

基于用户特征分析的微博谣言早期检测研究^{*}

尹鹏博 潘伟民 彭 成 张海军

(新疆师范大学计算机科学技术学院 乌鲁木齐 830054)

摘 要: [目的/意义] 开放的微博平台是谣言的重要发源地,研究谣言检测对于减少谣言的传播,降低或消除谣言的影响具有重要现实意义。[方法/过程] 提出了一种基于用户特征分析的微博谣言早期检测方法,通过对用户历史微博进行情感分析得到用户的发文行为特征,结合用户属性和微博文本,使用卷积-长短期记忆网络(Convolution-Long Short-Term Memory, C-LSTM)模型实现谣言早期检测。[结果/结论] 与目前大多数微博谣言检测方法不同的是,该方法基于微博用户的历史行为特征,不依赖待检测微博的评论转发信息,减少了等待产生评论转发信息的时间,实现了微博谣言的实时早期检测。

关键词: 微博谣言; 谣言检测; 行为特征; 深度学习; 卷积-长短期记忆网络

中图分类号: TP391.1

文献标识码: A

文章编号: 1002-1965(2020)07-0081-06

引用格式: 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.

DOI: 10.3969/j.issn.1002-1965.2020.07.014

Research on Early Detection of Weibo Rumors Based on User Characteristics Analysis

Yin Pengbo Pan Weimin Peng Cheng Zhang Haijun

(School of Computer Science and Technology, Xinjiang Normal University, Urumqi 830054)

Abstract: [Purpose/Significance] The open-ended Weibo is an important source of rumors. The study of rumor detection is of great practical significance to reduce or eliminate the spread of rumors. [Method/Process] This paper proposes an early detection method of Weibo rumors based on user characteristics analysis. Through the emotional analysis of the user's historical Weibo, we can get the user's behavior characteristics of posting. Combined with the user attributes and Weibo text, we use Convolution-Long Short-Term Memory (C-LSTM) deep learning model to realize early detection of rumors. [Result/Conclusion] Different from most of the current Weibo rumors detection methods, this method is based on the historical behavior characteristics of Weibo users, does not rely on the comment and forwarding information of the Weibo to be detected, reduces the waiting time for the generated comment and forwarding information, and realizes the real-time early detection of Weibo rumors.

Key words: Weibo rumors; rumor detection; behavior characteristics; deep learning; Convolution-Long Short-Term Memory

0 引 言

随着互联网技术的发展,网络社交平台如雨后春笋一般不断涌现,新浪微博就是典型代表之一。微博在给人们的日常生活带来便利的同时也促进了谣言的

产生和传播,一方面是由于用户可以在没有监督和审核的情况下发布任何信息,另一方面是由于接收信息的用户通常缺乏相关能力来鉴别他们所接收信息的真伪^[1]。微博官方辟谣平台使用人工审核方式鉴别谣言,效率低下。因此,微博谣言自动检测研究逐渐受到

收稿日期:2020-03-19

修回日期:2020-04-21

基金项目:国家自然科学基金-新疆联合基金项目“网络谣言检测与舆论引导算法研究”(编号:U1703261)研究成果之一。

作者简介:尹鹏博(ORCID:0000-0003-1707-9483),男,1996年生,硕士研究生,研究方向:自然语言处理、数据挖掘(网络舆情方向);潘伟民(ORCID:0000-0002-7441-2791),男,1963年生,硕士,教授,院长,研究方向:计算机应用技术、网络信息安全;彭 成(ORCID:0000-0003-0619-2771),男,1971年生,博士,副教授,副院长,研究方向:教育信息化、娱教技术;张海军(ORCID:0000-0002-6823-7077),男,1973年生,博士,教授,硕士生导师,研究方向:自然语言处理、信息抽取、人工智能。

