \quad (13)$$

$$s_i = \sum_j \alpha_j h_j \quad (14)$$

式中: h_i 表示在第 i 时刻隐藏层的输出状态包括前馈输出 \vec{h}_i 和反馈输出 \overleftarrow{h}_i ; u_i 是 h_i 的隐层表示; a_i 表示注意力分配系数; σ_w 表示上下文信息的参数矩阵(随机初始化), 之后会作为模型的参数一起被训练; W 是权重转化矩阵; b_h 为偏置量; \tanh 为非线性激活函数。

3.4 输出层

本文方法的最后部分是输出层, 首先使用全连接网络, 计算式如下:

$$y_i = w_z s_i + b_z \quad (15)$$

式中: w_z 为权重矩阵; b_z 为偏置项; s_i 是经过注意力层得到的特征向量。然后使用 Softmax 激活函数输出分类的概率, 计算公式为:

$$p_j = \frac{\exp(y_j)}{\sum_{j=1}^n \exp(y_n)} \quad (16)$$

4 实验

4.1 实验数据集

为了验证本文方法的有效性, 实验采用 Ma 等^[19] 2017 年公开的用于谣言检测研究的数据集 Twitter15 和 Twitter16, 它们均来自谣言揭穿网站(例如 snopes.com, Emergent.info 等)。两个数据集中每个事件的标签被分为四类: 非谣言类(non-rumors, NR)、假谣言类(false rumors, FR)、真谣言类(true rumors, TR)和未经证实谣言类(unverified rumors, UR)。为了便于本文方法与文献[20]方法进行比较, 本文采用与之相同的两个数据集和样本分类进行实验。表 1 为数据集 Twitter15 和 Twitter16 的统计信息。

表 1 实验数据集统计表

数据集统计信息	Twitter15	Twitter16
用户数	276 663	173 487
源推文数	1 490	818

续表 1

数据集统计信息	Twitter15	Twitter16
推文串数量	331 612	204 820
NR	374	205
FR	370	205
TR	372	205
UR	374	203

4.2 实验设置

将本文方法与文献 [20] 方法进行比较:

1) 文献 [20] 方法: 基于树结构递归神经网络的 Twitter 谣言检测, 在 Twitter15、Twitter16 数据集上实现了自底向上的树结构递归神经网络(Bottom-up Recursive Neural Network ,BU-RvNN) 和自顶向下的树结构递归神经网络(Top-down Recursive Neural Network ,TD-RvNN) 等模型, 并取得了良好的效果。

2) 本文方法: 基于 BERT 模型的增强混合神经网络的谣言检测方法(BC3BLA) ,并提出以下 3 个方法用于对比:

(1) Word2vec + CNN-3-BiLSTM(WC3BL): 使用 Word2vec 训练词向量, 利用 CNN-3 层提取高级短语特征, 利用双向 LSTM 捕捉两个维度的推文语义特征。

(2) Word2vec + CNN-3-BiLSTM-Attention(WC3BLA): 在 WC3BL 的基础上, 引入注意力机制, 计算输出信息的注意力分值。

(3) BERT + CNN-3-BiLSTM(BC3BL): 采用与 WC3BL 相同的分类模型, 在语言预处理阶段采用 BERT 模型训练词向量。

BC3BLA 方法对每一条推文进行预处理后, 送入 BERT 模型训练, 将单词转化为词向量。BERT 的向量维度为 768; Word2vec 的维度设置为 256; CNN-3 层卷积核的窗口高度分别采用 3、4、5, 每个尺寸的卷积核数目设置为 100 个; BiLSTM 层隐藏层大小为 100; Dropout 设置为 0.3, 学习率为 0.001, 优化函数采用 Adam, 训练轮数为 50 轮, 损失函数采用多分类交叉熵。为了便于与文献 [20] 方法进行比较, 本文同样对数据集进行五折交叉验证, 对四个类别分别使用准确率和 F1 值作为评价指标, 以评估模型性能。

