分类号: 密级:

学号: 107622017210742

单位代码: 10762

新疆师范大学 2019 届专业硕士学位论文

基于评论情感和 CNN-LSTM 结合模型的微博谣言检测 Microblog Rumor Detection Based on Comments Sentiment and CNN-LSTM

研究生姓名: 吕胜

学位类别:专业硕士

专业领域: 计算机技术

院系、年级: 计算机科学技术学院 2017 级

指导教师:张海军(教授)

新疆师范大学 2019年5月17日

摘要

随着社交网络的兴起,新浪微博成为中国活跃用户最多的网络信息平台,也成为谣言滋生和爆发的传播媒介。谣言检测能够服务于预防和监管等谣言清理任务,如何自动高效的检测谣言,成为了自然语言处理领域研究人员的重点研究方向。国内外科研工作者的研究大多数都是基于传统机器学习方法,如特征工程。这种方法需要大量专业背景知识,难度大,费时费力。深度学习模型可通过大数据学习自动发现规律,提取有用特征,在具体的谣言检测任务中取得不错的效果。

研究常见的 CNN、LSTM 和二者结合的 CNN-LSTM 在谣言检测任务中的应用,按分类问题方法判别检测的文本是否为谣言。但是,微博的用户页面结构不仅有微博内容文本还包括点赞、转发和大量的评论文本信息。评论中含有对微博文本的态度,而这种深层次的特征神经网络往往难以铺捉到。因此,可以把微博的评论情感极性看作一个极其重要的特征加入到微博谣言检测模型中。具体工作如下:

采集了微博辟谣中心发布的谣言微博和等量的正常微博构建训练集,并将LSTM 连接到 CNN 的池化层和全连接层之间组合成 CNN-LSTM 模型,通过设置与CNN、LSTM 模型的对比实验,证实了 CNN-LSTM 模型在微博谣言检测任务上良好表现。最后,本文采用 LSTM 模型做情感分类器,挖掘谣言微博评论的情感特征,提出将评论整体情感强度和情感方差特征融入 CNN-LSTM 模型中,构建了融合情感极性的 CNN-LSTM 微博谣言检测模型。通过实验,证明了该模型的有效性。

关键词: 谣言检测; 深度学习; 情感分析; 词向量

Abstract

With the rise of social networks, Sina Weibo has become the most active user of the network information platform in China, and also the media for rumors to breed and erupt. Rumor detection can serve rumor clearing tasks such as prevention and supervision. How to detect rumors automatically and efficiently has become the focus of researchers in the field of natural language processing. Most of the research of domestic and foreign researchers is based on traditional machine learning methods, such as feature engineering. This method requires a lot of professional background knowledge, which is difficult and time-consuming. Deep learning model can automatically discover rules and extract useful features through large data learning, and achieve good results in specific rumor detection tasks.

This paper studies the application of common CNN, LSTM and CNN-LSTM combined with them in rumor detection task, and distinguishes whether the detected text is a rumor according to the classification method. However, the user page structure of Weibo includes not only the content text of Weibo, but also a lot of praise, forwarding and comment text information. Comments contain attitudes toward microblog texts, and such deep-seated feature neural networks are often difficult to find. Therefore, we can regard the emotional polarity of microblog comments as an extremely important feature and add it to the rumor detection model of microblog. The specific work is as follows:

The training set is constructed by collecting rumor micro-blog and normal micro-blog published by Weibo Anti-rumor Center, and combining LSTM with pooling layer and full connection layer of CNN to form CNN-LSTM model. The comparative experiment with CNN and LSTM model proves that CNN-LSTM model performs well in the task of rumor detection of microblog. Finally, this paper uses LSTM model as an emotional classifier to mine the emotional characteristics of rumor microblog comments. It proposes to integrate the overall emotional intensity and emotional variance features of comments into CNN-LSTM model, and constructs a CNN-LSTM microblog rumor detection model that integrates emotional polarity. The validity of the model is proved by experiments.

Key Words: Rumor Detection, In-depth Learning, Affective Analysis, Word Vector

目录

摍	寶		l
Ak	ostra	ct	
1	绪论		1
	1 1	研究背景	1
		研究意义	
		国内外研究现状	
	1.5	1.3.1 微博短文本情感分析研究现状	
		1.3.2 微博谣言检测研究现状	
	1 Δ	研究内容	
		论文组织结构	
2		世论和技术	
_			
		微博谣言的特点	
	2.2	文本信息处理技术	
		2.2.1 文本预处理	
		2.2.2 词向量化	
		情感倾向分析技术	
	2.4	深度学习算法	
		2.4.1 卷积神经网络	
	2.5	2.4.2 循环神经网络	
	2.5	本章小结	15
3	基于	深度学习的微博谣言检测模型构建	16
	3.1	微博数据的采集和预处理	16
		3.1.1 微博语料库的构建	16
		3.1.2 微博垃圾评论过滤	19
		3.1.3 词向量训练	20
	3.2	基于卷积神经网络的微博谣言检测模型构建	
		3.2.1 卷积神经网络	
		3.2.2 CNN 神经网络实验	
	3.3	基于 LSTM 神经网络的微博谣言检测模型构建	25
		3.3.1 循环神经网络	
		3.3.2 LSTM 模型的门控结构	
		3.3.3 LSTM 神经网络实验	
	3.4	基于 CNN-LSTM 的微博谣言检测模型构建	
		3.4.1 CNN-LSTM 模型	
		3.4.2 CNN-LSTM 神经网络实验	
		实验对比与分析	
	3.6	本章小结	36

4	基于	情感分析和 CNN-LSTM 模型的微博谣言检测	.37
	4.1	基于情感分析和 CNN-LSTM 的谣言检测方法	.37
		4.1.1 文本预处理	
		4.1.2 微博评论的情感分析	.38
	4.2	基于评论情感的谣言检测	.39
		4.2.1 基于 LSTM 的情感分析	.39
		4.2.2 评论情感特征的挖掘	.40
		4.2.3 基于评论情感的谣言检测模型	.42
	4.3	融合情感极性的 CNN-LSTM 的谣言检测模型	.42
		4.3.1 模型的构建	.42
		4.3.2 实验设计	.44
	4.4	实验结果与对比分析	.45
		4.4.1 谣言检测结果	.45
		4.4.2 实验对比分析	.46
	4.5	本章小结	.49
5	总结	· 三与展望	.50
	5.1	总结	.50
	5.2	展望	.50
参	考文	献	.51
致	谢		.55

1 绪论

1.1 研究背景

时至今日,互联网快速发展,技术急速普及,对我们的生活产生全方面的影响,从娱乐到信息交流都产生翻天覆地的变化。互联网哺育了社交平台,社交平台一经诞生就风靡世界各地,受到数以亿计的全球网民的喜爱。微博是目前最优秀的社交产品之一,促进了人与人之间的沟通交流。基于广播技术的社交网络平台,平台中的每个用户既可以是信息的发布者也可以是信息的接受者,使信息共享成为一个人人可以参与其中的事。与传统的博客网站相比,微博的注册和运营门槛更低,每个人都可以在该平台上发挥自己的智慧,实时的分享自己的想法,信息交流传播方便快捷。

在十几年前,国内就已经诞生了许多类似微博的信息平台,在市场的竞争中 Twitter 和微博不断壮大,如今已成为最火热的基于社交关系的网络平台。2006年 Twitter 诞生于美国,由 Evan Williams等人创建,随着不断创新与升级,目前已成为美国用户基数最大的社交平台,其用户人数超过 5 亿,网站访问量排名全球前十。在国内,新浪微博也同样取得巨大的突破,成为信息与社交服务产品的领头羊,占领最高市场份额。据微博 2017年第三季度财报显示,当前月活跃用户超 3.76 亿,日活跃用户超 1.65 亿人,每天发布分享的微博数数亿条[1]。随着技术发展创新和用户体验的完善,微博的用户数必将继续增长。

微博提供了一个便捷的信息交流平台,用户可以在 140 字内自由编辑分享 自己的意见想法,这些文本和图片信息通过转发分享机制在社交网络中自由扩 散和传播。但微博基于自由平等的产品理念,致使用户门槛极低,人人都可以 注册,人人都能分享,虽然大大丰富了平台信息量,同时也造成了许多的不真 实信息的传播。

1.2 研究意义

微博是一个提供展示、分享、学习海量信息服务的全新的网络平,微博获取信息和分享信息只需动动手指头便可,因而微博谣言的危害和影响也更大。根据中国新媒体发展报告显示,网络谣言的首次出现有 60%是出现微博平台,微博成为了不实信息传播的最大媒介^[2]。如何提取谣言特征并设计有效的谣言检测方法,不仅具有显著的理论价值,而且也有重大应用价值和意义。

(1)谣言检测的研究具有显著的理论价值。社科院的新媒体发展报告表明, 2013年至2014年,由广大社交媒体公布的前一百大舆情话题事件中,谣言事件 约占 6%, 针对谣言的预测和监督任务,仍然困难重重^[3]。目前我们还是以人工举报辟谣为主,费时费力;如下图 1-1 所示,是关于 NASA 森林覆盖率数据的微博,经人工举报鉴定为谣言微博。如何自动的提取谣言特征并设计有效的谣言检测方法,具有显著的理论价值。目前,国内外相关领域的科研工作者已取得了大量显著成果。针对这些问题,本文利用深度学习自动提取特征的同时构建了谣言的情感特征,提出建立基于情感分析和深度学习的微博谣言检测模型。

- (2)谣言检测能有效的净化网络空间。微博上的不实信息对我们造成误判。 影响我们真实信息的获取,轻则影响我们的购物旅行,重则影响我们的医疗健 康。微博平台上大量的不实信息不仅影响生活还会造成社会恐慌,不利于国家 和谐发展。谣言检测就是希望利用技术手段将不实信息扼杀,在还未大范围扩 散造成巨大影响之前将其清除,减少个人因相信不实信息造成财产损失,避免 并阻止谣言传播刻不容缓。
- (3)谣言检测能提高网民的监督意识和安全意识。随着网络信息平台的发展,人人都可以通过网络发声,行使自己的表达权。与此同时,网民也可以通过微博辟谣中心举报谣言行使自己的监督权。谣言检测,能让网民更好的识别谣言,不传播谣言,减少上当受骗。
- (1) 网传"NASA的森林面积覆盖率的分析数据"不实,@人民日报已辟谣:实为PS的假地图,然后说成美国NASA航拍的地图。



图 1-1 NASA 森林覆盖率虚假数据人工举报

1.3 国内外研究现状

1.3.1 微博短文本情感分析研究现状

情感分析的概念最初由 Bo Pang^[4]等人在 2002 年提出来。之后,文本信息的情感分析开始被更多人关注并研究,目前已成为了 NLP 领域的重点研究问题。文本情感判别的任务主要是对文本中包含的情感词、情感态度等信息进行研究,然后判断文本整体情感倾向性^[5]。当前,文本情感分析研究以传统机器学习为主,

而深度学习是近年的研究热点,从算法层面主要以监督和非监督来区分,总体上可分为四类。

(1)情感词典的相关方法。该类法相对简单直接,利用文档中的情感词、副词、否定词等,判断情感倾向,但需要已标记的情感词典。目前中文领域的情感词典主要有台湾大学的 NTUSD 词典、大连理工大学的中文情感词汇库,其中大连理工的词汇库是基于 Ekman^[6]的六大情感分类形成的。知网 HowNet 公布的情感分析词语集配合其他词典使用,可做到很好的扩充,效果不错。朱嫣岚^[7]等人通过选定基准词,由当前词汇与基准词的相关性来判断情感极性。王素格^[8]等人为了更好的选取基准词,设计了情感词表和类别区分的结合方法,并考虑了同义词间的情感关联性。桂斌^[9]等人利用义群来划分微博并将程度词、否定词、标点符号的影响考虑其中。

利用情感词典进行情感分析可快速的判断词语的情感极性和计算句子的情感值且不依赖复杂的算法。中文微博存在严重的口语化、流行语、表情符号多和错别字问题,单一的用情感字典分析会难以全面的提取特征。一旦数据清洗不干净,对最终的实验结果影响相对较大。

- (2) 有监督学习的相关方法。Pang^[10]等人通过大量数据统计分析词性特征 同时结合 N-Gram 词语特征,在电影评论情感分类任务上分别选择最大熵、朴素 贝叶斯和支持向量机算法进行试验,发现 SVM 的综合分类性能最好。Davidiv[11] 等人在 twitter 平台的海量数据中重点研究了网络文本中的表情符号问题,并根 据符号特征设计了基于监督学习的分类实验。Go[12]等人在监督学习的基础上增 大监督距离进行分类,并将表情符号加到微博文本数据集中,准确率超过80%。 Huang^[13]等人充分挖掘句子的相似度特征,并在此基础上结合大量用户的反馈 信息去判断情感倾向。唐慧丰[14]等人在做了大量实验后发现, SVM 分类器通过 合理的构建特征当训练集的数据越多时,情感分类的效果越好。谢丽星[15]等人 在层次结构的基础上提出了多策略组合的情感分析方法。孙建旺[16]等人设计了 情感词典与机器学习算法相结合的方法,以形容词和动词为辅助特征,根据情 感词典计算特征情感极性值。李文[17]等人发现阻塞现象会严重影响模型的分类 性能,针对该现象,他们设计了阻塞分布预估模型,然后将阻塞先验知识与分 类相结合,解决了分类模型中文本阻塞问题。有监督学习算法一般有坚实的数 学基础,收敛快,不依赖大数据也能做到不错的分类效果,面对孤立的噪点和 缺失数据不太敏感。监督学习依赖标签数据,面对大数据的情感分析问题时需 要解决数据标注问题。
- (3) 半监督学习的相关方法。该方法只需少量的标注即可。Dai^[18]等人利用自编码读入输入的向量再预测输入序列,实现监督学习的"预培养"。Johnson^[19]

等人将少量的无标记文本输入到监督学习CNN网络中,达到了良好的分类效果。 Mintz^[20]等人通过距离监督的方法产生有标注的数据,但标注错误的数据会成为 噪声干扰实验结果。Tatemura^[21]等人对数据标记过程进行分析,并建立数学模型, 大大降低了标记错误;在松弛版基础上构建了非线性推荐模型(PROP)。半监 督学习是监督学习和无监督学习的结合,使用大量的未标记数据以及同时使用 标记数据来进行模式识别工作。半监督学习既减少了标记工作量,又能带来比 较高的准确性,但需要将监督学习和无监督学习两种技术相结合。