通信作者:潘伟民

广泛关注。

谣言在传播过程中,会有用户不断地提出对谣言信息的质疑与反对的意见^[2],这是众多谣言检测方法的检测依据。由于谣言需要经过一定范围的传播才能产生足够的评论信息,因此这种检测方法被称为“滞后检测”,与之对应的是在谣言产生之初就进行检测的“早期检测”。早期检测具有更大的实践意义,但在谣言产生的早期,没有太多的用户参与讨论,故无法通过其他用户的意见进行早期检测。通过对大量微博用户的历史数据进行分析,我们发现发布谣言的用户往往具有很多共同特征,例如他们的历史微博通常带有争议性观点等,因此可以通过对用户的历史微博进行分析,获取用户行为特征,进而实现谣言的早期检测。

因此,本文提出基于用户特征分析的微博谣言早期检测方法,通过对用户历史行为特征分析,结合用户属性和微博文本,使用 C-LSTM 深度学习模型融合文本特征和用户行为特征实现微博谣言的早期检测。由于本文采用的方法不使用评论转发信息进行检测,能在谣言产生之初将其检测识别,因此比现有的检测方法都要“早”。

1 相关工作

目前,微博谣言检测研究按检测时效性可以分为早期检测和滞后检测,研究的主要内容是高效检测算法的应用和深层次检测特征的挖掘。

在检测算法应用方面,主要由早期的机器学习算法,如 SVM、决策树等向深度学习算法过渡,Yu 等^[3]等提出使用卷积神经网络(Convolutional Neural Networks, CNN)进行谣言检测;Jin 等^[4]提出使用循环神经网络(Recurrent Neural Network, RNN)进行谣言检测;Chen 等^[5]提出引入注意力机制的 RNN 检测模型。从算法的效果来看,深度学习算法比传统机器学习算法更好,而复合深度学习算法模型又比单一深度学习算法模型效果好。

在检测特征挖掘方面,主要是在微博文本及其评论信息的基础上引入少量的其他特征。Yang 等人^[6]引入了发布微博的设备类型和地理位置信息;Wu 等人^[7]引入了微博的主题特征;Morris 等人^[8]引入了用户画像特征;还有研究引入了谣言的传播特征,如 Nourbakhsh 等^[9]研究了不同用户对谣言传播的影响,Vosoughi 等人^[10]通过谣言信息传播过程中的语言风格、参与者特征以及传播动态来进行谣言检测。这些特征的引入在一定程度上提高了检测的效果,但还是无法脱离对微博的评论转发数据的依赖。

相关研究中的早期检测和滞后检测的区别不在于检测算法和检测特征的根本性差异,而是相对于滞后

检测使用数据集中全部的评论数据进行检测的情况,早期检测会设置检测窗口,根据窗口大小,只选取谣言发生后一段时间内的评论数据进行检测,使其能实现在谣言产生后较短一段时间内将其检测出来。不过更少的评论信息并不意味着不需要评论信息,而评论信息的产生是以谣言的传播以及其他用户的参与为代价的,这就决定了基于谣言微博的评论进行检测的方法不能在谣言产生之初就将其识别出来。

通过对已有研究方法的分析发现:要实现真正的谣言早期检测,就要摆脱对微博评论信息的依赖;基于神经网络的检测算法相较于传统机器学习算法具备更好的效果,但其中的 CNN 和 RNN 又有不同:CNN 在处理短文本上效果比 RNN 好,而 RNN 在处理长文本上又比 CNN 更有效。

2 微博用户特征分析

本文所设计的谣言模型特征由微博文本分析和微博用户行为分析两部分组成。对于微博谣言信息的检测,微博博文内容在一定程度上可以表征该微博是否为谣言^[11],同时在微博博文内容分析的基础上引入用户特征分析,大大提高了检测效果。

2.1 用户行为分析 通过对大量谣言信息和发布谣言信息的博主进行分析,发现这些博主和谣言微博信息拥有诸多共性,例如:大多数谣言博主发布过不止一条谣言信息;大多数谣言博主的日常微博观点颇具争议性;谣言微博信息具有相较于普通微博更多的评论、转发;大多数的谣言微博评论中会有更多的意见不一致、争执的情况;发布谣言的博主中,有很大一部分为自媒体人、微商等需要依靠粉丝进行盈利的微博用户,很多博主会在个人说明中介绍自己的业务并且留下商业合作的联系方式,具体如表 1 所示。