4.3 实验结果分析

表 2、表 3 为各方法实验结果对比。可以看出, 本文提出的 BC3BLA 方法, 在两个数据集上的表现均优于其他方法。

表 2 Twitter15 数据集上谣言检测结果

模型	准确率	F ₁			
		NR	FR	TR	UR
BU-RvNN	0.708	0.695	0.728	0.759	0.653
TD-RvNN	0.723	0.682	0.758	0.821	0.654
WC3BL	0.698	0.736	0.606	0.829	0.575
WC3BLA	0.714	0.648	0.657	0.832	0.704
BC3BL	0.742	0.734	0.632	0.859	0.768
BC3BLA	0.775	0.697	0.761	0.878	0.793

表 3 Twitter16 数据集上谣言检测结果

模型	准确率	F ₁			
		NR	FR	TR	UR
BU-RvNN	0.718	0.723	0.712	0.779	0.659
TD-RvNN	0.737	0.662	0.743	0.835	0.708
WC3BL	0.725	0.675	0.659	0.859	0.731
WC3BLA	0.732	0.724	0.636	0.831	0.779
BC3BL	0.754	0.733	0.656	0.819	0.811
BC3BLA	0.786	0.768	0.751	0.835	0.811

BC3BLA 与文献 [20] 方法中表现最好的 TD-RvNN 相比, 在 Twitter15 数据集上的准确率提高了 5.2 个百分点, 类别 NR、FR、TR 和 UR 的 F₁ 值分别提高了 1.5、0.3、5.7 和 13.9 个百分点; 在 Twitter16 上准确率提高了 4.9 个百分点, TR 的 F₁ 值与 TD-RvNN 模型相同, NR、FR 和 UR 的 F₁ 值分别提高了 10.6 个百分点、0.8 个百分点和 10.3 个百分点。

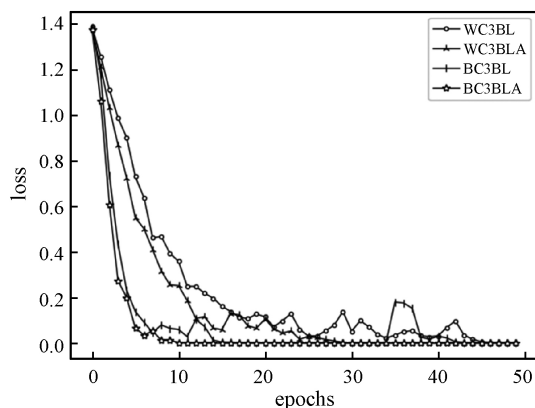
BC3BLA 与 WC3BLA 相比, 在 Twitter15、Twitter16 两个数据集上准确率分别提高了 6.1 个百分点和 5.4 个百分点, NR、FR、TR 和 UR 的 F₁ 值均有相应的提高。由实验结果分析可知, 采用 BERT 进行预训练后, 模型准确率得到了大幅度提升, 证明使用 BERT 模型训练的词向量的表现优于 Word2vec 模型。这是因为在谣言检测任务中, 无论词语在什么语境下具有什么含义, 采用 Word2vec 模型训练后都只有一个静态的词向量表示, 而 BERT 模型则会根据语料中上下文语境的不同, 动态地改变词语对应的词向量。通过上述实验, 发现在谣言检测任务中, BERT 模型优于 Word2vec 模型。

BC3BLA 与 BC3BL 相比, 在数据集 Twitter15 上准确率提高了 3.3 个百分点, 类别 FR、TR 和 UR 的 F₁ 值分别提高了 12.9、1.9 和 2.5 个百分点; 在 Twitter16 上准确率提高了 3.2 个百分点, NR、FR 和 TR 的 F₁ 值分别提高了 3.5、9.5 和 1.6 个百分点。实验结果分析可知, 引入注意力机制后, 四个类别的 F₁ 值均有较大提升。这是因为通过计算输出数据的注意力概率分布