(4)无监督学习的相关方法。当前国内的研究者主要关注邻近算法。此外,Le^[22]等人基于词袋模型提出了段向量算法(Paragraph Vector),能够在不定长文本上提取定长特征。Turney^[23]等人通过互信息(Pointwise mutual information),然后扩展基准情感词汇用于情感分析。Hasegawa^[24]等人提出利用上下文实体的相似度执行聚类分析。无监督学习可以不依赖标签数据和先验知识,直接从原始训练数据中提取特征。面对小规模的数据时特征提取往往比较困难,但能在大规模的无规则的数据中提取特征。

基于情感词典的方法,能充分挖掘情感词的特征,但容易受到网络文本中的垃圾数据的影响,对数据纯净度要求较高。基于有监督的微博文本情感分析方法,依赖已经标记好的语料,语料标记是一项费时费力的工作。无监督学习方法无需标注的特性很适合微博平台的海量数据,但为了提高情感分析的准确率,采用基于半监督的情感分析方法更好。微博文本存在话题杂、流行语、表情符号和极性转移多等问题。因此,本文通过构建用户词典过滤大量的垃圾信息,并利用 LSTM 模型实现情感分类。

1.3.2 微博谣言检测研究现状

当前,许多科研人员都利用新浪微博和 Twitter 平台进行自己的课题研究,并引入不同的研究方向。大多数科研工作者视其为文本分类,在分类框架下设计不同特征。有的研究人员注意到不实信息的传播特征,以及可信度特征。

分类问题离不开特征工程,好的特征工程是分类问题的关键,试验检测的准确率和特征选择的好坏关系极大。因此,谣言检测的过程简而言之就是的如何提取特征的过程。首先收集大量的语料,建立语料库,通过对海量语料的充分分析,提取有用的特征来设计分类模型,实现良好的检测效果。谣言特征主要从文本内容、谣言传播、可信度三个方面来提取。

(1) 基于微博文本的谣言检测相关研究

Vahed Qazvinian^[25]等人对 twitter 上的谣言检测问题进行了研究,谣言检测 是一个复杂的自然语言处理问题。通过选择微博文本特征、用户特征以及相关 字符标签特征,选取不同的特征在贝叶斯分类器上实验,实验表明微博文本内容特征的谣言检测效果最优。Carlos Castillo^[26]等人选取了信息特性、话题特征和传播特征等利用机器学习算法对 twitter 谣言进行检测,实现了较高的准确率。Kwon^[27]等人研究谣言在不同时期的特点,从语言特征、用户、时间特征、网络结构几个方面进行了分析总结。分析的结果显示:从较长的时间段上看,可利用传播特征和时间特征来区分,但该方法不适合谣言传播初期。在传播初期可通过微博内容和用户特征来分析。基于以上研究,Kwon设计了基于时间窗口的谣言检测模型。陈亮^[28]等人基于 BP 神经网络并改进激活函数,对微博谣言的传播过程进行检测。

路同强等人^[29]为了解决微博谣言数据的不均衡问题,对 Co-Forest 算法进行了改进设计了新的 ImCo-forest 算法。ImCo-forest 算法通过 SMOTE 算法与分层抽样加大数据集中数量较少数据的量,同时利用代价因子提高未标注数据预测的正确率。毛二松^[30]等人把深层特征和集成学习相结起来进行微博谣言检测。微博的深层特征主要包括情感特征、用户影响力特征和评论特征等。陈彦舟^[31]等人为了解决海量数据中的舆情分析问题,利用分布式系统处理微博的海量数据。Yang^[32]等人研究并总结了新浪微博谣言的特征。他们构建了微博内容、关注数、客户端类型、谣言时间等特征,并将这些特征输入到 SVM 模型中进行分类。实验证明,新加入的特征能明显提高实验结果。Sun^[33]等人对微博平台上的重大事件进行了研究,他们从微博文本、微博视频图片和微博用户三个方面提取特征,利用贝叶斯和决策树等机器学习算法进行试验,取得了不错的效果。

当前微博谣言检测的研究还是以传统机器学习方法为主,这些浅层的学习方法在文本表示方面还不理想。一方面是由复杂的语义和时序结构决定的,另一方面在提取微博的谣言特征时相对单一、浅层次,特征提取不够完备。

(2) 基于传播的谣言检测相关研究

网络平台的谣言传播是一个复杂且有难度的研究课题,它涉及多个学科领域包括信息学、社会学和计算机科学等。

有些研究者基于复杂网络理论,在无标度的网络中进行仿真实验,分析仿真网络的结构和参数,来探索谣言传播的特点。Mendoza^[34]等人关注了了 2010年的智利大地震,对 Twitter 该话题下的用户行为进行了研究。爬取 Twitter 上的相关数据,重点研究网民在智利地震发生后的转发行为并构建了地震谣言和相关新闻的传播结构。实验表明,谣言与正常新闻的传播路径和结构有很大不同。谣言的讨论明显激烈些,传播也相对快一些。Tetsuro Takahashi^[35]等人采集了日本海啸发生后 twitter 上用户的评论数据,分析发现可以从谣言微博的数量增长和转发率等方面寻找特征。Armineh^[36]等人采集了几百条谣言,研究其中的语义

特征,同时研究了谣言转发时用户扮演角色不同对谣言传播结构的影响。

赵金楼^[37]等人根据社会网络的分析方法,研究了突发事件在社交平台上的传播结构,同时研究了微博网络自身的结构对微博文本信息的传播影响。SNA 是探索社会结构与节点关系的研究方法,重点研究节点关系特征和关系对社会结构产生的影响。他们研究了"四川雅安地震",通过计算和测量网络中的节点,发现受难者和政府工作人员在舆情传播中有较大话语权。因此,这些关键点对信息的传播影响巨大。在谣言传播中最常见的是流行病传播模型,该模型有 SIS和 SIR 两种。根据 SIS 模型,用户可分为感染者和易感染者。易感染者可能会受到感染者的信息,感染者则有可能脱离信息转换成易感染者。在 SIR 模型中,S表示未接受到信息的人群,I表示已接收且还在传播的人群,R表示接受但未继续散播的人群。

当前关于谣言传播的研究中,许多学者只关注谣言传播的拓扑结构和传播 网络中的关键节点对整个谣言传播的影响。目前许多谣言传播模型都是基于流 行病模型的改进,然后在不同的影响因素细分研究。很多传播结构的研究都是 基于仿真网路,真实的网络谣言传播数据较少,对谣言传播中每个个体用户的 研究较少。

(3) 基于可信度的谣言检测相关研究

当前针对微博可信度的研究工作中,Yu Suzuki^[38]等人设计了评价推特信息可信度的方法,通过推特内容转发保留率判断信息可信度。若转发后原推文信息几乎保持不变,则可信度较高;反之,当推文可信度较低时,转发时用户常常会加入自己的见解。通过这个特征,可以根据推文在转发时被修改的情况判定推文信息的可信度。Aditi Gupta^[39]等人研究热门事件发现,推特上与热门事件相关的虚假信息较多。卞先华^[40]等人构建了社会结构与文本信息相结合形成的信任模型,基于此模型可以快速的计算可信度。吴树芳^[41]等人以基于 HHTS 构建了用户可信度计算方法。HITS 计算网页的重要性主要参考权威度与中心度,他们除此之外还将文本内容和用户行为综合起来计算用户的可信度。可信度研究,主要考虑用户的可信度和影响力,同时考虑信息来源可信度。

当前基于文本的研究多以传统机器学习为主,特征单一、层次浅;基于传播的研究很多都是仿真网路,缺乏真实传播数据;基于可信度的研究,来源可信度需要人工建立标准,直接从内容上判断为谣言难度比较大。针对这些问题,本文提出通过深度学习模型学习大量的谣言数据自动提取微博谣言的特征,同时融入微博评论情感等深层次的特征,来进行谣言检测。

1.4 研究内容

微博是一个开放、自由、免费的网络信息交流分享平台,人人可以发布和分享信息,海量的用户和超高的活跃度致使微博上的信息传播速度极快,一旦产生不真实信息便会迅速传播出去,造成严重的个人财产损失和社会影响。

人工谣言检测需要巨大的人力和金钱,耗时长。目前谣言检测模型大多从谣言内容、谣言传播和用户属性三个方面构建特征。为了解决传统费时费力的问题,设计了基于 CNN 的谣言检测模型、基于 LSTM 的谣言检测模型和基于 CNN-LSTM 的谣言检测模型。将微博谣言文本向量化,微博文本中隐藏的深层特征通过神经网络的学习训练来挖掘。它避免了特征构造的问题,并且可以找到不容易找到的特征,从而产生更好的结果。

微博的开放互动让朋友们热情地参与微博话题。基于这种高参与度使得微博谣言传播得更迅速,在大量激烈的在线讨论中使谣言走向澄清、破灭。因此,本文认为微博用户的反馈行为可以作为谣言检测的一个重要特征。与简单的转发数量和评论数量相比,微博评论可以提供更多信息。

通过观察谣言发现公众相信并传播谣言通常是由心理因素和情感因素决定的,刺激负面情感比积极情感更容易。因此,谣言通常营造更煽情的情节,煽动公众的愤怒,恐惧,焦虑,实现快速扩散。正是由于这种故意谣言,谣言不仅仅是一个谣言,即使是同一主题的真实信息,也引发了更为笼统和更为激烈的负面评论。同时,微博谣言评论包含大量的质疑、反驳的文本信息,大大增加了负面情感词汇在评论中的占比。

基于这种思路,本文将新浪微博谣言事件作为研究对象,在前人研究基础上,把神经网络和评论情感作为研究重心,提出用 CNN-LSTM 模型提取基本谣言特征和将评论整体正负倾向比例和情感方差作为新的谣言特征,构建了融合情感极性的 CNN-LSTM 谣言检测模型用于谣言检测。主要研究如下:

- (1)本文采集了新浪微博辟谣中心的谣言微博,同时采集等量的正常微博数据并构建完善的微博语料库。直接从微博平台爬取数据会存在大量的垃圾信息,干扰实验结果。本文重点研究了数据的预处理和垃圾评论的过滤,为实验奠定良好的基础。
- (2)为了减少人工构建特征和标注的工作量,文本对比了常见的 CNN 模型、LSTM 模型以及基于两者设计的 CNN-LSTM 模型在微博谣言检测上的效果,通过设置不同词向量模型、不同词向量维度和模型参数的对比实验,设计出在谣言分类表现最好的神经网络模型。
 - (3) 谣言捏造者刻意散布的谣言包含了自己的情感态度,加上谣言在公众

的讨论中产生大量的疑虑和反驳,所有这些都使得谣言微博的评论比普通的微博更倾向于负面情感且情感波动较大。因此,本文提出将评论整体情感强度和情感方差作为一个新的谣言检测特征,使用基于 LSTM 神经网络的分类器来标注评论情感值,计算谣言文本的整体情感强度和情感方差作。

最后,将新特征结合到 CNN-LSTM 模型中,形成融合情感极性的 CNN-LSTM 谣言检测模型,通过实验证明该模型的有效性。

1.5 论文组织结构

本文的结构安排如下:

第一章,绪论。细致描述了谣言检测历史和现状,同时概述了情感分析的研究状况。最后,说明了文章的主要工作内容和研究方向。

第二章,相关理论和技术。首先总结了微博谣言的特点,详细介绍了文本预处理的相关操作。然后从当前主流深度学习的方法构建分类器等方面进一步详细介绍,主要包括卷积神经网络和循环神经网络等方法,最后重点讲解的两种神经网络的结构和特点。

第三章,基于深度学习的微博谣言检测模型构建。研究对比了 CNN 和 LSTM 两种模型的实现算法和优缺点。为了结合 CNN 在局部特征上的提取能力和 LSTM 在时序信息上的记忆能力,尝试将 CNN 的全连接层替换成 LSTM,构建了 CNN-LSTM 模型。最后通过三组基于 LSTM,CNN 和 LSTM+CNN 模型的对比实验,分析各模型的优劣,证明了 CNN-LSTM 模型的有效性。

第四章,基于情感分析和 CNN-LSTM 模型的微博谣言检测。描述了微博谣言检测的整体方法框架以及具体实现过程。利用爬虫构建微博的文本数据集;基于 LSTM 模型建立情感分类器,对评论进行情感标注。最后将基于情感的谣言识别模型与 CNN-LSTM 进行融合,得到融合情感极性的 CNN-LSTM 谣言检测模型。设置对比实验,证明了本文构建的方案的有效性。

第五章,总结与展望。总结目前研究的不足之处及以后可以改进的地方。

2 相关理论和技术

2.1 微博谣言的特点

微博是国内目前最优秀的社交平台之一,以用户关系为基础,进行人与人的连接。微博自由平等的产品理念大大的促进了微博信息的传播。五湖四海的人汇聚在微博,快捷的沟通和获取信息方式,大大提高了用户的活跃度,积极参与各种讨论交流中。微博发布数以亿计的信息,巨大的信息洪流下,微博也成了谣言滋生的好土壤。通过统计分析,总结出微博谣言有如下特点:

- (1) 微博内容来源广。微博门槛低,只要通过简单的注册流程就可以成为 微博用户,便能自由发布信息。用户通过电脑或者手机无需审核便可发布信息, 发布内容没有限制,其真实性往往无法保证。而且微博用户的职业背景和教育 经历各不相同,微博的来源和内容都极其广泛。
- (2)内容夸张,吸引眼球。在微博这个媒体平台中,人人都可以是记者,每天都能记录故事和创造新闻,满足大家分享故事和窥探世界的欲望。要想获得大家的关注,便需要制造爆炸性的新闻,每个大新闻的背后都是道德的考验,很难做到不夸大或歪曲事实。微博具有转发功能,很多微博会分享一些生活常识的总结,大多数用户在面对这种一眼看上去有用的信息时,往往不仔细考究,便先进行转发,这样未经核实的所谓生活常识便被传播开来,很多谣言也是如此。通常,只要微博内容有用或者吸引眼球,便能很快获得传播。微博谣言能吸引到大量关注,往往是因为内容劲爆夸张、煽动性强,在热论中渐渐传播出去。
- (3)能够快速成为热门话题引起广泛热议。微博谣言有时是个人或团队刻意策划的营销炒作事件,是常见的网络营销方法。为了获得流量,推广产品或者宣传品牌,往往会通过制造谣言成为热门话题,引发关注、轰动一时。例如,一些即将上映的电影为了获取好票房,经常设计电影男女主因戏生情的谣言,炒 CP 来获取关注,达到宣传电影的目的。通常热门话题,都有大量的微博评论,表达了群众对话题事件的态度看法。评论对微博来说十分重要,有些评论者添油加醋,将事情越描越黑,扩大了争议并引起轰动效应。