表 1 谣言用户的用户名、自我介绍与发布谣言微博数

用户名	谣言数	个人说明
草根原创精选	3	精选情感语录,指点情感迷津。商务合作请加 QQ1803065009
史上第一最搞笑	3	个人微博,合作 QQ 3544258637
靠谱影评	2	关注我,节省你找寻经典的时间。商务合作加 Q:3325415544
全球热门排行榜	6	合作请加 qq251335674。有态度、有热情、有意思的热门榜单
五行犯二	14	商务合作 QQ297906068
上海热门资讯	7	关注上海,乐享资讯。合作联系+Q:2209273419
韩国 me2day	42	联系邮箱: me2day@ qq.com 与韩国媒体同步的全方位韩流报道。

谣言信息的观点通常具有较大的争议性,并且常常拥有激发大众情绪的能力,能抓住大众的心理进行广泛的传播,进而获得较大的社会关注度^[11]。例如抓

住大众对食品安全的担忧,大肆炒作转基因食品食用后会导致基因突变等问题的谣言。

很多微博用户发布谣言的动机是希望通过谣言信息的广泛传播为自己带来大量的粉丝,进而通过粉丝获取利益,譬如自媒体人可以通过粉丝转发微博扩大影响力,还可以依靠粉丝量获取广告赞助费等;而有些用户发布谣言则是为了宣泄;还有用户是为了猎奇。总的来说,这些用户的日常微博更具有吸引关注的能力,因此可通过结合此特点进行谣言早期检测。下面先对引战言论进行定义:

引战言论:具有较大争议性的微博。微博的争议性计算见公式(1)。

2.2 引战言论的计算 谣言博主的日常言论能吸引较多的用户参与其中,评论中较大的争论说明博文观点具有引战性,较多的转发意味着博文的影响不仅仅局限于该博主的粉丝群体之中,还通过粉丝进行了扩大。基于此特点,在对用户历史微博进行特征提取时采用情感倾向分析法提取历史微博的特征。在这里我们对历史微博的评论信息进行情感倾向分析以确定它的引战性。步骤如下所示:

(1)定义 pos 为正向情感评论数, neg 为负向情感评论数,且均不为 0。

(2)定义 C 为情感激烈程度,当一种情感数比另一种多出一数量时为情感激烈,否则不激烈,因此可

ID: 大鹏看天下
location: 北京朝阳区
verify: 媒体人
description: 淘宝
店——良品海购,代购
海外精品,平价路线
status: 30521
followers: 454300
follow: : 1065
rumors: 2 strips: 2

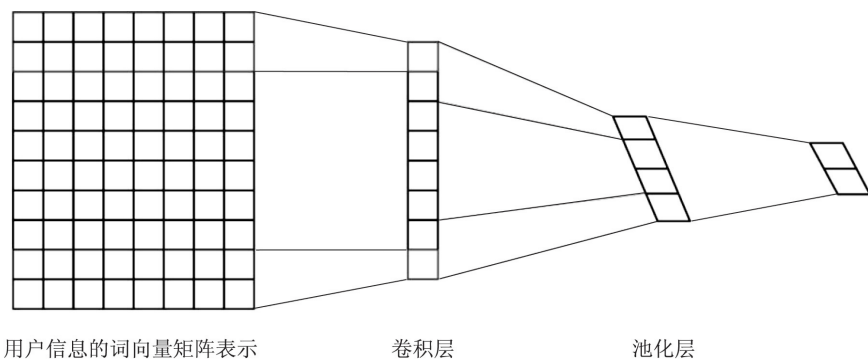


图1 文本处理中的一维卷积过程

对于微博文本特征提取任务,使用 RNN 可挖掘其中深层的隐含关系以及语义信息^[12],但是 RNN 存在梯度爆炸的问题,长短记忆网络(Long Short-Term Memory, LSTM)则通过设置输入门、输出门和遗忘门解决了梯度消失和梯度爆炸的问题,进而在文本语义挖掘中的表现较好,因此本文选取 LSTM 进行微博文本深层语义信息挖掘的算法。