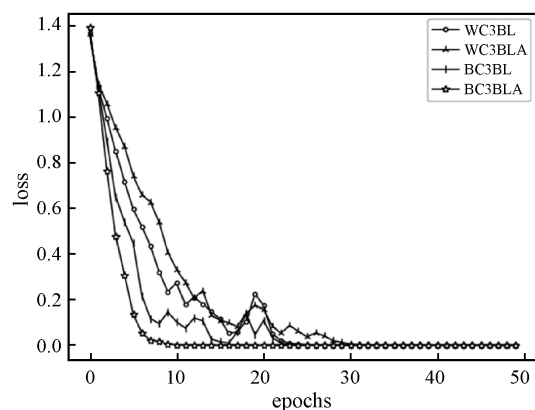
值,可以提取出对谣言检测任务影响更大的关键词,从而优化了特征提取过程。实验表明注意力机制消除了输入数据对输出数据的不合理影响,有效提高了谣言检测的准确率。

BC3BLA 与 WC3BL 相比,在 Twitter15、Twitter16 两个数据集上准确率分别提高了 7.7 和 6.1 个百分点, Twitter15 上 UR 的 F_1 值提升了 21.8 个百分点。实验数据分析可知,使用 BERT 模型预训练并在分类模型中引入 Attention 机制后,实验效果大幅度提升,表明了 BERT 与 Attention 机制结合的优越性。

图 6 所示为本文和三种对比方法分别在两个数据集训练过程中的收敛情况。可以看出,在利用批量梯度下降法进行训练过程中,函数损失值逐渐下降,最终趋于稳定收敛状态,相较于其他三种方法收敛时间最短、收敛速度最快,本文提出的 BC3BLA 方法有利于谣言的早期检测。



(a) Twitter15



(b) Twitter16

图 6 模型训练损失函数变化图

综上所述,本文提出的基于 BERT 模型的增强混合神经网络的谣言检测方法解决了文本多义词的向量表示问题,且综合考虑到推文的词语局部特征、全局上下文语义特征以及关键词对谣言检测结果的影响程度,模型结构更加全面,能高效准确地检测网络中的谣言。

5 结 语

本文提出了一种基于 BERT 模型的增强混合神经网络的谣言检测方法。该方法通过使用 BERT 模型训练词向量,结合 CNN-3、BiLSTM 和注意力机制对谣言进行识别检测,充分发挥各网络模型的优势。实验结果表明,使用 BERT 模型训练的词向量解决了自然语言处理领域多义词的识别问题,从而大幅度提高了谣言检测准确率;引入的注意力机制可以捕获输入数据对输出数据的影响力度,减少不合理影响以及特征提取过程中的信息丢失问题,丰富了关键词对文本的语义表达。目前网络谣言不仅限于文本信息,还包含大量音频、图片和视频等多媒体信息,如何将多媒体信息应用到谣言检测中是今后研究的重点。

参 考 文 献

- [1] Kwon S, Cha M, Jung K. Rumor detection over varying time windows[J]. PLoS ONE 2017, 12(1): e0168344.
- [2] Zhao Z, Resnick P, Mei Q Z. Enquiring minds: Early detection of rumors in social media from enquiry posts[C]// Proceedings of the 24th International Conference on World Wide Web. ACM 2015: 1395 - 1405.
- [3] Ma J, Gao W, Wong K F. Detect rumors on Twitter by promoting information campaigns with generative adversarial learning[C]//2019 International Conference on World Wide Web. ACM 2019: 3049 - 3055.
- [4] 程亮,邱云飞,孙鲁. 微博谣言检测方法研究[J]. 计算机应用与软件 2013, 30(2): 226 - 228.
- [5] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. ACM 2013: 3111 - 3119.
- [6] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[EB]. arXiv: 1810.04805 2018.
- [7] Qazvinian V, Rosengren E, Radev D R, et al. Rumor has it: Identifying misinformation in microblogs[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. ACM 2011: 1589 - 1599.
- [8] Castillo C, Mendoza M, Poblete B. Information credibility on twitter[C]//Proceedings of the 20th International Conference on World Wide Web. ACM 2011: 675 - 684.
- [9] Yang F, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo[C]//Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. ACM 2012: 13.