2.2 文本信息处理技术

2.2.1 文本预处理

利用爬虫程序的直接采集的网络文本通常包含大量的垃圾信息,若不进行清理,将严重干扰实验结果。因此,实验前的文本预处理操作十分重要,通常

会对文本进行分词、垃圾过滤、去停用词等操作。

(1) 分词

词作为最小的情感单位在文本情感倾向性分析中是决定句子情感的重要因素,因此,在挖掘评论情感时,能否准确分词变得至关重要。在英文文本中,每个单词由空格分隔,而在中文文本中每个词往往联系在一起,虽然句子、段落之间有相应分隔符,但是单词之间没有明显的分隔符。因此,在进行中文文本分类的相关实验时,分词不仅是必须的而且分词的好坏直接影响实验的结果。中文分词问题已经被许多科研者所关注,目前的相关算法有基于统计的方法和根据字符串匹配的方法,也有学者尝试从理解文本的角度进行分词[42]。国内关于分词系统的研究已经取得了显著成果,例如,jieba分词系统、中科院 NLPIR 汉语分词系统等。因此,本文的分词工作使用第三方工具来完成。

(2) 词性标注

词性 (POS) 简单来说就是由单词的自身特点来分类词语的标准。常见的词性有名词,动词,形容词,连词,副词等,共 14 种。词性标注基于分词,标注后可以知道文本中每个词的词性。前文介绍的 NLPIR, FudanNLP 系统都能实现词性标注的,本文的词性标注也是借助第三方工具完成的。

即使是相同的词在不同的语境中也会出现不同词性,例如,"这个护士正在服务病人"中"服务"是动词,而"这家店的服务很好"中"服务"是名词。因此,同一个词并不能同一标注而要结合语境不同位置的标注可能不同。

(3) 去停用词和特殊字符

在中文情感分析任务中,停用词并不影响文本的情感倾向,因此为了减少数据噪声,更好的训练模型,通常将这些情感词去除掉。汉语常见停用词有代词、助词、介词和语气助词等,一般使用停用词表将其去除,本文使用通用停用词集来构造停用词表。

除了停用词,爬取的微博文本数据中还充斥着大量数字、网络字符、表情包等。这些字符不影响评论情感判断,因此,可以利用匹配中文字符的正则表达式过滤非汉字字符。

2.2.2 词向量化

在自然语言处理任务上,构建一个优秀的词向量模型至关重要,是取得良好实验结果的基础。文本探讨了两种词向量模型,Word2vec^[43]模型和 Glove^[44]模型。Word2vec 包含了对 CBOW(Continuous Bag-of-Words Model)和 Skip-gram(Continuous Skip-gram model)两种模型的训练。

连续词袋模型(CBOW)是根据周围的词计算中心词出现的概率,该模型的

预测次数约等于词向量个数;而 Skip-Gram 根据中心词计算周围词的出现概率,模型的预测次数多于 CBOW 模型。Glove 基于整体语料的统计信息完成训练,是全局词向量模型。

(1) Glove 模型

Glove 是由 Pennington 等人设计的一种无监督的词向量训练模型,该模型基于矩阵分解,不仅包含统计信息而且能学习到不同词向量间的关系。该模型通过遍历语料库得到共现矩阵,然后通过该矩阵中非零元素来训练词向量 [46]。 Glove 训练效率高,当语料库小或者短向量时,依然能取得不错的效果。其中,词共现的次数用矩阵 X 表示,词 w_i 在上下文中出现的次数用 X_i 表示,词 w_j 在上下文中出现的次数用 X_i 表示,词 w_j 在词 w_i 的上下文中的共现次数用 X_{ij} 表示,则 $p_{ij} = p(w_j \mid w_i) = X_{ij}/X_i$ 是词 w_j 出现在词 w_i 上下文中的概率值。因此,可基于共现概率表示两个单词之间的关联性,只需计算这些词与随机选取词汇 w_k 的共现概率比值便可判定词汇 w_k 与词汇 w_i 、 w_i 间的关联性,其关联性模型的计算公式为:

$$F(w_i, w_j, w_k) = \frac{p_{ik}}{p_{jk}}$$
 (2-1)

Glove 由单词词共现概率的 p_{ik}/p_{jk} 值来转换成的最小化目标函数如下:

$$\sum_{i,j=1.X_{ij}\neq 0}^{|V|} f(X_{ij}) (p_i^T q_j + b_i + b_j - \log(X_{ij}))^2$$
 (2-2)

在具体计算时,单词 w_i 的目标词向量为 p_i ,单词 w_j 的上下文词向量为 q_j , p_i 的偏移量为 b_i , q_j 的偏移量为 b_j ,词表规模为|V|, $f(x_{ij})$ 为权值变换函数,故 f(x)可定义为:

$$f(x) = \begin{cases} (x/x_{\text{max}})^{\partial}, & x < x_{\text{max}} \\ 1, & x \in \mathbb{Z} \end{cases}$$
 (2-3)

其中, x_{max} 由语料库规模而定。

(2) Word2vec 模型

Word2vec 是谷歌发布的一个能实现将单词以向量形式表示的工具,通常由 三层神经网络组成。将单词表示为向量的实现可以说是将深度学习算法引入自 然语言领域的关键技术,自然语言处理领域问题转换成传统机器学习问题首要问

第 11 页 共 55 页

题就是要利用 Word2vec 等工具把文本用数字特征表示。Word2vec 把文本转换成神经网络模型能处理的词向量。词向量模型常用于输入文本的处理,在模型训练之前,单个单词的特征数量和向量维度数相一致。

Word2vec 模型在生成语料文本的向量会自动学习到文本的特征。如果一个单词出现在足够多的文本数据中,Word2vec 可以根据单词的出现准确地预测单词的含义,对情感分析而言意义重大。Word2vec 主要有以下几个基本功能:

- a. 利用单词向量的余弦值判断不同单词间的相似性;
- b. 将输入文本生成词向量;
- c. 利用聚类算法可从海量数据集中导出词类。

2.3 情感倾向分析技术

情感倾向分析通常被看作是文本分类问题,一般文本情感可分为正负两类 或者正负中三类。情感取向分析可以根据研究文本对象的粒度分为多个层次, 如词级,词组级,句子级,段级以及章级。

(1) 情感词判别

情感词汇指的是包含情感取向的词,表示喜的喜悦,欢乐等; 表示怒的生气、恼怒等;表示哀的悲伤、伤心等; 表示乐的乐不思蜀、乐此不疲等; 表示愁得烦恼、愁眉不展等; 表示恐的胆怯、畏缩等等。情感词一般划分为正向、负向和中性三种,情感词挖掘和判断是短文本情感分析的基础。

无论是深度学习还是传统机器学习,情感词极性一般在[-1,1]之间取值,当极性值大于零则为正向,反之则为负向,极性值的绝对值大小表示情感强度。

情感词识别方法一般有两种:基于统计规则和基于情感词典。

基于统计方法是利用统计手段在大规模语料库中挖掘情感词特征。主要用于文本中的连词,比如:而且,但是,然而,是,等等。连词前后的文本具有某种关系,如并列、因果、转折、比较等关系。我们可以基于这些连词以及统计特征来区分词语的情感极性。由连词连接的两个单词极性之间存在相关性,如由"和"连接的单词在情感极性上表现出一致性(如美丽和漂亮,两者都包含好看的意思)。

此外,还有基于情感词典的方法。英文词典以 WordNet 为主,中文主要以 HowNet 以及大连理工大学的情感词汇本体库为主。朱嫣岚^[7]等人提出了一种基于语义相似度和语义相关场方法。实验证明,采用基于知网语义相似度的方法有更好的效果。乔春庚^[45]等人利用在扩展的情感词基础上选取情感倾向的词,然后将选择的词加入到种子情感词中。继续这样操作,直到没有出现具有情感极性的新情绪词语。

(2) 语句级情感分类

句子的情感倾向分类比单词倾向分类粒度粗,单词的倾向分类也是句子倾向分类的基础。句子情感分类就是判别句子级文本的整体情感倾向,对句子文本的主客观性要判别,同时要提取句子文本中的细粒度情感(例如评论持有者,评论对象,与表达情绪相关的评论倾向和优势)。例如,句子"小明认为,苹果的手机系统流畅,时尚,并且具有出色的相机性能,每个人都应该拥有它。"此句子中评论者是"小明";评价对象是"手机","操作系统","相机";"流利","时髦"和"优秀"都是正向情感词,整个句子的感情取向被归类为积极的分类。

J.Wiebe^[46]等人利用形容词来判断句子主客观性; Hovy^[47]等人基于情绪词、使用词和情绪词出现的位置三大特征来分类情感句。叶强^[48]等人提出了基于汉语的主观情感词类的结合模型,可以自动判别中文句子在主观性上的相似程度,该方法属于监督学习。无监督学习目前主要基于句法特征,可以表示情感词与评价对象间的修正关系。但是,这种方法也有缺点,往往准确率高而召回率不高。

2.4 深度学习算法

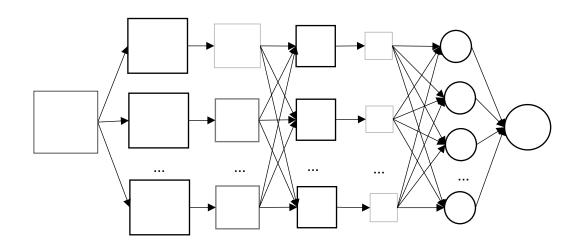
随着深度学习的快速兴起,引发了研究热潮,许多科研人员将相关研究应用于各个领域。但是在谣言检测任务上,基于深度学习的谣言检测相关研究还比较少。谣言检测问题和情感分析问题在某些方面比较相似,都可以基于文本分类这个框架进行研究。因此我们不仅可以将基于深度学习的情感分析研究的相关工作借鉴到谣言检测工作上,尝试用深度学习模型解决谣言问题,还可以利用情感分析技术挖掘谣言的情感特征用于谣言检测。本节将重点讲解深度学习相关算法。

2.4.1 卷积神经网络

卷积神经网络(CNN)是图像识别中常用的深度模型结构,在图像识别、智能搜索、自然语言处理、智能机器人等方面广泛应用。随着深度学习的兴起,CNN网络成了识别领域的研究热点。科学家 Hubel 和 Wiesel^[49]在 1960 年左右,提出了一种卷积神经网络模型,该模型受到了猫的大脑神经元的启发。目前,卷积神经网络已经赢得了研究人员的研究热情,并在不同领域都用相关应用。

Kim 等人^[50]设计了基于卷积神经网络的情感分析方法,通过调节模型参数优化模型便在情感分类任务上取得了良好的效果,证明了 CNN 网络用于情感分类任务是有效的。

CNN 网络是在反向传播神经网络 (BP 网络) 的基础上不断改进而来的,两者之间的区别在于 BP 网络是全连接的,相邻层的所有节点必定都是相互连接的,而 CNN 网络中每个神经元都是单个节点,邻层间的节点并不是全部连接而是部分连接。CNN 结构如图 2-1 所示。



输入 卷积层 C1 池化层 S2 卷积层 C3 池化层 S4 输出层 输出

图 2-1 卷积神经网络的基本原理结构图

CNN模型一般由五种结构组成:输入层,卷积层,池化层,完全连接层和Softmax层。卷积层最为重要,是提取特征的关键,池化层本质上就是采样层,在保留数据特征的同时减少数据维度,以便计算和进行非线性操作,池化层能扩展感知域并保持平移,旋转和缩放不变。

2.4.2 循环神经网络

循环神经网络(Recurent Neural Network,RNN)能够对时序信息建模分析且可以通过反馈结构学习前一时刻记录的信息。循环神经网络的结构特点,使得可以对时序信息进行更深度的学习^[51]。因此,RNN 网络近年来被广泛应用于自然语言处理领域,尤其在文本分类、语义抽取等方面。

循环神经网络能够学习到相关联序列的信息。传统的神经网络,在学习时数据层层向后传递,每层内的节点没有全连接,这样的网络结构在复杂问题的学习中会丢失很多重要信息。而循环神经网络层与层之间以及隐藏层各个节点之间都是全连接的,同时当前时刻的输入信息由上一层的输出信息和当前层前一时刻的输出信息共同决定,图 2-2 为循环神经网络按时间的展开。

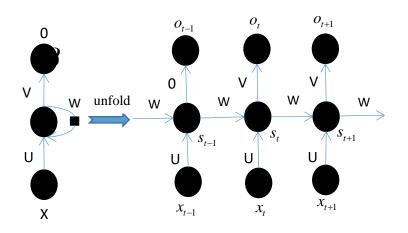


图 2-2 循环神经网络按时间的展开图

在展开图中可以观察到,RNN 网络各个时刻都有信息输入,该网络可以结合当前状态给出该时刻的输出,t时刻的网络状态是由t-1时刻的状态和t时刻的输出一起决定的。RNN 网络特别适合时序信息的处理,在实际任务往往有良好效果。

目前,随着循环神经网络的研究获得新突破,该网络已被大范围的应用在各个方面,如语言翻译和语音识别等,在这些方向上都取得喜人的进步,成果丰硕。RNN与其他的网络相比,其独特之处就是可以分析更长时间段的信息。但是面对一些极其复杂问题时,时间跨度会极其长,可能会导致现梯度爆炸和梯度消失^[52]。针对这类问题,科学家构建许多方案,例如在 RNN 模型的基础上增加门控结构,即 LSTM 模型。这样能很好解决梯度爆炸(消失)问题,使得 LSTM 比 RNN 在语言识别、文本分类等相关领域的实际表现更好。