综合以上研究分析,本文采用 CNN 和 LSTM 相结合的方式对谣言检测模型的构建。使用 CNN 提取用户特征的深层次表示,使用 LSTM 提取微博文本的深层隐含信息,最后将两个神经网络的特征进行融合。图 2 所示是本文采用的神经网络模型图。

以将其量化为: $C = \frac{neg}{pos} + \frac{pos}{neg} = \frac{pos^2 + neg^2}{pos * neg}$

(3)定义归一化函数 f 。当 pos 和 neg 是小数和是大数时,情感倾向的不同无法由 C 表达,因此需要为 C 加上 pos/neg 在数量上的权重。令 $C = \frac{pos^2 + neg^2}{pos * neg} * f(|neg + pos - a|)$, a 是偏置系数。

(4)由于 $C \in [2, +\infty)$,不便量化,因此将 C 用归一化函数 f 处理得到公式 1:

$$C = f\left(\frac{pos^2 + neg^2}{pos * neg} * f(|neg + pos - a|)\right) \quad (1)$$

通过公式(1)计算微博评论的情感激烈程度。当 $C=0$ 时,评论中无争论,当 $C=1$ 时,评论争论最为激烈。当 C 大于一定阈值时,判定该条微博为引战微博,引战言论数将是用户的一个重要特征。由于 C 阈值与样本数据有较高关联性,因此通过实验确定较为合理。

3 基于 C-LSTM 的检测模型

3.1 检测算法的选取 在用户特征提取方面,由于用户特征信息并不是由一段完整文本构成的信息,而是由不同特征描述构成的不连贯的文本,卷积神经网络具有关注局部语义信息的特点,因此 CNN 更能有效的提取用户特征(如图 1 所示),展现了 CNN 对文本的特征提取过程。

3.2 使用 CNN 提取用户特征 本模型对输入的用户信息进行一维卷积,设卷积核 W 为 $width * height$ 大小的矩阵。不同于图像处理中的二维卷积,一维卷积的卷积核只需要设置它的高度 $height$,而它的宽度 $width$ 则设置为输入数据矩阵的宽度。这是由于在图像处理中需要在连续的像素值之间提取局部特征,故需要有不同的卷积核的宽和高(决定特征提取的粒度),而在用户特征的提取中,需要整体特征而不是词向量之间的关系,故将 $width$ 指定为输入矩阵的宽。通过一维卷积产生的用户特征用公式(2)表达:

$$a_{user} = ReLU(W * X_{user} + b) \quad (2)$$

其中, $ReLU$ 为激活函数, W 为卷积核, X_{user} 为输

入的用户信息, b 为偏置项。

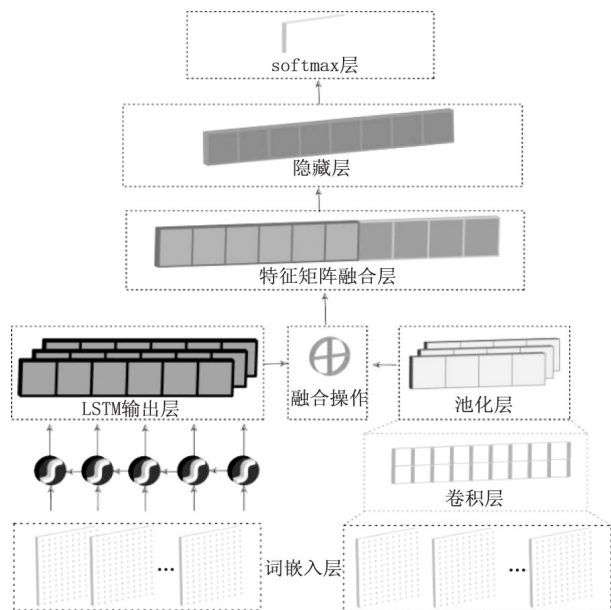


图2 C-LSTM融合模型图

3.3 使用 LSTM 提取微博文本特征 通过 LSTM 循环神经网络提取微博文本特征用公式(3)表示:

$$a_{weibo} = \text{sigmoid}(\text{last}(U^T * x^t + O^T * a^{t-1} + b)) \quad (3)$$