(下转第 189 页)

- graphs using physiologically plausible features [J]. *Investigative Ophthalmology & Visual Science* ,2007 ,48(4) : 1665 – 1673.
- [10] Cheng J , Liu J , Xu Y , et al. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening [J]. *IEEE Transactions on Medical Imaging* ,2013 ,32(6) : 1019 – 1032.
- [11] Ronneberger O , Fischer P , Brox T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention ,2015: 234 – 241.
- [12] Shelhamer E , Long J , Darrell T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence* ,2017 ,39(4) : 640 – 651.
- [13] Yu S , Xiao D , Frost S , et al. Robust optic disc and cup segmentation with deep learning for glaucoma detection [J]. *Computerized Medical Imaging and Graphics* ,2019 ,74: 61 – 71.
- [14] Abraham N , Khan N M. A novel focal tversky loss function with improved attention U-Net for lesion segmentation [C]//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) ,2019: 683 – 687.
- [15] Liu J , Wong D W K , Lim J H , et al. Optic cup and disk extraction from retinal fundus images for determination of cup-to-disc ratio [C]//2008 3rd IEEE Conference on Industrial Electronics and Applications ,2008: 1828 – 1832.
- [16] Xu Y , Liu J , Lin S , et al. Efficient optic cup detection from intra-image learning with retinal structure priors [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention ,2012: 58 – 65.
- [17] Alom M Z , Hasan M , Yakopcic C , et al. Recurrent residual convolutional neural network based on u-net (R2U-net) for medical image segmentation [EB]. arXiv: 1802.06955 ,2018.
- [18] Oktay O , Schlemper J , Folgoc L L , et al. Attention U-Net: Learning where to look for the pancreas [EB]. arXiv: 1804.03999 ,2018.
- [19] Jin Q , Meng Z , Sun C , et al. RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans [EB]. arXiv: 1811.01328 ,2018.
- [20] Lin T Y , Goyal P , Girshick R , et al. Focal loss for dense object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence* ,2020 ,42(2) : 318 – 327.
- [21] Gu Z , Cheng J , Fu H , et al. CE-Net: Context encoder network for 2D medical image segmentation [J]. *IEEE Transactions on Medical Imaging* ,2019 ,38(10) : 2281 – 2292.
- [22] Liu Y , Cheng M M , Hu X , et al. Richer convolutional features for edge detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) ,2017.
- [23] Szeliski R. *Computer vision: algorithms and applications* [M]. Springer Science & Business Media ,2010.
- [24] Orlando J I , Fu H , Breda J B , et al. REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs [J]. *Medical Image Analysis* ,2020 ,59: 101570.

(上接第152页)

- [10] Kwon S , Cha M , Jung K , et al. Prominent features of rumor propagation in online social media [C]//2013 IEEE 13th International Conference on Data Mining. IEEE ,2013: 1103 – 1108.
- [11] Kim Y. Convolutional neural networks for sentence classification [EB]. arXiv: 1408.5882 ,2014.
- [12] Ma J , Gao W , Mitra P , et al. Detecting rumors from microblogs with recurrent neural networks [C]//Proceedings of the 25th International Joint Conference on Artificial Intelligence. AAAI ,2016: 3818 – 3824.
- [13] Zhou C , Sun C , Liu Z , et al. A C-LSTM neural network for text classification [EB]. arXiv: 1511.08630 ,2015.
- [14] Yang Z , Yang D , Dyer C , et al. Hierarchical attention networks for document classification [C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies ,2016: 1480 – 1489.
- [15] Chen T , Li X , Yin H , et al. Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection [C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer ,2018: 40 – 52.
- [16] Wu X , Lü S , Zang L , et al. Conditional BERT Contextual Augmentation [C]//International Conference on Computational Science. Springer ,2019: 84 – 95.
- [17] Dai Z , Yang Z , Yang Y , et al. Transformer-xl: Attentive language models beyond a fixed-length context [EB]. arXiv: 1901.02860 ,2019.
- [18] Liu W Y , Wen Y D , Yu Z D , et al. Large-margin softmax loss for convolutional neural networks [C]//Proceedings of the 33rd International Conference on International Conference on Machine Learning. ACM ,2016: 507 – 516.
- [19] Ma J , Gao W , Wong K F. Detect rumors in microblog posts using propagation structure via kernel learning [C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics ,2017: 708 – 717.
- [20] Ma J , Gao W , Wong K F. Rumor detection on twitter with tree-structured recursive neural networks [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics ,2018: 1980 – 1989.