2.5 本章小结

本章主要介绍相关理论和技术,回顾了微博谣言检测的相关技术。细致的描述了网络谣言的发展历程,同时讲述了谣言检测应该注意的问题和谣言的特征,对其中涉及的情感分析技术和预处理方法做了介绍。最后,结合目前最新的研究,探讨深度学习在谣言检测研究上的应用及 CNN 和 RNN 两种常见神经网络。

3 基于深度学习的微博谣言检测模型构建

本文考虑到人工检测谣言费时费力,速度慢。目前的谣言检测主要以传统的机器学习为主,通常从谣言的文本内容、传播结构以及可信度三方面人工建立特征,人工建立特征不仅费时而且单一。针对这些问题,本文提出了基于深度学习的微博谣言检测模型,通过谣言微博数据训练神经网络自动提取谣言的特征,实现微博谣言的检测,整体方法流程如图 3-1 所示。

方法思路是先从微博平台爬取谣言和非谣言微博数据,建立语料库,过滤垃圾评论并进行分词、词向量化等预处理操作。其次,将文本向量输入神经网络,为了得到更好的实验结果,本文分别了设计了基于 CNN、LSTM 和 CNN-LSTM 三种神经网络的谣言检测模型。最后,设置对比实验,检测本文提出的 CNN-LSTM 模型在谣言检测上的有效性。

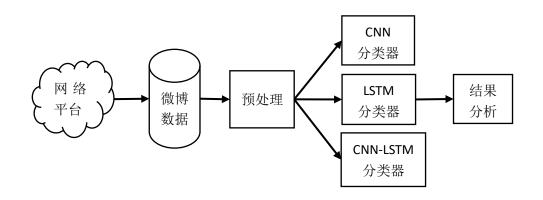


图 3-1 基于深度学习的谣言检测流程图

3.1 微博数据的采集和预处理

3.1.1 微博语料库的构建

微博是国内最活跃最优秀的公众社交平台,凭借信息交流的便捷性这些年取得飞速发展,用户数和月活量都快速增长,每天发布的微博数以亿计,这些微博构成了海量的文本信息。针对微博的研究已成为当下的热点。利用微博的海量数据研究谣言,就必须做好数据采集和清洗工作。在本节中主要描述微博数据如何搜集以及如何构建相应的语料库。

(1) 基于 API 的采集方法

微博平台为了便于用户采集数据,本身提供了相应的数据服务。微博开放平台提供了两百多个 API 接口,可以下载微博内容、微博评论、点赞数、转发

第 16 页 共 55 页

数等相关信息。用户可以利用OAuth2.0获得授权后便可以进行相关信息的收集,构建本地数据库。

利用 API 搜集数据就必须在微博平台上申请并获得认证,简而言之就是取得第三方接入的许可,这样就可以直接从微博服务器中取得想要的文本数据。以前版本的授权必须提供用户密码与名称,这种方式使得用户的信息安全不能被保障,第三方服务在获得用户密码与名称之后有意或无意将其泄露出去,这样极不安全。为了避免安全事故发生,2007 年微博发布新版本的 OAuth 协议完美的解决了泄密问题,保障了用户的安全。

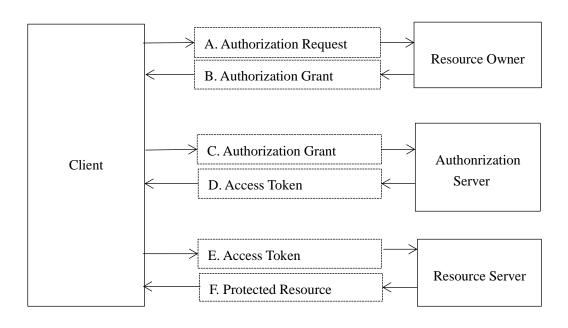


图 3-2 OAuth 协议的授权过程

随后,微博又在 OAuth1.0 版本后经过改良发布了 OAuth2.0 版本,它是比前代更加安全的标准。该协议的详细结构如上图 3-2 所示。通过协议的授权流程图,在客户端取得访问码,就可以向服务器验证并获取资源了。获取的数据一般有两种格式,XML或 JSON。当然除了利用 API 获取资源,也可以直接利用爬虫技术进行数据搜集。

(2) 基于爬虫技术的采集方法

网络爬虫程序就是按照网络协议能够自动采集网页信息,从初始 URL 爬取,爬虫程序利用搜索引擎及网页超链接获取网页信息。利用模拟登陆技术获取页面上的更深层次的信息,直到爬取的数据满足要求了便可以停止爬取数据。在爬取数据的过程中,一般要关注一些问题,如怎样过滤页面数据,爬取的策略该如何设计等。然后对页面内容进行分析,内容分析就是利用现有的数据内容特征,使用分类、聚类等算法并结合关联网页去分析页面数据信息。爬虫相

关的搜寻策略常见有:深度优先搜索、广度优先搜索以及优先级搜索。

爬虫程序要完成数据采集工作就必须用用户名和密码申请微博服务器授权并完成目标页面的模拟登陆。爬虫获取页面数据的详细过程如下图所示:

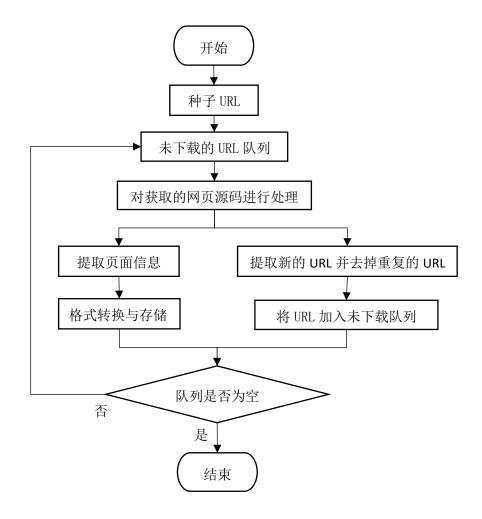


图 3-3 爬虫程序爬取过程流程图

(3) 语料库的构建

由当前想要搜集海量的微博文本数据并不简单,而针对微博的相关公开文本语料比较少。因此在解决微博的谣言检测问题时,通常需要自己采集自己的研究语料。文本选择 API 和第三方爬虫程序相结合的方案,微博数据集包含谣言信息与非谣言信息。

由于新浪公司直接在微博社区管理中心公布不实微博信息,因此我们可以直接去该中心页面上爬取谣言文本构建谣言微博数据集。这种方式既简单又快速还避免人工标注的问题,减少了大量人力和时间。本文章主要搜集了从 2015 到 2018 年间管理中心公布的谣言微博信息共 10229 条。

同时采集这些微博的举报者微博名称、举报者微博链接、发布谣言者的微

第 18 页 共 55 页

博名称、谣言评论。然后爬取相同数量的正常微博,通过限制至少是蓝色 V 用 户发布的信息来确保信息质量。最后,获得了 10200 个常规微博,以及相关的 评论与转发。

最终, 谣言微博与非谣言微博各采集一万多条, 共计 20429 条, 训练集和 测试的比例为 9:1, 相关评论近 15 万多条。

3.1.2 微博垃圾评论过滤

因为微博的自由与开放,导致注册门槛低,用户质量层次不齐,发布的信息 也比较随意不严谨,信息质量无法保障。有些用户为了增加人气,会购买许多 "水军",使得自己的评论信息含有大量垃圾信息。垃圾评论主要来自营销账 号、重复评论、网络垃圾字符等方面。微博平台中存在大量垃圾文本,如果不加 以过滤,这些噪声数据会研究的实验数据造成影响。因此,将这些噪声数据过 滤掉极其重要。

进行预处理后,便可以设置一些特征来进一步过滤文本信息。过滤特征的 选择我们可以从以下几个方面进行研究。

(1) 判断相似度

垃圾文本一般在内容上与微博原文表现出不相关, 所以可以根据与原微博 是否具有相似性来判断垃圾文本。只需要分别将原微博文本和相关评论数据向 量化,比较余弦距离即可。距离越远,则为垃圾文本的概率越大。

(2) 超链接数

由于很多垃圾评论来自广告营销账号,因此超链接数较多的评论往往为垃 圾评论的概率较大,通常利用正则表达式来识别评论中的超链接。

(3) 评论的回复数量

有的评论会被评论者继续评论, 评论数越多说明探讨的越激烈, 该评论为 垃圾评论的概率也越小。

(4) 统计名词比例

基于统计数据属于发现,评论一般都是发表看法,情感词汇较多。而垃圾评 论则相反,往往为陈述句,名字所占比例较高。我们只需要统计评论中的名词 占比就能判别是否为垃圾评论。计算名词占比公式如下:

$$P(Noun) = \frac{N_{noun}}{N_{sum}} \times 100\%$$
 (3-1)

评论中还包含评论回复数、评论点赞数、评论长度以及重复次数等特征,我们可以经过统计分析其中的规律。其中评论点赞数和回复数越多,表明参与讨论和认同该信息的人越多,该评论被认为是正常评论的概率越大。而评论文本越长且重复次数越多的评论,一般复制粘贴而来的较多,被判断为"水军刷评论"的概率较大。最后可以根据相关特征,利用 Logistic 回归模型作为分类器,识别垃圾评论。

3.1.3 词向量训练

词向量就是将文本信息数字化实现可计算的的方法,将文本映射到高维空间,用向量表示。本章实验数据均采用 Word2vec 与 Glove 训练,主要介绍 Word2vec 的训练方法。词向量训练首先要进行预处理,先创建词汇表并统计词频,由高往低排。然后将每个词都生成向量,最后确定向量维数即可。

词向量处理步骤:

- (1) 设置窗口大小,对每个词生成训练样本;
- (2) 设置 batch size, 保证每个 batch 含有所有样本;
- (3)训练算法,设置迭代次数,得到参数矩阵,参数矩阵转置后每一列即 是对应词的词向量。

图 3-4 词向量

本小节生成的词向量将作为后面深度学习的输入,在深度学习模型中进一步学习,最终实现谣言微博与非谣言微博的分类。

3.2 基于卷积神经网络的微博谣言检测模型构建

3.2.1 卷积神经网络

卷积神经网络诞生于 20 世纪 60 年代,在 1980 年被 K.Fukushima^[53]第一次 实现。目前 CNN 网络在计算机视觉方向,如图像识别等问题上获得巨大突破,被越来越多的人研究。本节研究利用 CNN 网络来训练文本,进行微博谣言检测。

简单的 CNN 结构包括卷积层、池化层和全连接层。CNN 网络的特点是,权值共享、下采样和多核卷积等^[54]。权值共享指,用同一个卷积核对图片进行卷积,卷积核的值不变。CNN 网络在卷积层可以设置多个卷积核同时执行卷积操作,每个卷积核都会生成一个对应的 feature map,卷积核不同生成的 feature map 也不同,而两者的数量一般相同。下采样指利用池化操作降低数据维度。

简而言之, CNN 利用卷积操作来提取特征,采用池化操作降低参数维度,最后通过连接层的神经网络实现分类,具体 CNN 结构图如 3-5 所示。

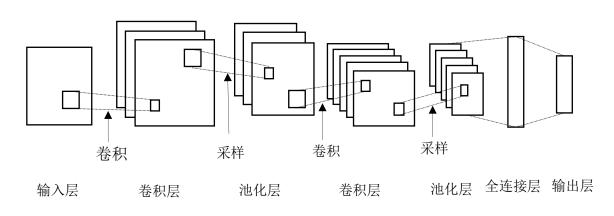


图 3-5 CNN 网络结构图

在 CNN 网络中的 filter 设置不同,卷积出来的 feature map 也不同,feature map 包含大量的局部特征。本章训练 CNN 网络自动提取谣言特征,根据提取的特征判断微博内容是否为谣言。

3.2.2 CNN 神经网络实验

(1) 评估标准

本文的谣言检测属于分类问题,可直接使用二分类问题中使用的评价标准评价所提出的谣言识别模型,我们采用二分类问题中常用的指标精确度、准确率、召回率和 F1 值作为评价标准。具体标准如表 3-1 所示:

第 21 页 共 55 页

表 3-1 模型预测结果分类

	实际为谣言	实际为非谣言
被分类为谣言	TN	FN
被分类为非谣言	FP	TP

准确率:
$$Accuracy = \frac{TN + TP}{TN + FN + FP + TP}$$
 (3-2)

以谣言微博为例给出精确度、召回率和F1值的计算公式如下所示:

精确度:
$$Precision = \frac{TN}{TN + FN}$$
 (3-3)

召回率: Re
$$call = \frac{TN}{TN + FP}$$
 (3-4)

F1 值:
$$F1-Measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
 (3-5)

其中,非谣言的相关计算公式可由上述公式推出。

(2) CNN 实验

在本节实验中我们分别基于 Word2vec 和 Glove 模型生成词向量,选择 TensorFlow 框架进行试验。

基于 Word2vec 模型进行试验:

首先,设置 CNN 网络超参数初始值,如表 3-2 所示。其次,将微博语料训 练成词向量并输入 CNN 模型中进行训练。最后,设置不同维度的词向量对比实 验,调节参数使模型性能最优。

表 3-2 模型超参数初始值

dropout	激活函数	学习率	批处理大小	迭代次数
0.5	relu	0.001	64	20

其中, Precision、Recall、F1-Measure 对应的是谣言微博的精确度、召回率、 F1 值,实验数据如表 3-3 所示。

评估指标				
词向	Precision	Recall	F1-Measure	Accuracy
量维度				
50	0.8318	0.9131	0.8706	0.8661
100	0.8273	0.9105	0.8669	0.8623
150	0.8382	0.9132	0.8741	0.8675
200	0.8401	0.9143	0.8756	0.8691
250	0.8439	0.9199	0.8803	0.8761
300	0.8485	0.9157	0.8808	0.8725
350	0.8470	0.9149	0.8797	0.8692

表 3-3 word2vec 词向量下 CNN 谣言识别性能表现

由表 3-3 的实验数据可知,不同纬度的词向量对模型的准确率有较大影响。在 50 纬词向量时最初准确率只要 86.61%,然后随着纬度的增加准确率也有所增加并达到最高准确率 87.61%,最终在 350 纬时准确率又回到了 86.92%。由以上数据可知,在 250 纬度词向量下 CNN 神经网络取得了最高准确率:87.61%,模型测试结果如图 3-6 所示。

```
For train_samples: step 14200, loss 0.172192, accuracy 0.9375
Testing loss: 0.501615, Testing accuracy: 0.867287
For train_samples: step 14300, loss 0.189703, accuracy 0.9375
Testing loss: 0.447957, Testing accuracy: 0.875612
For train_samples: step 14400, loss 0.0460306, accuracy 1
Testing loss: 0.403775, Testing accuracy: 0.880509

Testing loss: 0.452609, Testing accuracy: 0.876102
TN: 931 FN: 172 FP: 81 TP: 859
rumor
f1: 0.8803 precision: 0.8439 recall: 0.9199
```