其中, sigmoid 是激活函数, last 指取序列的最后一个输出作为序列的结果, U 和 O 为输入门和遗忘门权值矩阵, b 为偏置项。

3.4 特征融合 通过 CNN 和 LSTM 分别对用户信息和微博文本进行特征提取之后得到特征矩阵 a_{user} 和 a_{weibo} , 最后一步通过将 a_{user} 和 a_{weibo} 融合产生最终的特征矩阵送入神经网络的全连接层进行学习。 a_{user} 和 a_{weibo} 为二维矩阵, 分别用形如 $1D(a)$ 和 $2D(a)$ 来表示特征矩阵的两个维度, 因为 $1D$ 维度代表着不同用户或者不同微博文本, 而 $2D$ 维度代表着同一用户的各种不同特征集合或者同一微博文本的不同词向量集合, 特征向量的拼接需要将用户特征与其发表的微博相对应, 为保证将用户所属的微博与其个人信息融合在一起, 就需要沿着 $2D$ 维度将两个向量集合拼接起来形成融合向量集合, 但这里的拼接并不是指向量乘积运算, 而是将两个向量集合叠加在一起, 用符号 \oplus 表示。将融合之后的特征矩阵 a 输入全连接层, 最后使用 softmax 分类器进行谣言检测。融合过程可由下式(4)表达:

$$a = \sum 2D(a_{user}) \oplus 2D(a_{weibo}) \quad (4)$$

4 实验

4.1 数据集的选择 本文选取 Ma 等^[13] 公开的微博谣言数据集。在对数据集进行整理后发现数据集中

有 164 名用户已经被封停账号, 他们的数据已经无益于后续研究, 因此将封停账号涉及到的数据全部删除, 整理后数据集共包含 2 145 名用户, 4 346 条微博, 其中谣言数据 2 313 条, 非谣言数据 2 033 条。由于该数据集不包含用户的历史微博信息, 因此需要另行爬取此数据集所涉及到的用户的历史微博及评论, 进行用户分析。本文设计了微博爬虫在微博网站上爬取这 2 145 名用户的个人信息和历史微博, 这些用户平均拥有 70 条历史微博, 因此对于历史微博数超过 70 条的只爬取最新 70 条, 不足 70 条的全部爬取, 同时还要爬取历史微博的评论、点赞、转发等信息, 共爬取这些用户的 135 991 条历史微博和各项最新个人信息作为用户特征。

4.2 数据预处理 微博的数据中包含很多话题符号、表情符号等无关信息, 需要进行数据清洗将它们去除, 这里采用正则表达式进行数据清洗。然后对每名用户的历史微博进行情感分析, 情感分析采用 SnowNlp 工具, 它方便使用且具备二次训练的能力。由于 SnowNlp 自带的情感分析并非训练自微博语料, 应用到微博情感分析会有偏差, 因此本文选择了部分热点事件的 400 万条微博评论数据对 SnowNlp 进行二次训练, 训练采取半监督的方式进行, 先使用 SnowNlp 的情感分析功能对训练数据进行情感分析, 将情感很积极或者很消极的数据保存为训练数据。最后将训练数据用于新的情感分析模型的训练。

在通过二次训练得到新的情感分析模型后, 使用其对用户历史微博的评论进行情感倾向分析, 根据前文所述, 将情感分析的结果通过公式(1)计算, 结合 C 的阈值可得该微博是否为引战微博, 将结果计入用户属性之中, 得到引战微博数。

4.3 检测特征选取 根据相关研究, 选取微博文本、发布设备为微博特征; 选取关注数、粉丝数、微博数、性别、微博认证、用户昵称、用户自我描述、引战言论数作为用户特征; 其中引战言论数需要对用户的历史微博进行情感分析得到, 除此之外的信息都可以直接使用爬虫技术从微博平台获取。

