

南 京 理 工 大 学

毕业设计（论文）任务书

学 院：计算机科学与工程学院

专 业（方 向）：智能科学与技术

学 生 姓 名：李庆贺 学 号：915106840425

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

起 迄 日 期：2019 年 2 月 26 日 ~ 2019 年 5 月 18 日

设计(论文)地点：南京理工大学

指 导 教 师：夏睿 教授（计算机科学与工程学院）

专 业 负 责 人：任明武

发任务书日期：2018 年 12 月 7 日

任务书填写要求

1. 毕业设计（论文）任务书由指导教师根据课题的具体情况填写，经学生所在专业的负责人审查、学院领导发布后生效；

2. 任务书内填写的内容，必须和学生毕业设计（论文）完成的情况相一致，若有变更，应当经过所在专业及学院主管领导审批后方可修改；

3. 任务书内“主要参考文献”的填写，必须按照国标 GB 7714—2005《文后参考文献著录规则》的要求书写；

4. 如果主要在校内开展毕业设计（论文）工作，“地点”填写“南京理工大学”；主要在校外开展毕业设计（论文）工作，“地点”填写实际单位名称。

5. 有关年月日等日期的填写，应当按照国标 GB/T 7408—2005《数据元和交换格式、信息交换、日期和时间表示法》规定的要求，一律用阿拉伯数字书写。如“2014年3月15日”或“2014-03-15”。

毕 业 设 计（论 文）任 务 书

1. 本毕业设计（论文）课题应达到的目的：

基于机器学习的微博突发事件分析与谣言检测的主要任务是要对社交媒体的微博流数据进行自动的谣言识别并且尽可能早的探测到谣言的产生。本课程设计应达到的目的是：掌握文本处理的基础方法，并能够在谣言检测上有效运用分类、序列标注等深度学习方法，进一步完善突发事件和早期检测的深度学习模型。

2. 本毕业设计（论文）课题任务的内容和要求（包括原始数据、技术要求、工作要求等）：

1、查阅相关资料，完成开题报告。

2、完成外文翻译。

3、基于特征模板和常规深度学习方法，完成对社交媒体语料特点、任务特点的有效模型，并提高其效果。

4、基于突发事件和早期检测的要求，从深度学习神经网络方面进行探索，进行方法创新，提升早期检测的效果。

5、在更多的无标签数据下，从无监督领域探索新的方法来完成谣言检测任务。

毕 业 设 计（论 文）任 务 书

3. 对本毕业设计（论文）课题成果的要求〔包括毕业设计论文、图表、实物样品等〕：

- 1、构建基于机器学习的微博早期谣言检测系统。
- 2、构建基于无监督的有效谣言检测系统。
- 3、完成外文翻译、不少于两千字的开题报告、不少于一万字的毕设论文。

4. 主要参考文献：

- [1]Ma J, Gao W, Mitra P, et al. Detecting Rumors from Microblogs with Recurrent Neural Networks[C]//IJCAI. 2016: 3818-3824.
- [2]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.
- [3]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.
- [4]He Y, Li J, Song Y, et al. Time-evolving Text Classification with Deep Neural Networks[C]//IJCAI. 2018: 2241-2247.
- [5]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.
- [6]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.
- [7]Liu X, Nourbakhsh A, Li Q, et al. Real-time rumor debunking on twitter[C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM, 2015: 1867-1870.
- [8]Shu K, Sliva A, Wang S, et al. Fake news detection on social media: A data mining perspective[J]. ACM SIGKDD Explorations Newsletter, 2017, 19(1): 22-36.
- [9]Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews[C]//Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002: 417-424.
- [10]Liu L, Kang J, Yu J, et al. A comparative study on unsupervised feature selection methods for text clustering[C]//Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-KE'05. Proceedings of 2005 IEEE International Conference on. IEEE, 2005: 597-601.

毕 业 设 计（论 文）任 务 书

5. 本毕业设计（论文）课题工作进度计划（共 14 周）：	
起 迄 日 期	工 作 内 容
11-13 周	从无监督角度构建模型进行谣言检测
14 周	总结、结题
1-4 周	复现文献，构建特征模板和深度学习模型进行谣言检测
5-10 周	研究突发事件及早期检测模型构建，并有所改进
所在专业审查意见： 同意 负责人：任明武 2018 年 12 月 7 日	学院意见： 同意专业意见。 学院领导：陆建峰 2018 年 12 月 11 日