图 3-6 Word2vec 词向量 250 维时 CNN 网络各项评价指标结果

Word2vec 模型下, CNN 网络的最高准确率为 87.61%, 谣言文本的召回率较高, 真实文本的精确度较高。由图 3-6 可知, 在训练了 14400 次后模型在训练集上的准确率稳定在 93.75%以上。

基于 Glove 模型进行实验:

首先,设置 CNN 网络超参数初始值,如表 3-4 所示。其次,将微博语料训练成词向量并输入 CNN 模型中进行训练。最后,设置不同维度的词向量对比实

第 23 页 共 55 页

验,调节参数使模型性能最优。

表 3-4 模型超参数初始值

dropout	激活函数	学习率	批处理大小	迭代次数
0.5	relu	0.001	64	20

基于 Glove 词向量的 CNN 网络实验数据,如表 3-5 所示。

表 3-5 Glove 词向量下 CNN 谣言识别性能表现

次000mm/11重上四位,周日次21年间次2					
评估指标 词向 量维度	Precision	Recall	F1-Measure	Accuracy	
50	0.8526	0.9182	0.8842	0.8741	
100	0.8550	0.9197	0.8862	0.8762	
150	0.8597	0.9225	0.8900	0.8784	
200	0.8630	0.9244	0.8927	0.8813	
250	0.8553	0.9262	0.8893	0.8854	
300	0.8578	0.9245	0.8899	0.8783	
350	0.8502	0.9298	0.8883	0.8765	

由实验数据可知,并不是向量维度越大模型效果越好,250 维词向量时 CNN 模型的性能最好,准确率为 88.54%,模型测试结果如图 3-7 所示。

```
For train_samples: step 14200, loss 0.196949, accuracy 0.9375

Testing loss: 0.386361, Testing accuracy: 0.87855

For train_samples: step 14300, loss 0.220484, accuracy 0.921875

Testing loss: 0.366786, Testing accuracy: 0.889324

For train_samples: step 14400, loss 0.0534106, accuracy 1

Testing loss: 0.376327, Testing accuracy: 0.886876

Testing loss: 0.357548, Testing accuracy: 0.885406

TN: 941 FN: 159 FP: 75 TP: 868

rumor

f1: 0.8893 precision: 0.8553 recall: 0.9262
```

图 3-7 glove 词向量 250 维时 CNN 网络各项评价指标结果

结合上述两个实验的数据来看, Glove 词向量模型的效果优于 Word2vec 词

第 24 页 共 55 页

向量模型,由于 CNN 网络中设置了多个卷积核,模型有较好特征提取能力,多次试验模型准确率都在 85%以上。

3.3 基于 LSTM 神经网络的微博谣言检测模型构建

3.3.1 循环神经网络

LSTM 模型是在 RNN 的基础上通过改进网络结构而来的。该模型的优点在处理时序信息时表现较好,能够建立之前时刻的输入与当前时刻输入的联系。例如在具体的情绪分析检测中,我们人判断短文本的情感时不是通过单一的词汇来判断的,而是联系整个句子的语境来判断,只有综合前后信息才能正确判断出文本的整体情绪倾向。RNN 之所以在时序问题上表现不错,是因为它特殊的网络结构。具体结构如图 3-8 所示:

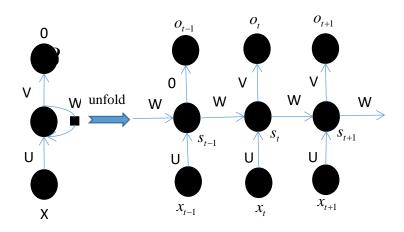


图 3-8 RNN 模型展开图

在展开图中,X 表示输入的词向量,S 是隐藏层的向量值,O 表示输出层输出的向量值。U 表示输入层到隐藏层的权值矩阵,W 表示隐藏层前一时刻的输出值作为当前时刻输入值的权值矩阵,V 表示从隐藏层的值到输出层值的权重矩阵。

网络中当前时刻隐藏层的值 St 由当前时刻的输入值 Xt 和前一时刻隐藏层的值共同决定。构建 RNN 网络输入的特征向量之间的关系。由 RNN 模型展开图可知,网络中各个时间都共享权值矩阵,因此,减少了训练参数数量。简单来说,就是不同的特征向量输入网络中执行相同的操作。由图 3-8 可知,在时间 t输出层计算公式如下:

$$O_t = g(Vs_t) \tag{3-6}$$

第 25 页 共 55 页

隐藏层的每个神经元都和输出层中的神经元相互连接,在时间 t,隐藏层和输入层间的计算关系如下:

$$S_{t} = f(Ux_{t} + Ws_{t-1}) (3-7)$$

式子中, f 表示激活函数。将公式 3-7 重复带入公式 3-6 中计算可知, 网络在 t 时刻的输出信息是由之前所有时刻的输入信息共同决定的。

我们用随时间反向传播算法[46]来训练 RNN 网络,算法计算过程如下:

- (1) 计算每个隐藏层向前的输出值;
- (2) 计算出每个神经元的误差项;
- (3) 通过梯度下降的算法重新计算权重。

现实中,RNN 在处理长序列信息时并不理想,训练中使用的链式求导会带来梯度爆炸和梯度消失。科学家针对该问题,提出了 LSTM 模型。LSTM 的解决思路就是在 RNN 的基础上添加门控结构。

3.3.2 LSTM 模型的门控结构

长短时记忆(LSTM)模型在标准的 RNN 模型上引入了记忆单元 cell,是具有门控结构的特殊神经网络,其单元结构如图 3-9 所示。

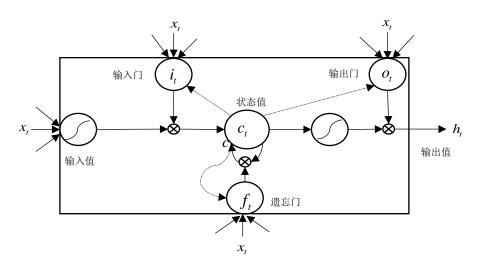


图 3-9 LSTM 结构单元

根据图 3-9 可知,一个 cell 中有三个"门":输入门 i、遗忘门 f 以及输出门 o。利用这些门控结构可以更新神经网络中的历史信息。遗忘门能让神经网络选择忘记部分无用的信息,由当前的输入和前一时刻的状态及输出等因素综合决定。输入们会在遗忘门遗忘部分信息之后,从输入信息中补充新的信息替

第 26 页 共 55 页

换遗忘部分的信息。输出门会在神经网络更新状态后根据当前时刻的状态、输入以及前一时刻的输出决定当前时刻的输出。若用函数 F 表示 LSTM 网络,则表达式为:

$$h_{t} = F(x_{t}, h_{t-1}) (3-8)$$

F 是由以下公式组成:

$$i_{t} = \sigma(w_{xi}x_{t} + w_{hi}h_{h-1} + w_{ci}c_{t-1} + b_{i})$$

$$f_{t} = \sigma(w_{xf}x_{t} + w_{hf}h_{h-1} + w_{cf}c_{t-1} + b_{f})$$

$$c_{t} = f_{t} \odot c_{t-1} + i_{t} \odot \tanh(w_{xc} + w_{hc}h_{t-1} + b_{c})$$

$$o_{t} = \sigma(w_{xo}x_{t} + w_{ho}h_{h-1} + w_{co}c_{t} + b_{o})$$

$$h_{t} = o_{t} \odot \tanh(c_{t})$$
(3-9)

3.3.3 LSTM 神经网络实验

与 CNN 网络实验相同,在本节实验中我们分别基于 Word2vec 和 Glove 模型生成词向量,选择 TensorFlow 框架进行试验。

(1) 基于 Word2vec 模型进行实验

首先,设置 LSTM 网络超参数初始值,如表 3-6 所示。其次,将微博语料训练成词向量并输入 LSTM 模型中进行训练。最后,设置不同维度的词向量对比实验,调节参数使模型性能最优。

 LSTM 堆叠层数
 dropout
 学习率
 批处理大小
 迭代次数

 2
 0.6
 0.001
 64
 20

表 3-6 模型超参数初始值

基于 Word2vec 词向量的 LSTM 网络实验数据,如表 3-7 所示。

次の / Word Zieo / 1/1 至 / Eb Tit 福音 / /// Ella / C//						
评估指标						
词向	Precision	Recall	F1-Measure	Accuracy		
量维度						
50	0.8104	0.9088	0.8568	0.8423		
100	0.8182	0.9131	0.8631	0.8491		
150	0.8218	0.9150	0.8659	0.8522		
200	0.8264	0.9174	0.8695	0.8561		
250	0.8248	0.9091	0.8649	0.8603		
300	0.8268	0.9102	0.8665	0.8619		
350	0.8290	0.9188	0.8716	0.8583		

表 3-7 word2vec 词向量下 LSTM 谣言识别性能表现

根据实验数据看到,在 300 维词向量时 LSTM 神经取得了最高准确率 86.19%,如图 3-10 所示。

```
For train_samples: step 14200, loss 0.388457, accuracy 0.875
Testing loss: 0.349699, Testing accuracy: 0.856023
For train_samples: step 14300, loss 0.244994, accuracy 0.890625
Testing loss: 0.33809, Testing accuracy: 0.864838
For train_samples: step 14400, loss 0.107714, accuracy 0.947368
Testing loss: 0.348663, Testing accuracy: 0.856023

Testing loss: 0.34434, Testing accuracy: 0.8619
TN: 916 FN: 192 FP: 90 TP: 845
rumor
f1: 0.8665 precision: 0.8268 recall: 0.9102
```

图 3-10 Word2vec 词向量 300 维时 CNN 网络各项评价指标结果

根据实验数据分析可知,LSTM 神经网络通过复杂的线性、非线性运算能较好的提取词向量中包含的谣言特征,且能在较高纬度的词向量下有不错的准确率,整体效果略差于 CNN 神经网络模型,但差距不大。

(2)基于 Glove 模型进行实验

首先,为 LSTM 神经网络设置超参数初始值,如表 3-8 所示。其次,将微博语料训练成词向量并输入 LSTM 模型中进行训练。最后,设置不同维度的词向量对比实验,调节参数使模型性能最优。

表 3-8 模型超参数初始值

LSTM 堆叠层数	dropout	学习率	批处理大小	迭代次数
2	0.6	0.001	64	20

基于 Glove 词向量的 LSTM 网络实验数据,如表 3-9 所示。

表 3-9 Glove 词向量下 LSTM 谣言识别性能表现

评估指标	评估指标					
词向	Precision	Recall	F1-Measure	Accuracy		
量维度						
50	0.8307	0.9196	0.8729	0.8641		
100	0.8340	0.9214	0.8755	0.8669		
150	0.8289	0.9187	0.8715	0.8693		
200	0.8352	0.9220	0.8765	0.8743		
250	0.8457	0.9209	0.8817	0.8775		
300	0.8476	0.9284	0.8862	0.8762		
350	0.8415	0.9253	0.8814	0.8711		

从实验数据可以看到,在 250 维词向量时模型的性能最佳,准确率达到最高 87.75%,模型测试结果如图 3-11 所示。

```
For train_samples: step 14200, loss 0.0883956, accuracy 0.984375
Testing loss: 0.388404, Testing accuracy: 0.87855
For train_samples: step 14300, loss 0.144912, accuracy 0.96875
Testing loss: 0.390982, Testing accuracy: 0.873653
For train_samples: step 14400, loss 0.0499061, accuracy 1
Testing loss: 0.389384, Testing accuracy: 0.873653

Testing loss: 0.397222, Testing accuracy: 0.877571
TN: 932 FN: 170 FP: 80 TP: 861
rumor
f1: 0.8817 precision: 0.8457 recall: 0.9209
```

图 3-11 Glove 词向量 250 维时 LSTM 网络各项评价指标结果

由实验数据可知,词向量模型直接影响实验结果,Glove 词向量下的模型准确率相对更高,准确率提高约 1%。

第 29 页 共 55 页

3.4 基于 CNN-LSTM 的微博谣言检测模型构建

3.4.1 CNN-LSTM 模型

CNN-LSTM 模型是一个融合模型,集合了两者模型的优势。和常见模型一样,首先将文本数据预处理后利用词向量技术生成对应向量。先利用的 CNN 的卷积层抽取局部特征,然后用池化层降低参数维度,接着进入 LSTM 层,再利用全连接层和 Softmax 层实现分类输出,CNN-LSTM 模型的结构如图 3-12 所示。

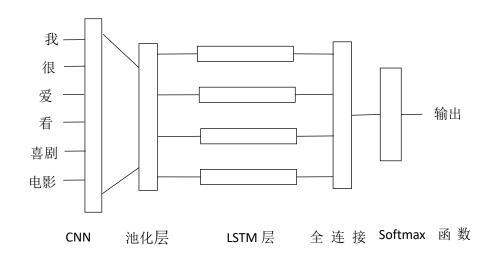


图 3-12 CNN-LSTM 模型结构图

3.4.2 CNN-LSTM 神经网络实验

本节采用 LSTM+CNN 结合模型对微博文本进行谣言检测。首先,CNN 网络中的卷积核用于在微博评论中抽取不同维度的特征。利用卷积层获取特征,利用池化层减少数据维度,将特征输入到 LSTM 层中,然后经过全连接层,最后由 softmax 层输出。

(1)基于 Word2vec 词向量进行实验

首先,为 CNN-LSTM 神经网络设置超参数的初始值,如表 3-10 所示。其次,将微博语料训练成词向量并输入 CNN-LSTM 模型中进行训练。最后,设置不同维度的词向量对比实验,调节参数使模型性能最优。