4.4 阈值选择实验 C 的最优阈值可以通过阈值选择实验确定, 因此在计算 C 时, 分别令阈值为 0.30, 0.35...0.70, 计算出该阈值所对应的数据, 然后分别在这 9 份数据之上进行实验, 取 F1-score 指标最好的实验数据所对应的阈值作为最佳阈值。

实验使用不同阈值所计算出的 9 组数据分别训练并测试 C-LSTM 模型, 9 组数据的实验结果如表 2 所示。

通过实验结果可以看出, 当阈值取 0.50 时, 模型的表现最佳, 因此使用阈值为 0.50 所计算的数据作为

下一步实验的实验数据。

表2 不同阈值下 C-LSTM 模型的表现

C 阈值	Precision	Accuracy	F1-score
0.30	0.785	0.779	0.761
0.35	0.815	0.813	0.801
0.40	0.853	0.848	0.838
0.45	0.890	0.882	0.875
0.50	0.917	0.917	0.913
0.55	0.871	0.882	0.878
0.60	0.844	0.848	0.840
0.65	0.828	0.825	0.813
0.70	0.805	0.790	0.771

4.5 实验结果 实验使用 5-折交叉验证法充分利用数据进行训练。本文复现了目前效果比较好的网络模型,如 Castillo 等人^[14]的 bayes 模型,DecisionT 模型,Ma 等人^[15]的 SVM 模型,Yu 等人^[3]等 CNN 模型,Chen 等人^[5]的 RNN 模型,Nguyen 等人^[16]的 CNN-RNN 模型,将这些模型应用在本文所构建的数据集上,分别比较了 Precision, Accuracy, Recall, F1-score 四项指标。表 3 展示了本文的 C-LSTM 模型与其他模型的效果对比。

表3 实验结果对比

模型名称	Precision	Accuracy	Recall	F1-score
Bayes	0.807	0.821	0.839	0.823
SVM	0.836	0.846	0.853	0.844
DecisionT	0.853	0.865	0.875	0.864
CNN-RNN	0.878	0.887	0.890	0.884
RNN	0.896	0.896	0.882	0.889
CNN	0.897	0.903	0.906	0.901
C-LSTM	0.917	0.917	0.908	0.913

通过对比实验可以看出本文提出的 C-LSTM 模型的检测效果在 Precision, Accuracy, Recall, F1-score 四个方面均优于其他模型。

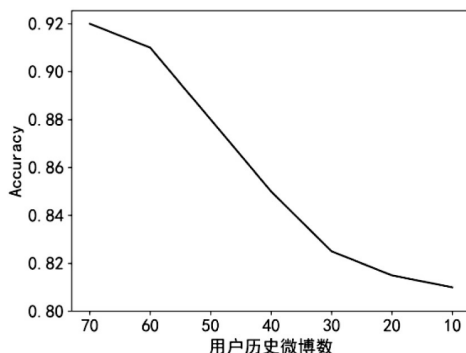


图3 历史微博数量对检测准确率的影响

本文还研究了在分析用户历史行为时不同历史微博数据量对于检测准确率的影响,结果如图3所示。更多的历史数据意味着能更好地刻画一个用户的行为特征,但也意味着需要耗费更多的时间,检测的时机也就要相对延迟,因此需要通过权衡,找出最佳的策略。

虽然本文所提出的检测方法依赖于用户的历史微

博数据,过少的历史数据将影响检测效果,但是这种影响是有限的。根据谣言传播理论的研究,谣言的传播范围和影响力和发布谣言的用户影响力成正相关,通常活跃度较高的用户发布的谣言容易造成广泛传播,而一些不活跃的用户,甚至于小号、水军号,他们的日常活跃度很低,使得这些账号发布的言论无法广泛传播,危害性低,不作为主要研究对象。因此,谣言检测研究中的大多数谣言发布者都是拥有相当数目的历史数据的,不会因为大量用户缺乏历史数据而显著影响检测效果。

