|  |  |  |  |
| --- | --- | --- | --- |
| **NUSTM研究组周报** | | | |
| **姓 名** | 李庆贺 | **日 期** | 2019/02/04– 2019/02/24 |
| **项目及内容** | 毕业设计 | | |
| **上阶段计划** | 寒假计划：  1基于前阶段实现的RNN模型来设计并实现早期谣言检测任务；  2.参照IJCAI2016的《Detecting Rumors from Microblogs with Recurrent Neural Networks》来实现早期谣言检测模型。 | | |
| **完成情况** | 1. 阅读了IJCAI2016的Detecting Rumors from Microblogs with Recurrent Neural Networks. 论文，对论文中的模型及方法进行了研究并结合自己实现的模型进行了比较，理清模型复现思路并初步实现了简单的早期谣言检测； 2. 针对Twitter数据集进行早期谣言检测模型进行测试并绘制Deadline-Accuracy图像，准确率达到83.3%； 3. 分析了模型的结构，并针对自己搭建的第一个模型进行分析优化改进方法，从数据预处理、分词方法、特征表达、网络模型等方面逐步改进，学习并采用了已有的成熟的方法及工具来提高各个环节的效率； 4. 在已实现的基于Twitter数据集、Bi-LSTM模型基础上，结合改进方法完成了基于Weibo中文数据集的GRU网络模型的搭建。 | | |
| **下周计划** | 将调试实现的GRU模型，并进行优化改进；  在GRU模型上进行早期谣言检测的实验。 | | |
| **收获/困难** | 假期里主要实现了自己动手搭建的双向LSTM模型，并简单的实现了早期谣言检测的要求，但结合结果来看效果并不是很好，通过分析，觉得是这个模型从数据预处理、分词方法、词向量的提取整合表达以及网络模型的搭建都是自己编写的代码，在实现上往往进行了简化处理，因此在后面的时间里主要分析改进的方式，并计划写第二版模型，在写第二板模型时，学习了很多相关的优秀的代码，并学习使用现在比较好的一些分词工具、机器学习工具等，并在编写代码时采用。  有了第一个模型的练习，在写第二个模型的代码的时候感觉就比较轻松了，遇到问题时也能更好地解决了。 | | |
| **参考文献** |  | | |
| **请假** | **请假时间** | **请假原因、地点** | |
|  |  | |
| **其他事项** |  | | |