表 3-10 模型超参数初始值

LSTM 堆叠层数	dropout	学习率	批处理大小	迭代次数
2	0.5	0.001	64	20

基于 Wordvec 词向量的 CNN-LSTM 网络实验数据,如表 3-11 所示。

表 3-11 word2vec 词向量下 CNN-LSTM 谣言识别性能表现

•	• •			- • · · -
评估指标				
词向	Precision	Recall	F1-Measure	Accuracy
量维度				
50	0.8536	0.9253	0.8880	0.8782
100	0.8428	0.9193	0.8794	0.8731
150	0.8592	0.9314	0.8939	0.8882
200	0.8612	0.9295	0.8941	0.8903
250	0.8248	0.9091	0.8649	0.8918
300	0.8592	0.9314	0.8939	0.8845
350	0.8493	0.9293	0.8875	0.8776

由实验数据可知在词向量为 250 维时 CNN-LSTM 网络模型取得了最高准确率 89.18%,实验结果如图 3-13、3-14 所示。

```
For train_samples: step 14200, loss 0.18115, accuracy 0.90625
Testing loss: 0.412524, Testing accuracy: 0.891826
For train_samples: step 14300, loss 0.12199, accuracy 0.984375
Testing loss: 0.449948, Testing accuracy: 0.87861
For train_samples: step 14400, loss 0.0453402, accuracy 1
Testing loss: 0.422921, Testing accuracy: 0.889378

Testing loss: 0.450921, Testing accuracy: 0.891826
TN: 947 FN: 150 FP: 71 TP: 875
rumor
f1: 0.8956 precision: 0.8631 recall: 0.9305
```

图 3-13 Word2vec 词向量 250 维时 CNN-LSTM 网络各项评价指标结果

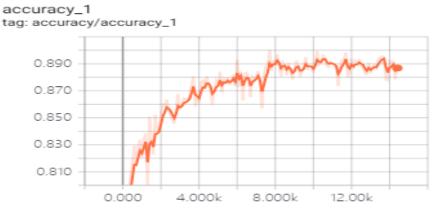


图 3-14 250 维 word2vec 词向量下 CNN-LSTM 训练中准确率变化

在 word2vec 下进行实验,三种模型中 CNN-LSTM 的性能最好,模型准确率比 CNN 提高 1.5%,比 LSTM 提高 3%,同时训练时损失函数收敛较快。

(2) 基于 Glove 词向量进行实验

首先,为 LSTM+CNN 网络设置超参数的初始值,如表 3-12 所示。其次,将微博语料训练成词向量并输入 CNN-LSTM 模型中进行训练。最后,设置不同维度的词向量对比实验,调节参数使模型性能最优。

表 5 12 侯主起多数仍和臣					
LSTM 堆叠层数	dropout	学习率	批处理大小	迭代次数	
2	0.5	0.001	64	20	

表 3-12 模型超参数初始值

基于 Glove 词向量的 LSTM+CNN 网络实验数据,如表 3-13 所示。

	表 3-13 Glove 词问重	r CNN-LSTM	谣言识别性能表现	
评估标准				
词向	Precision	Recall	F1-Measure	Accuracy
量维度				
50	0.8570	0.9333	0.8935	0.8841
100	0.8628	0.9362	0.8980	0.8962
150	0.8736	0.9416	0.9064	0.9013
200	0.8864	0.9431	0.9139	0.9061
250	0.8902	0.9451	0.9168	0.9138
300	0.8788	0.9391	0.9080	0.8964
350	0.8627	0.9303	0.8953	0.8934

表 3-13 Glove 词向量下 CNN-LSTM 谣言识别性能表现

第 32 页 共 55 页

由表 3-13 可知, 在词向量为 250 维时 CNN-LSTM 模型取得最高准确率 91.38%, 实验结果如图 3-15、图 3-16 所示。

```
For train_samples: step 14200, loss 0.12664, accuracy 0.984375
Testing loss: 0.369728, Testing accuracy: 0.906954
For train_samples: step 14300, loss 0.0434393, accuracy 0.984375
Testing loss: 0.348329, Testing accuracy: 0.907444
For train_samples: step 14400, loss 0.166378, accuracy 0.947368
Testing loss: 0.381476, Testing accuracy: 0.903036
Testing loss: 0.350989, Testing accuracy: 0.91381
TN: 971 FN: 120 FP: 56 TP: 896
rumor
f1: 0.9168 precision: 0.8902 recall: 0.9451
```

图 3-15 Word2vec 词向量 250 维时 CNN-LSTM 网络各项评价指标结果



图 3-16 250 维 glove 词向量下 CNN-LSTM 训练中准确率变化

在 LSTM+CNN 网络实验中, Glove 词向量的效果要好于 Word2vec 词向 量。训练中模型前期收敛较快,后期准确率有一定的波动。

3.5 实验对比与分析

(1) 词向量对比

为了使模型的分类性能最佳[12],分别设置了基于 Word2vec 和 Glove 的对比 实验。根据 3.2 和 3.3 小节的数据可知,各模型基本在 250 维词向量时取得了最 高准确率, 因此选择 250 维时各模型的数据进行对比, 结果如表 3-14 所示。

第 33 页 共 55 页

1 J-1+ L	万土工工23	で作りり主	
模型	CNN	LSTM	CNN-LSTM
词向量			
Word2vec	87.61%	86.19%	89.18%
Glove	88.54%	87.75%	91.38%

表 3-14 各模型在 250 维词向量下的性能表现

各模型在 250 维词向量下的实验对比,如图 3-17 所示。

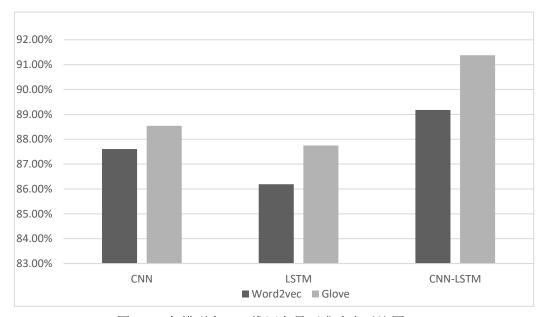


图 3-17 各模型在 250 维词向量下准确率对比图

比较表 3-14 和图 3-17, Glove 单词向量的效果略好于每个模型下的 Word2vec 单词向量。其中原因,主要是因为 Word2vec 是根据上下文窗口预测 单词,而 Glove 是根据全局信息来预测。

(2) Word2vec 词向量下结果对比

本章实验的三种深度学习模型在 Word2vec 词向量下的准确率对比,如图 3-18 所示。

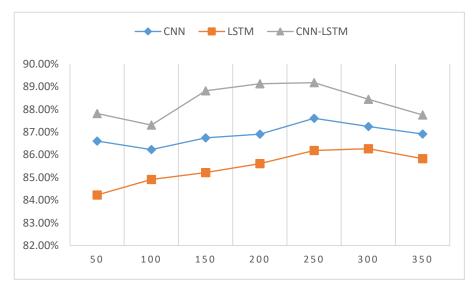


图 3-18 word2vec 词向量下各模型的准确率对比

CNN-LSTM 模型整体表现最优,CNN 模型比 LSTM 模型表现稍好一些。 分析其中的原因,通过合理的参数设置,CNN-LSTM 组合模型分别保留了 CNN 模型的局部特征提取能力和 LSTM 解决长序依赖问题的优势,因此表现最优。

(3) Glove 词向量下结果对比

本章实验的三种深度学习模型在 Glove 词向量下的准确率对比,如图 3-19 所示。

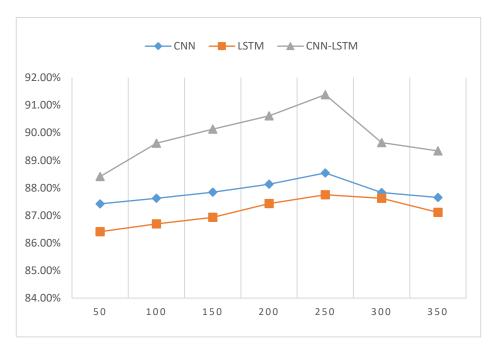


图 3-19 不同模型在 Glove 词向量下准确率对比

CNN-LSTM 模型整体好于 CNN 模型和 LSTM 模型, 在词向量为 250 维度

第 35 页 共 55 页

时最为明显。LSTM 模型表现稍逊与 CNN 和 CNN-LSTM。

3.6 本章小结

本章讲述了三种神经网络模型 LSTM、CNN 和 CNN-LSTM, 剖析这三种模 型的结构和原理。根据第二章介绍的预处理方法构造了训练集和测试集,并训 练了 Word2vec 和 glove 两种词向量模型。词向量模型和词向量维度为区分条件, 设置各种对比实验,比较了不同模型的检测表现和不同词向量模型对实验结果 的影响。Glove 词向量模型整体效果好于 word2vec 模型,且各个模型基本都在 250 维词向量下趋于稳定。在 glove 词向量下, CNN 模型的准确率最高为 88.54%, LSTM模型的准确率最高为87.75, CNN-LSTM组合模型准确率最高为91.38%。 实验证明,本章设计的 CNN-LSTM 模型有良好的谣言检测效果。

4 基于情感分析和 CNN-LSTM 模型的微博谣言检测

4.1 基于情感分析和 CNN-LSTM 的谣言检测方法

在第三章中,重点研究了基于神经网络模型的谣言检测,但是仅靠神经网络难以提取更深层次的谣言特征,比如微博评论的情感特征。因此,本章重点研究评论的情感特征,通过分析大量微博评论数据的情感倾向,根据评论的正负倾向分布比例和情感方差设计了基于评论情感的谣言检测模型。然后,将基于情感的谣言检测模型和基于 CNN-LSTM 的谣言检测模型相结合,构建融合情感极性的 CNN-LSTM 谣言检测模型,完成微博谣言检测。为了有效的挖掘评论情感特征,合理的文本预处理和精确的评论情感倾向判断至关重要,整体方法的具体流程如图 4-1 所示。

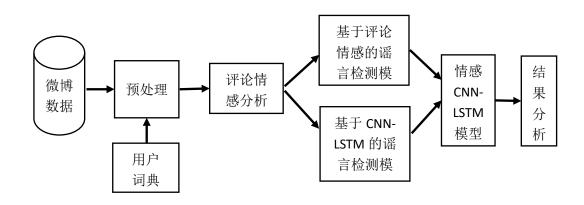


图 4-1 基于情感分析和 CNN-LSTM 的谣言检测流程图

4.1.1 文本预处理

爬取收集相关微博数据并建立语料库,对于采集的数据首先要采取分词、 去特殊字符以及垃圾信息过滤等操作,过滤方法一般采用正则化。本文将评论 情感作为一个重要的特征,分词的好坏直接影响句子情感倾向性判断,因此分 词至关重要。为了更好的实验效果,我们通过构建用户词典来提高分词工具的 分词准确率。

基于以上问题,为了设计出更好的分类模型,提高模型对情感词的适应性,本文加入情感词典,形成用户词典。

用户词典构成如下所示:

(1) 知网 Hownet 词典

第 37 页 共 55 页

- (2) 台湾大学 NTUSD 词典
- (3) 褒贬义词典(清华大学李军)
- (3) 大连理工中文情感词汇库

本文章的情感词典主要基于大连理工的情感词汇库,并以NTUSD词典、知网词典和褒贬义词典为辅。

4.1.2 微博评论的情感分析

情感分析是对文本中包含的情感词进行挖掘,并提取特征判别文本的情感倾向性。一般将文本分成正负两类或者正负中性三类。对微博文本情感分析可以挖掘谣言文本和非谣言文本的情感差异,作为谣言检测的特征。当前,情感分析方法可分为三大类:情感词典方法、深度学习方法、机器学习方法。

情感词典就是利用统计规则构建词典并将处理过的文本结合,看文本是否有情感词典中的词汇来判断文本情感。分别统计积极和消极情感词出现的频率,加权计算文本的整体情感倾向。其计算公式如下:

$$sentiment_{text} = \begin{cases} positive, 正向的情感词多 \\ neutrality, 正负情感词等量 \\ negative, 负向的情感词多 \end{cases}$$
 (4-1)

机器学习的方法也是常见的情感判别方法,情感分析问题可归为分类问题。 即将文本分为正负两类。机器学习分类效率高准确率也不错,但多数算法都是 监督学习算法,语料依赖人工标注。

深度学习是最新的而技术,在早期的传统神经网络基础上发展而来,逐渐变成当下的研究热点。CNN 和 RNN 是使用最多最常见的两种模型,不需要人工,通过训练模型,自动提取文本特征,但是需要海量的训练数据。

微博具有海量数据,因此采用深度学习算法进行情绪分析,具体分析流程如图 4-2 所示。其步骤为:

- (1) 下载 COAE2014 任务 4 微博数据集
- (2) 将文本信息转换成词向量;
- (3)使用训练后的文本词向量作为神经网络的输入,不断优化参数,训练模型,构建情感分类器。
 - (4) 利用神经网络情感分析模型,计算评论的情感倾向并对其进行标注。

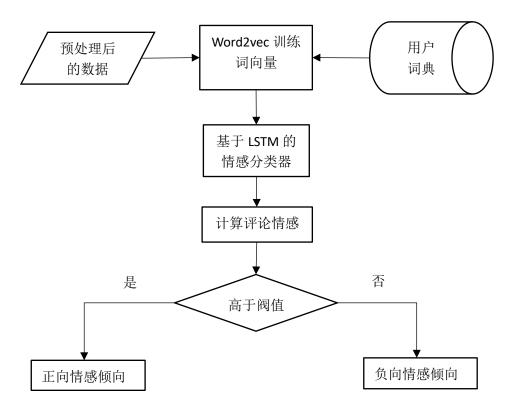


图 4-2 微博评论情感分析流程图

4.2 基于评论情感的谣言检测

本文将使用第3章介绍的LSTM建基于LSTM的情感分类模型。通过挖掘微博评论的情感特征,实现了基于评论特征的谣言检测。

4.2.1 基于 LSTM 的情感分析

本文将基于 LSTM 模型分析情感。one-hot 编码生成的向量维度太高,参数矩阵十分稀疏,容易出现词义丢失现象。Word2vec 词向量方法很好的避免了以上问题,因此文本中使用 Word2vec 训练的词向量作为输入,构建的基于 LSTM 网络的情感分析模型如图 4-3 所示。