4.6 早期检测效果 在检测时效性方面,我们从上述研究所使用的数据集中,随机抽取 100 条微博样本分析了它们的评论随时间的产生情况:评论最少的有 146 条,最多的 1 189 条,平均每条微博拥有 506 条评论;从微博发布到产生第 100 条评论平均用时 14 小时;不计评论数不足 300 的微博,产生第 300 条评论平均用时 23 小时;不计评论数不足 500 的微博,产生第 500 条评论平均用时 31 小时;对于使用微博全部评论数据作为样本的研究而言^[11],其平均检测时效就是微博发布后的 31 小时。而对于 Ma^[15]等采取划分时间序列进行早期检测的研究而言,需要每条微博平均拥有大约 300 条评论时,检测效果才趋于稳定,也就是说平均检测时效为 23 小时,使检测时效提前了 8 个小时。而本文提出的基于用户历史数据进行谣言检测的方法无需依赖评论信息,可以在谣言产生之初便将其识别,使检测时效能提前 31 小时。

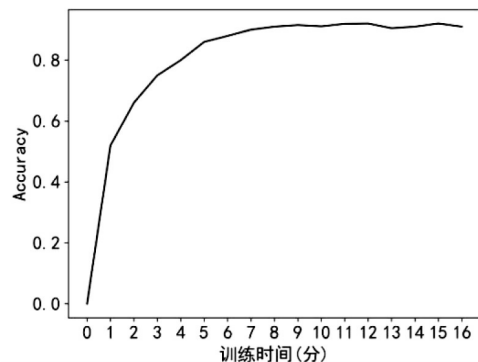


图4 验证集的准确率随训练时间的变化

图4展示了本文所提方法在训练中验证集的准确率随时间的变化,训练在 16G 内存,8 核 CPU,7 代 i7 处理器,GTX 960 显卡的机器上进行。可以看出模型在训练大约 8 分钟以后就能达到比较稳定的准确率。用户历史微博分析是单独进行的,对所有的 13 万余条历史微博进行分析大约耗时 4 小时。但是这并不意味着模型需要在谣言发生 4 小时后才能将其检测出来,对于单个用户的 70 条历史微博进行分析可以在数十秒的时间内完成。因此,在检测模型已经训练好和微博数据爬取不受限制的情况下,对于谣言的检测时间

主要取决于发布谣言用户的历史微博数据的爬取与分析。

通过实验结果和分析可以看出本文所提出的用户特征分析方法能大大提高检测时效,实现谣言的早期检测,同时,基于 C-LSTM 的检测模型相对于其他的检测模型拥有更好的表现。

5 结 论

本文所提出的基于用户历史特征分析的微博谣言早期检测方法,不依赖微博的评论转发信息,而是借助历史信息特征分析以及情感分析技术实现谣言的早期检测,可以在谣言产生之初将其检测识别出,通过对比实验进一步证明了本文所提出的方法的可行性和有效性。本文的不足之处是该方法需要大量的用户历史微博及其评论信息支撑,而在历史数据不足的情况下,检测效果就会有所下降。为实现有效早期检测,需要深度挖掘用户特征和检测文本的深层次特征,进一步研究更高效的检测算法,在不断降低历史数据依赖量的情况下保持较好的检测效果。

