

南 京 理 工 大 学

毕业设计(论文)开题报告

学 生 姓 名: 李庆贺 学 号: 915106840425

专 业: 智能科学与技术

设计(论文)题目: 基于机器学习的微博突发事件分析与
谣言检测

指 导 教 师: 夏睿

2019 年 01 月 06 日

毕 业 设 计（论 文）开 题 报 告

1. 结合毕业设计（论文）课题情况，根据所查阅的文献资料，每人撰写 2000 字左右的文献综述：

随着微博等网络社交媒体的快速发展，关于微博谣言检测的研究工作越来越多。在研究早期，国外主要基于机器学习的经典分类方法进行研究。

Castillo 等[1]研究了 Twitter 话题可信度评估的方法，获取 Twitter 上与热点话题相关的数据集，从中提取微博文本、用户信息和传播特征，采用 J48 决策树分类学习方法来预测话题的可信性，实现了一定的效果。Qazvinian 等[2]通过分析 Twitter 帖子的浅层文本特征、行为特征和元素特征，构建了多个贝叶斯分类器和集成分类器以识别 Twitter 上的谣言。Mendoza 等[3]利用 Twitter 数据分析了突发事件下的用户发帖的行为特征，并追踪事件相关的谣言与新闻的传播特征，从而提出了微博转发形成的拓扑网络结构。Takahashi 等[4]针对具体的事件在 Twitter 上产生的谣言分析提出了基于爆发点、转发率和词分布差异等特征的谣言检测系统；相对于国外研究，国内针对谣言检测的相关研究起步较晚，早期研究较少。其中，程亮等[5]分析了微博谣言产生的原因及传播特点并利用 BP 神经网络模型并引入冲量项进行建模。许晓东等[6]从谣言传播角度对微博谣言进行了研究，采用计算机仿真技术发现微博谣言传播受有效传染率和网络分布熵影响。

然而上述方法均未较好地利用微博的文本属性特征。Yang 等[7]对微博进行了语料标注并基于 SVM 机器学习方法对微博进行了自动检测研究，但其未挖掘微博中隐含的深层特征信息，导致识别率并不高，因此贺刚等[8]在其研究基础上，基于 SVM 机器学习方法并引入了识别谣言的符号特征、链接特征、关键词分布特征和时间差等新特征来预测微博是否为谣言，获得了改善的效果。

近年来，随着自然语言处理以及深度学习的快速发展，国内外对于谣言检测的相关研究逐渐从谣言事件本身转向更细粒度的文本属性特征等方面，模型方法也开始采用能够获取更高层次特征水平的深度学习模型。

Friggeri 等[9]和 Hannak 等[10]利用谣言传播及级联特征结合事件的相关评论进行建模。Kwon 等[11]根据推文的数量推出了一种时间序列配置模型。Ma 等[12]在模型中加入了时间序列的社会背景特征。这些方法通常需要进行大量的数据预处理和特征工程。Zhao 等[13]通过使用一组正则表达式来挖掘用户对论文的态度。Ma 等[14]提出

使用循环神经网络模型来自动学习相关推文帖子上下文信息随时间的变化；最近他们[15]又提出了两种基于自下而上和自上而下的树状结构的递归神经网络模型用于谣言的表示学习和分类，通过整合结构特征和语义特征来生成更好的表示；同时 Ma 等[16]又首先提出基于多任务学习的统一方法将谣言检测和立场分类结合在一起的工作，该方法成功地学习为两个核心任务共同表示和分类数据。

此外，Wu 等[17]利用一些基于内核的方法来模拟传播结构，提出了一种混合 SVM 分类器，它结合了 RBF 内核和基于随机游走的图形内核，以捕获新浪微博上的传闻和传播模式。Ma 等[18]使用树核通过计算它们相似的子结构来捕获传播树的相似性，以便在 Twitter 上识别不同类型的谣言。最近，He 等[19]认为数据是长期演变和短期波动的，并首次尝试探索用于时间演化文本分类的进化神经网络模型，通过使用时间平滑框架将任意神经网络扩展到进化学习的简单方法，提出了一个历时传播框架，通过历时连接将历史影响纳入当前学习的特征。Yavary 等[20]将用户反馈视为谣言检测的重要特征，提出了用户会话树的模式和基于子树的模式，并将提取的模式用作极限学习机的特征，构建谣言检测模型。

值得注意的是，Guo 等[21]将谣言事件建模为包含不同语义信息水平的层次时间序列，提出了一种包含社会注意力机制的谣言检测的新型层次网络（HSA-BLSTM）；它将一个事件分为几个包含几个帖子的子事件。每个帖子进一步分为几个单词。然后将结构化事件输入分层 Bi-LSTM 网络，社会特征被用作识别谣言的突出部分的另一条线索并通过 Bi-LSTM 中的注意机制实现这些特征，获得谣言检测的准确表示。该模型不仅考虑了源微博及其评论转发的特征信息，也将他们的层次结构信息、社会特征等利用注意力机制来实现其表示。

参考文献:

- [1] Castillo C, Mendoza M, Poblete B. Information credibility on Twitter [C] // Proceedings of the 20th International Conference on World Wide Web. New York:ACL,2011:675-684.
- [2] Qazvinian V, Rosengren E, Radev D R, et al. Rumor has it: Identifying misinformation in microblogs [C] // Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. Edinburgh: ACL, 2011:1589-1599.
- [3] Mendoza M, Poblete B, Castillo C. Twitter under crisis: Can we trust what we RT ? [C] // Proceedings of the First Workshop on Social Media Analytics. New York: ACL, 2010: 71-79.
- [4] Takahashi T, Igata N. Rumor detection on Twitter [C] // 2012 Joint 6th International Conference on Soft Computing and Intelligent Systems(SCIS) and 13 th International Symposium on Advanced Intelligent Systems(ISIS). Kobe: IEEE, 2012: 452-457.
- [5] 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.
- [6] 许晓东, 肖银涛, 朱士瑞.微博社区的谣言传播仿真研究[J].计算机工程, 2011, 37(10): 272-274.
- [7] Yang Fan, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo [C] // Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. Beijing: ACM, 2012: 1-7.
- [8] 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.
- [9] Adrien Friggeri, Lada A Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor cascades. In Proceedings of ICWSM.
- [10] Aniko Hannak, Drew Margolin, Brian Keegan, and Ingmar Weber. 2014. Get back! you don't know me like that: The social mediation of fact checking interventions in twitter conversations. In Proceedings of ICWSM.
- [11] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang.2013. Prominent features of rumor propagation in online social media. In Proceedings of

ICDM. pages 1103–1108.

- [12] Jing Ma, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong. 2015. Detect rumors using time series of social context information on microblogging websites. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. CIKM '15, pages 1751–1754.
- [13] Zhe Zhao, Paul Resnick, and Qiaozhu Mei. 2015. Enquiring minds: Early detection of rumors in social media from enquiry posts. In Proceedings of the 24th International Conference on World Wide Web. WWW '15, pages 1395–1405.
- [14] Ma J, Gao W, Mitra P, et al. Detecting Rumors from Microblogs with Recurrent Neural Networks[C]//IJCAI. 2016: 3818-3824.
- [15] 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 (Volume 1: Long Papers). 2018, 1: 1980-1989.
- [16] Ma J, Gao W, Wong K F. Detect rumor and stance jointly by neural multi-task learning[C]//Companion of the The Web Conference 2018 on The Web Conference 2018. International World Wide Web Conferences Steering Committee, 2018: 585-593.
- [17] Ke Wu, Song Yang, and Kenny Q Zhu. 2015. False rumors detection on sina weibo by propagation structures. In Data Engineering (ICDE), 2015 IEEE 31st International Conference on. IEEE, pages 651–662.
- [18] Jing Ma, Wei Gao, and Kam-Fai Wong. 2017. Detect rumors in microblog posts using propagation structure via kernel learning. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). volume 1, pages 708–717.
- [19] He Y, Li J, Song Y, et al. Time-evolving Text Classification with Deep Neural Networks[C]//IJCAI. 2018: 2241-2247.
- [20] Yavary A, Sajedi H. Rumor detection on Twitter using extracted patterns from conversational tree[C]//2018 4th International Conference on Web Research (ICWR). IEEE, 2018: 78-85.
- [21] Guo H, Cao J, Zhang Y, et al. Rumor Detection with Hierarchical Social Attention

Network[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. ACM, 2018: 943-951.

毕 业 设 计（论 文）开 题 报 告

2. 本课题要研究或解决的问题和拟采用的研究手段（途径）：

一、课题研究的目的是和意义

随着在线社交媒体的快速发展，大量不可靠的信息得以快速和广泛地在人群中传播。社交媒体上谣言泛滥可能导致人们难以从纷繁的信息中甄别得到可信的信息，进而影响人们正常的生活秩序。研究微博谣言识别的方法，有助于用户判断信息的真假，营造健康的互联网环境，使微博在信息传播引导、舆情监控过程中起到积极作用。

目前在谣言检测技术方面，存在过度依赖人工特征、模型结构单一、突发检测成功率低等问题，因此构建有效的谣言检测模型来实现早期谣言检测具有重要意义。

二、课题整体解决方案及其研究思路

1. 第一阶段将针对课题所涉及的理论方法与原理进行系统全面的学习。本课题需要掌握机器学习、深度学习、表示学习和文本分类的相关原理及方法，在这一阶段将进行针对性的理论学习与知识补充。

2. 第二阶段将对研究课题进行调研。通过搜集并阅读国内外相关的研究论文，对课题整体的研究发展趋势有一定了解，并对模型方法的演化进行分析研究，从而整体上把握最新进展，选择并确定研究思路与模型。

3. 第三阶段将进行实际性的建模分析与研究，结合任务书逐步推进课题研究进程：

- （1）挖掘对谣言检测有效的特征模板，并在传统机器学习方法上构建有效的模型。
- （2）引入深度学习的方法，从序列标注、分类等不同角度解决谣言检测任务。
- （3）研究在深度学习模型下，如何有效地进行谣言早期识别及突发事件检测。
- （4）获取更多的微博语料资源，引入无监督学习的方法来完成谣言检测任务。

4. 第四阶段将不断完善模型和优化算法，提高模型的实用性。将对本课题所涉及的问题进行系统分析，按照任务书时间节点按时完成、定期与指导老师沟通，发现并改善模型存在的问题，提高模型的鲁棒性。

毕 业 设 计（论 文）开 题 报 告

指导教师意见：

1. 对“文献综述”的评语：

2. 对本课题的深度、广度及工作量的意见和对设计（论文）结果的预测：

指导教师：

年 月 日

所在专业审查意见：

负责人：

年 月 日