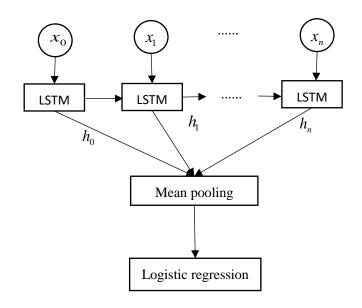


图 4-3 基于 LSTM 网络的情感分析模型

相关算法介绍如下:

第一步:词向量的训练

- (1) 清洗语料库并将词向量模型初始化。
- (2) 将语料信息输入模型, 计算目标词条件概率。
- (3) 算出目标词汇最大化概率。
- (4) 预测目标词汇,并输出词向量。

第二步:模型训练

- (1) 设置 LSTM 网络初始化超参数
- (2) 生成文本向量集合
- (3) 将向量输入模型
- (4) 给输出样本打上情感标签
- (5) 确定损失函数
- (6) 保存模型,导出模型参数

第三步:测试

- (1) 生成测试文本的词向量集合
- (2) 将测试文本的向量集合输入模型,生成测试文本情感标签;
- (3) 输出模型结果。

4.2.2 评论情感特征的挖掘

与前面涉及的其他功能不同,微博的整体评论情绪倾向并不容易获得,每个评论的情绪倾向要加以区分。无论情绪强度如何,文本都将划分为正负两类。

微博语言自由化,适合采用简单有效的[56]深度学习分类方法;也就是说,

第 40 页 共 55 页

构建基于 LSTM 的情感分析模型,并且在训练集上训练 LSTM 情感分类器,并且对所有评论执行情感趋势分类。具体步骤如下(评论数据的预处理已在 3.1.2 小节中完成):

(1) 训练集标注。本文通过 LSTM 情感分类器标注微博语料库中谣言微博和非谣言微博两部分的评论数据情感倾向性。最终标注集情况如表 4-3 所示。

表 4-3 评论标注情况

	正	负
谣言评论	33.9%	66.1%
非谣言评论	48.3%	51.7%
所有评论	41.1%	58.9%

(2) 根据评论文本情感值, 计算整体评论情感强度, 计算公式为:

$$text_{sentiment} = \frac{n_{pos} - n_{neg}}{N}$$
 (4-2)

其中 n 表示微博评论中具有积极(消极)情感倾向的评论数,评论的总数用 n 表示。可以看出,当特征值是区间[-1,1]中的值时,当它大于 0 时,这表明微博引起的一般观众情绪是积极的,反之则代表消极。绝对值越大,微博引起的听者的积极(消极)情绪就越普遍。

(3) 特征集描述

VIP: 表示经过新浪微博认证过的用户

POS: 评论的正向情感值

NEG: 评论的负向情感值

text_{sentiment}: 整体评论情感强度

sentiment_{variance}: 所有和微博相关的评论的情感方差值

4.2.3 基于评论情感的谣言检测模型

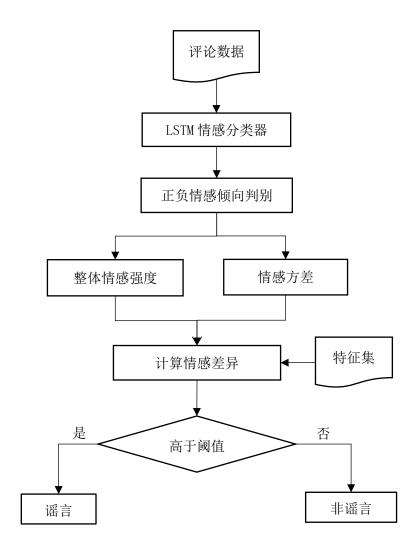


图 4-4 基于评论情感的谣言检测模型

4.3 融合情感极性的 CNN-LSTM 的谣言检测模型

4.3.1 模型的构建

融合情感极性的 CNN-LSTM 谣言检测模型的具体工作流程如下:

- (1) 构建微博语料库,数据集分为谣言微博和正常微博。
- (2)清洗数据,过滤垃圾评论、分词;训练词向量模型来向量化微博文本。
- (3) 微博文本的情感分析。计算微博评论的情感方差,其公式为:

$$sentiment_{\text{variance}} = \frac{1}{n} \sum_{i=1}^{n} \left(text_{sentiment} - s \right)^{2}$$
 (4-3)

第 42 页 共 55 页

(4)构建微博评论的情感特征集,重点根据评论情感强度特征和情感方差特征来计算测试集微博的情感差异值,计算情感差异值公式如下:

$$sentiment_{diff} = \alpha |text_{sentiment} - a| + \beta |sentiment_{variance} - b|$$
 (4-4)

- (5) 训练基于 CNN-LSTM 的谣言识别模型,并预测文本为谣言(和非谣言)的概率值。
- (6) 将情感差异值和模型预测结果代入权重公式,计算结果大于阀值则为谣言,小于则为非谣言。rumor_{predict}表示谣言检测值, λ 、 γ 为权重参数,分别为 0.4、0.6,则谣言计算公式为:

$$rumor_{predict} = \lambda * sentiment_{diff} + \gamma * cnn - lstm_{predict}$$
 (4-5)

设定谣言检测阈值为 k,若 f(w)>k,则判断为谣言;若 f(w)<k 时,则判断为非谣言。

模型框架图如下图 4-5 所示:

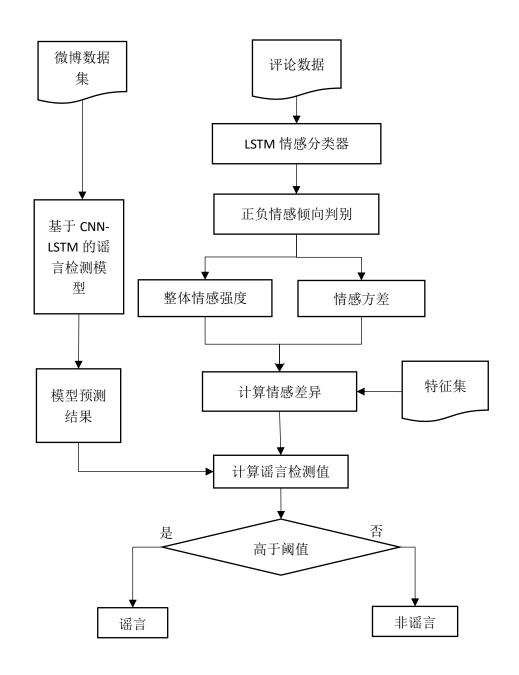


图 4-5 融合情感极性的 CNN-LSTM 谣言检测模型

4.3.2 实验设计

根据 4.3.1 小节中的方法,将疑似谣言案例的微博经过清洗和向量化后输入 CNN-LSTM 模型中,预测该案例为谣言的可信度。同时采集该案例微博下的评论,利用评论的情感特征判断该话题为谣言的可信度。最后输入两个模型的可信度预测值,根据谣言检测公式判断是否为谣言。

利用流言百科提供的谣言话题和微博辟谣平台提供的谣言微博, 采集 300

第 44 页 共 55 页

条疑似谣言案例的微博和每条微博相关的500条评论,作为本章实验的测试集。

4.4 实验结果与对比分析

4.4.1 谣言检测结果

谣言检测阈值取 0.3 时,得到的谣言检测结果如表 4-1 所示。

表	4-1 谣言检测结果	
	实际为谣言	实际为非谣言
被分类为谣言	182	10
被分类为非谣言	12	96

本章实验测试集微博共计 300 条,实际和预测都为谣言 182 条,实际和预 测均不是谣言的96条,其中22条微博判断错误。谣言检测阈值设定为0.3,各 项评价标准的实验结果如表 4-2 所示。

PC : = 1137	L. CI (I C 20 II) I	DC 13 11 //	1 N 1 1 H 1 1 1 1 1 1 2 2 2 2 2 2 1 1	•	
评估指标模型	Precision	Recall	F1-Measure	Accuracy	
融合情感极性的 CNN-LSTM	0.9479	0.9192	0.9333	0.9266	
CIVIN-LS I WI					

表 4-2 情感 CNN-LSTM 模型的各项评价指标实验结果

将谣言检测遇到的情景可分为三类, 典型判断正确的、典型判断错误的和 存在争议的。

(1) 典型判断正确案例

判断正确,即模型预测结果与实际人工鉴定一致。例如"日本人经常吃所以 三文鱼体内没有寄生虫"微博案例, CNN-LSTM 神经网络判断为谣言, 该话题下 评论负面情感较多,且情感差异较大,最终谣言检测值为 0.42>0.3,故综合判断 该微博话题为谣言。容易检测的微博案例,一般具有话题热、讨论多、情感倾向 明显等特点。对于情感倾向明显的真实信息微博,如 "xx 地方发生五级地震" 信息,可以根据评论情感方差特征判断。

(2) 典型判断错误案例

对于一些符合群众认知的略偏冷门的话题往往微博相关数据量不大且话题 下的评论情感相对一致,情感差异不大,往往容易形成误判。如:"专门研制的 '儿童酱油'有益健康,增强儿童食欲"谣言微博,可能受广告影响符合人们

第 45 页 共 55 页

认知,且相关微博数量不大。CNN-LSTM 神经网络判断为真实,该话题下评论 正面情感较多,且情感差异较小,最终谣言检测值为 0.22<0.3。故模型综合判断 该微博案例为真实,判断错误。

(3) 存在争议案例

例如"转基因不安全,长期吃转基因食品会影响生育"微博案例,虽然讨论很激烈,评论的情感方差也很大,符合谣言的情感特征。但是由于,转基因食品安全是一个学术问题,目前在学术上和临床还未论证所有副作用,并不能给出一个准确判断。对于这些无法鉴定的微博信息,我们直接将其从测试集中删除,不做判断。

4.4.2 实验对比分析

(1) 谣言检测阈值对实验结果的影响

谣言检测阈值的大小决定的模型的好坏,为了设置合理的阈值,本文设置对比实验,将谣言检测阈值在[0.05,0.8]区间类划分成 16 档进行实验,实验结果如图 4-6 所示。

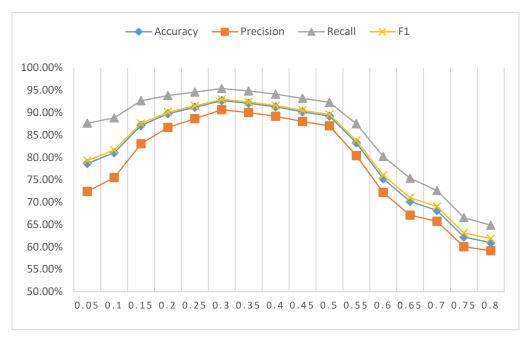


图 4-6 不同谣言检测阈值对实验各项指标的影响

谣言检测阈值的大小和模型的准确率高度相关,谣言检测阈值过高,模型容易将谣言误判为真实,谣言检测阈值过低,模型容易将真实误判为谣言。过高或过低的谣言检测阈值下,实验结果都不太理想。模型在谣言检测阈值取 0.3 时准确率达到最高 92.66%。

(2) 加入情感特征对实验的影响

微博评论中包含大量的情感特征,而神经网络对于这种深层特征的提取表现不佳。因此,我们通过 4.2 小节设计的情感特征提取方法,将情感特征加入到 CNN-LSTM 模型中,验证加入情感特征对模型检测结果的影响。选用 Glove 模型,词向量为 250 维,谣言检测阈值为 0.3 进行实验,结果如图 4-7 所示。

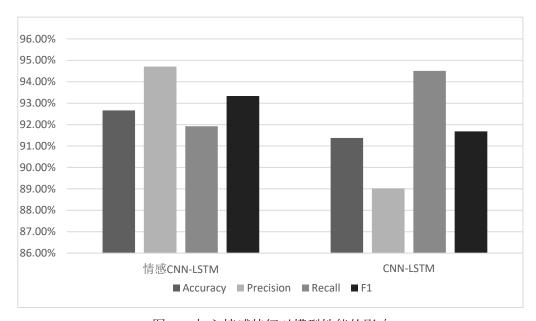


图 4-7 加入情感特征对模型性能的影响

根据图 4-7 实验结果可知,将情感特征加入模型后,谣言检测的准确率提升 1.28%,精确度提升 5.69%,F1 值提升了 1.65%,召回率稍有下降。实验数据证明,加入评论情感特征可以提高模型的谣言检测效果。

(3) 情感 CNN-LSTM 模型与其他模型的实验对比

为了验证本文所构建的融合情感极性的 CNN-LSTM 模型在微博谣言检测任务上的有效性,本文将情感 CNN-LSTM 模型与传统的机器学习方法如 NaiveBayes、SVM 分类器模型以及常见的深度学习模型进行对比实验。基于本文的数据集,词向量模型为 Glove,词向量维度为 250 维,谣言检测阈值为 0.3,实验结果如表 4-3 和图 4-8 所示。

	- 11404	**************************************	14614: 4:-	
评估指标 模型	Precision	Recall	F1-Measure	Accuracy
NaiveBayes	0.7428	0.7995	0.7701	0.7631
SVM	0.7737	0.8252	0.7986	0.7922
GRU	0.8636	0.9047	0.8837	0.8796
CNN	0.8553	0.9262	0.8893	0.8854
LSTM	0.8457	0.9209	0.8817	0.8775
CNN-LSTM	0.8902	0.9451	0.9168	0.9138
情感 CNN-LSTM	0.9479	0.9192	0.9333	0.9266

表 4-3 不同模型间的各项评价指标对比

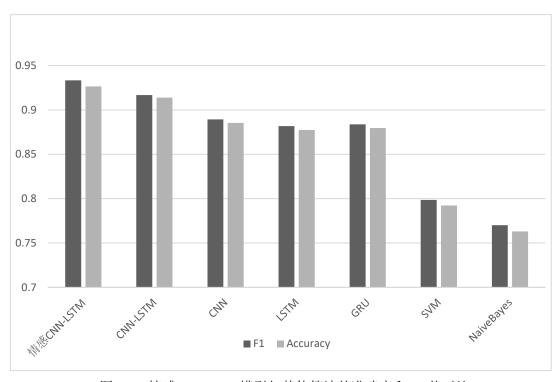


图 4-8 情感 CNN-LSTM 模型与其他算法的准确率和 F1 值对比

由图 4-8 可知,在以上七种模型的对比实验中,本文所构建的情感 CNN-LSTM 模型的微博谣言检测效果最好。基于深度学习的方法明显好于传统的机器学习方法,在深度学习方法中 CNN 模型、LSTM 模型以及 GRU 模型三者的谣言分类性能相近,而 CNN-LSTM 模型的表现均稍好于这三类模型。分析原因,主要是因为深度学习对大数据有更好的学习效果;而在深度学习模型中 CNN-LSTM 组合模型既保留 CNN 的卷积层池化层,有较好的局部特征提取能力,同时 LSTM 替换了 CNN 的全连接层,使该网络具有 LSTM 适合解决长序依赖问题的特性。最后,将评论情感特征融入 CNN-LSTM 模型,进一步提高了模型的