参 考 文 献

- [1] Sejeong K, Meeyoung C, Kyomin J, et al. Rumor detection over varying time windows[J]. Plos One, 2017, 12(1): E0168344.
- [2] Zhao Z, Resnick P, Mei Q Z, et al. Enquiring minds: Early detection of rumors in social media from enquiry posts[C]//ACM. Proceedings of the 24th international conference on World Wide Web. New York: Assoc computing machinery, 2015: 1395-1405.
- [3] Yu F, Liu Q, Wu S, et al. A convolutional approach for misinformation identification[C]//IIIA-CSIC. Proceedings of the 26th international joint conference on artificial intelligence. Melbourne: IJCAI, 2017: 3901-3907.
- [4] Jin Z W, Cao J, Guo H, et al. Multimodal fusion with recurrent neural networks for rumor detection on microblogs[C]//ACM. 25th ACM international conference on multimedia. New York: Assoc computing machinery, 2017: 795-803.
- [5] Chen T, Li X, Yin H, et al. Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection[C]//Ganji M. Pacific-Asia conference on knowledge discovery and data mining. Melbourne: Springer, 2018: 40-52.
- [6] Yang F, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo[C]//Ding Y. Proceedings of the ACM SIGKDD workshop on mining data semantics. Beijing: ACM, 2012: 1-7.
- [7] Wu K, Yang S, Zhu K Q. False rumors detection on sina weibo by propagation structures[C]//ICDE. 2015 IEEE 31st international conference on data engineering. Seoul: IEEE, 2015: 651-662.
- [8] Morris M R, Counts S, Roseway A, et al. Tweeting is believing?: Understanding microblog credibility perceptions[C]//Morris R. Proceedings of the ACM 2012 conference on computer supported cooperative work. Washington: ACM press, 2012: 441-450.
- [9] Nourbakhsh A, Liu X, Shah S, et al. Newsworthy rumor events: A case study of Twitter[C]//Cui P. IEEE 15th International conference on data mining workshops. Atlantic: IEEE, 2016: 27-32.
- [10] Vosoughi S, Mohsenvand M N, Roy D. Rumor gauge: Predicting the veracity of rumors on Twitter[J]. ACM transactions on knowledge discovery from data, 2017, 11(4): 1-36.
- [11] Lin D, Ma B, Cao D, et al. Chinese microblog rumor detection based on deep sequence context[J]. Concurrency and computation: Practice and experience, 2019, 31(23): E4508.
- [12] Lan T, Li C, Li J. Mining semantic variation in time series for rumor detection via recurrent neural networks[C]//IEEE. 20th IEEE international conference on high performance computing and communications. Exeter: IEEE, 2018: 282-289.
- [13] Jing M, Wei G, Mitra P, et al. Detecting rumors from microblogs with recurrent neural networks[C]//Arizona State University. Proceedings of the 25th international joint conference on artificial intelligence. New York: AAAI, 2016: 3818-3824.
- [14] Castillo C, Mendoza M, Poblete B. Information credibility on Twitter[C]//Romero D. Proceedings of the 20th international conference on World Wide Web. Hyderabad: ACM press, 2011: 675-684.
- [15] Ma J, Gao W, Wei Z, et al. Detect rumors using time series of social context information on microblogging websites[C]//ACM. Proceedings of the 24th ACM international on conference on information and knowledge management. Melbourne: Association for computing machinery, 2015: 1751-1754.
- [16] Nguyen T N, Li C, Nieder E C. On early-stage debunking rumors on Twitter: Leveraging the wisdom of weak learners[C]//Ciam-paglia G L. International conference on social informatics. Oxford: Springer, 2017: 141-158.

(责编/校对:王平军)

(上接第 58 页)

头:内蒙古科技大学, 2017.

- [11] 江 屏, 张瑞红, 孙建广, 等. 基于 TRIZ 的专利规避设计方法与应用[J]. 计算机集成制造系统, 2015, 21(4): 914-923.
- [12] 成思源, 王 瑞, 杨雪荣, 等. 基于 TRIZ 的专利规避创新设计[J]. 包装工程, 2014, 35(22): 68-72.
- [13] 梁艳红. 基于专利挖掘的创新设计关键技术研究[D]. 天津: 河北工业大学, 2011.
- [14] 杨玉美. 基于 TRIZ 的专利布局方法研究及应用[D]. 天津: 河北工业大学, 2017.
- [15] 祝凤金. TRIZ 技术系统进化法则在专利布局中的应用研究[J]. 科技管理研究, 2010, 30(19): 148-151.
- [16] 白光清. 医药高价值专利培育实务[M]. 北京: 知识产权出版社, 2017.
- [17] 马天旗. 高价值专利培育与评估[M]. 北京: 知识产权出版社, 2018.
- [18] 支苏平. 高价值专利培育路径研究[M]. 北京: 知识产权出版社, 2018.

(责编/校对:刘影梅)