谣言检测性能。通过对比实验,证明了本文提出的情感 CNN-LSTM 模型的有效性。

4.5 本章小结

本章通过各种对比实验,检验了本文提出的谣言检测模型的有效性。首先对谣言检测的方法从文本预处理、评论情感分析和 CNN-LSTM 模型三个方面做了基本介绍。然后设计了 LSTM 情感分析模型,挖掘了微博评论的情感特征,并基于情感特征设计谣言检测模型。最后,将情感特征和 CNN-LSTM 模型相结合,设计了融合情感极性的 CNN-LSTM 模型。通过设置对比实验,选取合适的谣言检测阈值提高模型的分类效果。接着设置加入情感特征和没加情感特征的对比实验,证明了情感特征的加入能明显提升模型分类性能。实验最后,设置不同模型在谣言检测任务上的对比实现,结果表明本文提出的模型分类效果都优于其他模型,证明了融合情感极性的 CNN-LSTM 模型在微博谣言检测任务上是有效的。

5 总结与展望

5.1 总结

互联网还在不断发展,人们的生活还在继续网络化,网络社交成为生活不可或缺的一部分。微博作为社交平台的领跑者,会继续提供信息获取和分享服务,微博的海量信息会继续服务于我们生活的方方面面。微博不真实信息的发现、预防,是继续提供高质量社交信息服务的前提。我们要继续关注谣言的检测,未来的微博平台信息将更加繁杂更加庞大。

本文总结了微博情感分析和谣言检测的发展历史和现状,探讨不同的情感极性判别方法,如情感词典、机器学习方法以及深度学习算法等。目前将深度学习应用于谣言检测是一个不错的尝试,实验证明深度学习方法效果不错。为了进一步提升深度学习模型的谣言分类效果,我们设计了融合情感极性的CNN-LSTM模型,并通过设置对比实验证明了该方法的有效性。

5.2 展望

本文构建了谣言检测方法,利用模型融合方法和 CNN-LSTM 模型充分发挥了深度学习在分类问题上的优势,最终取得了很好的实验结果。但一个人的力量毕竟有限,从微博上搜集的语料也有待进一步扩展,我的计算机硬件目前还不能进行更大规模的数据训练。总之,这次研究还有诸多不足,有待后面进一步改进:

- (1)本文章的研究只是结合了深度学习和微博情感特征,但忽略了谣言信息来源可信度和谣言传播方面的研究,可以继续扩大特征集。
- (2)本文检测的是特定领域内的微博不实信息,怎样基于现有的研究成果设计一个新的模型使之适应更多领域的不真实信息检测。
- (3)目前我的研究都是基于过去的信息进行的,也就是说分析的都是静态的数据,无法实时的做出判断。但不真实信息的产生是时时刻刻都在发生,如何基于动态数据实时的检测谣言使得不实信息扼杀在萌芽状态。这将是谣言检测研究以后值得努力的方向,也是谣言检测取得重大突破的标志。

参考文献

- [1] 新浪 2017Q3 微博财报统计数据[EB/OL]. [2017-12-26].
- [2] 新媒体蓝皮书: 中国新媒体发展报告(2015)[M]. 北京:社会科学文献出版社, 2015 (07): 1.
- [3] 唐绪军,吴信训,黄楚新等. 中国新媒体发展报告 No.5(2014)[M]. 2014.
- [4] Pang B, Lee L, Uaithyanathan S. Thumbs Learning Techniques[C]. IN PROCEEDINGS up? OF Sentiment Classification using Machine 2002.
- [5] 任远,巢文涵,周庆. 基于话题自适应的中文微博情感分析[J]. 计算机科学, 2013, 40(11): 31-235.
- [6] Ekman P.Emotions revealed: Recognizing faces and feelings to improve communication and emotional life[J]. Holt Paperback, 2007, 128(8): 140-140.
- [7] 朱嫣岚, 闵锦, 周雅倩. 基于 HowNet 的词汇语义倾向计算[J]. 中文信息学报, 2006, 20(1): 16-22.
- [8] 王素格,李德玉,魏英杰.基于同义词的词汇情感倾向判别方法[J].中文信息学报,2009,23(5):68-74.
- [9] 桂斌,杨小平,朱建林.基于意群划分的中文微博情感倾向分析研究[J].中文信息学报,2015,29(3):100-105.
- [10] Pang B, Lee L. Opinion mining and sentiment analysis[J]. Foundations and trends in information retrieval. 2008, 2(1-2): 1-135.
- [11] Davidov D, Tsur O, Rappoport A. Enhanced Sentiment Learning Using Twitter Hashtags and Smileys[C]//International Conference on Computational Linguistics: Posters. 2010:241-249. Turney P D. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews[J]. 2002: 8.
- [12]Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision [J]. CS224N Project Report, Stanford, 2009:1-12.
- [13] Huang P S, He X, Gao J, et al. Learning deep structured semantic models for web search using clickthrough data[C]//Proceedings of the 22nd ACM international conference on Conference on information&knowledge management. ACM, 2013: 2333-2338.
- [14] 唐慧丰, 谭松波, 程学旗. 基于监督学习的中文情感分类技术比较研究[J]. 中文信息学报, 2007, 21(6): 88-94.
- [15]谢丽星,周明,孙茂松.基于层次结构的多策略中文微博情感分析和特征抽取[J].中文信息学报,2012,26(1):73-84.

- [16]孙建旺, 吕学强, 张雷瀚. 基于词典与机器学习的中文微博情感分析研究[J]. 计算机应用与软件, 2014, 31(7): 177-181.
- [17]李文,苗夺谦,卫志华.基于阻塞先验知识的文本层次分类模型[J].模式识别与人工智能,2010,23(4):456-463.
- [18] Dai A M. and Le Q V. Semi-supervised sequence learning[C]// NIPS: Conference on Neural Information Processing Systems. Montreal, Quebec, 2015:5949-5958.
- [19] Johnson R and Zhang T. Semi-supervised convolutional neural networks for text categorization via region tembedding [C]// NIPS: Conference on Neural Information Processing Systems. Montreal, Quebec, 2015:5849-5858.
- [20] Mintz M, Bills S, Snow R, et al. Distant supervision for relation extraction without labeled data[C]//Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume2-Volume2. Association for Computational Linguistics, 2009: 1003-1011.
- [21] Tatemura J. Virtual reviewers for collaborative exploration of movie reviews [C]// Proceedings of the 5th international conference on Intelligent user interfaces. ACM, 2000: 272-275.
- [22] Le Q V, Mikolov T. Distributed representations of sentences and documents[J]. arXiv preprint arXiv: 1405. 4053, 2014.
- [23] Turney P D, Littman M L. Measuring praise and criticism: Inference of semantic orientation from association[J]. ACM Transactions on Information Systems (TOIS), 2003, 21(4): 315-346.
- [24] Hasegawa T, Sekine S, Grishman R. Discovering relations among named entities from large corpora[C]//Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004: 415.
- [25] Qazvinian V, Rosengren E, Radev D R, et al, Rumor has it: identifying misinformation in microblogs[C]// Conference on Empirical Methods in Natural Language Processing, EMNLP 2011, 27-31 July 2011, John Mcintyre Conference Centre, Edinburgh, Uk, A Meeting of Sigdat, A Special Interest Group of the ACL. DBLP, 2011: 1589-1599.
- [26] Castillo C, Mendoza M,Poblete B..Information credibility on twitter[C]// International Conference on World Wide Web, WWW2011,Hyderabad, India, March 28-April. DBLP; 2011:675-684.
- [27] Kwon, S., Cha, M., Jung, K., Chen, W., Wang, Y:Prominent features of rumor 第 52 页 共 55 页

- propagation in online social media. In: 2013 IEEE 13th International Conference on Data Mining (ICDM), pp. 1103-1108 (2013).
- [28]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.
- [29]路同强, 石冰, 闰中敏, 等. 一种用于微博谣言检测的半监督学习算法[J]. 计算机应用研究, 2016, 33 (3): 744-748.
- [30]毛二松,陈刚,刘欣,等. 基于深层特征和集成分类器的微博谣言检测研究 [J]. 计算机应用研究,2016,33 (11):3369-3373.
- [31]陈彦舟,曹金漩.基于 Hadoop 的微博舆情监控系统[J]. 计算机系统应用, 2013 22 (4): 18-22.
- [32] Yang F, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo[C]//ACM, 2012: 1-7.
- [33]S. Sun, H. Liu, J. He, and X. Du, Detecting event rumors on sina weibo automatically, in Web Technologies and Applications. Springer, 2013, pp. 120-131.
- [34]乔春庚,孙丽华,吴韶,等. 基于模式的中文倾向性分析研究[J]. 第一届中文倾向性分析评测研讨会. 2008: 21-31.
- [35] Takahashi T; Igata N.Rumor detection on twitter[C]// Joint, International Conference on Soft Computing and Intelligent Systems. 2012:452-457.
- [36] Nourbakhsh A, Liu X, Shah S, et al. Newsworthy Rumor Events: A Case Study of Twitter[C]// IEEE International Conference on Data Mining Workshop. IEEE, 2015:27-32.
- [37]赵金楼, 成俊会. 基于 SNA 的突发事件微博與情传播网络结构分析—以"4. 20 四川雅安地震"为例[J]. 管理评论, 2015, 27 (1): 148.
- [38] Yu S. A Credibility Assessment for Message Streams on Microblogs[C]l/ International Conference on P2p, Parallel, Grid, Cloud and Internet Computing. IEEE Computer Society, 2010:527-530.
- [39] Gupta A, Kumaraguru P Credibility ranking of tweets during high impact events [C]//2012: 2-8.
- [40] 卞先华,陈亮,郑倩冰.基于文本内容和社会结构的可信度[J].重庆理工大学 学报自然科学版,2013,27(1):57-61.
- [41]吴树芳, 徐建民.基于 HITS 算法的微博用户可信度评估[J]. 山东大学学报 江学版), 2016, 46 (5): 7-12.
- [42] 张彬. 面向中文网络信息检索的自动分词系统设计与算法实现[D]. 华东师范大学,2007.

- [43]周顺先, 蒋励, 林霜巧, 等. 基于 Word2vector 的文本特征化表示方法[J]. 重庆邮电大学学报(自然科学版), 2018, 30(2): 272-279.
- [44]陈珍锐,丁治明. 基于 GloVe 模型的词向量改进方法[J]. 计算机系统应用, 28(1).
- [45] HATZIVASSILOGLOU V, MCKEOWN K R. Predicting the Semantic Orientation of Adjectives[C]. The 8th Conference on European Chapter of the Assoc iat ion for Computat ional Linguistics, Madrid, Spain, 1997: 174-181.
- [46] WIEBE J, WILSON T, BRUCER, et al. Learning Subjective Language[J]. Computational linguistics. 2004, 30(3): 277-308.
- [47] HOVY E. Generating Natural Language Under Pragmatic Constraints[J]. Journal of Pragmatics. 1987, 11(6): 689-719.
- [48] 叶强, 张紫琼, 罗振雄. 面问互联网评论情感分析的中文主观性自动判别方法研究[J]. 信息 系统学报. 2007, (01): 79-91.
- [49] Hubel DH, Wiesel T N. Receptive fields and functional architecture of monkey striate cortex[J]. The Journal of physiology, 1968, 195(1): 215-243.
- [50] Kim Y. Convolutional Neural Networks for Sentence Classification[C]//
 Proceedings of Empirical Methods in Natural Language Processing, 2014:17461751
- [51] Sathasivam S.Logic Learning in Hopfield Networks[J]. Modern Applied Science, 2009.
- [52] Zhu X Sobihani P Guo H.Long short-teen memory over recursive structures[C]. In:Proceedings of the 32nd International Gonference an Machine Learning (IGML15) 2015: 1604-1612.
- [53] Fukushima K, Miyake S,Ito T. Neocognitron: A neural network model for a mechanism of visual pattern recognition[J]. IEEE Transactions on Systems, Man and Cybernetics, 1983 (5): 826-834.
- [54] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT Press, 2016.

致谢

时间如白驹过隙,两年的研究生学习转瞬即逝。在新疆师范大学读研期间,我收获了很多,不仅习得了专业技能,更懂得了很多为人处事的道理。在师大的两年,是我最难忘、最充实和成长最多的两年。由衷的感谢在我读研期间出现在我生活中的每一个亲人、朋友和老师,谢谢你们在我学习和生活中对我鼓励和帮助。

首先,真心的感谢我的导师张海军教授在我两年的研究生生活中对我的科研指导和帮助。张老师平易近人,就像我在新疆的亲人一样,为我创造了温馨融洽的科研环境。从研究课题的选定到研究深入分析和论文撰写都得到了张老师的指导与帮助,不仅使我在研究方向打开了思路,提高了科研水平,更让我对自然语言处理相关知识有了更深次的理解。张老师学识渊博、治学态度严谨,是我科研路上的榜样。我会谨记张老师的教诲,继续我的学习之路。

感谢潘伟民、陈炳才、年梅等老师两年来在我的专业学习上给予的帮助和指导;感谢学校提供的良好的实验环境;感谢实验室的师兄师姐、师弟师妹给予我的大大小小的帮助;特别感谢我的室友,我们曾在宿舍为一个问题讨论几个小时,也曾为一行行代码激动不已。因为你们的出现,我的研究生生活变得丰富多彩。

感谢我的家人给予我的支持与理解,是你们的辛勤工作扛起了家庭负担,让我可以全身心投入到学习研究中,你们的期望是我最大的学习动力。

最后,衷心感谢于百忙之中评阅论文和参加论文答辩的专家老师!再次向所有给予我关心和帮助的老师、同学和朋友表示真心的感谢!