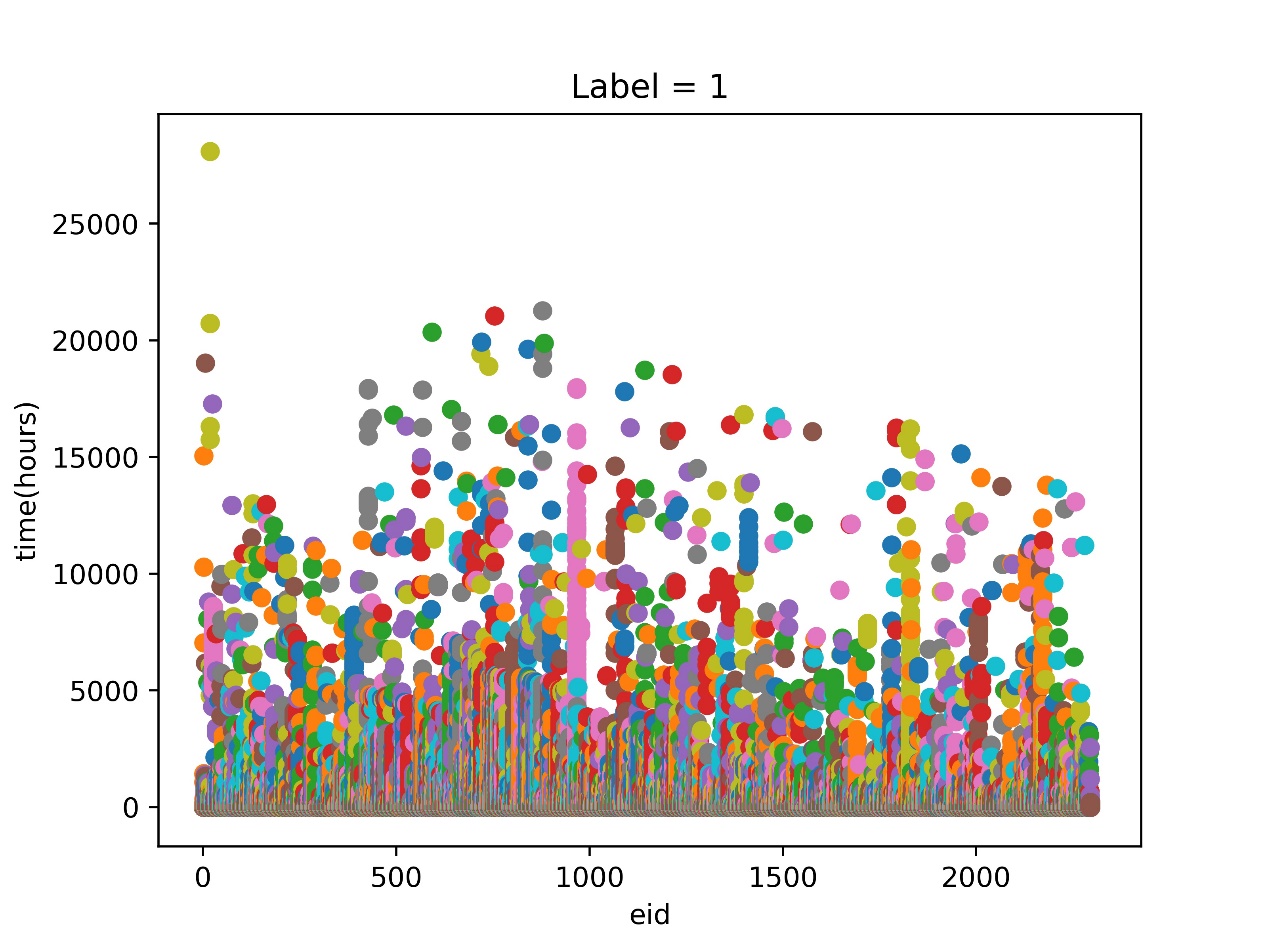
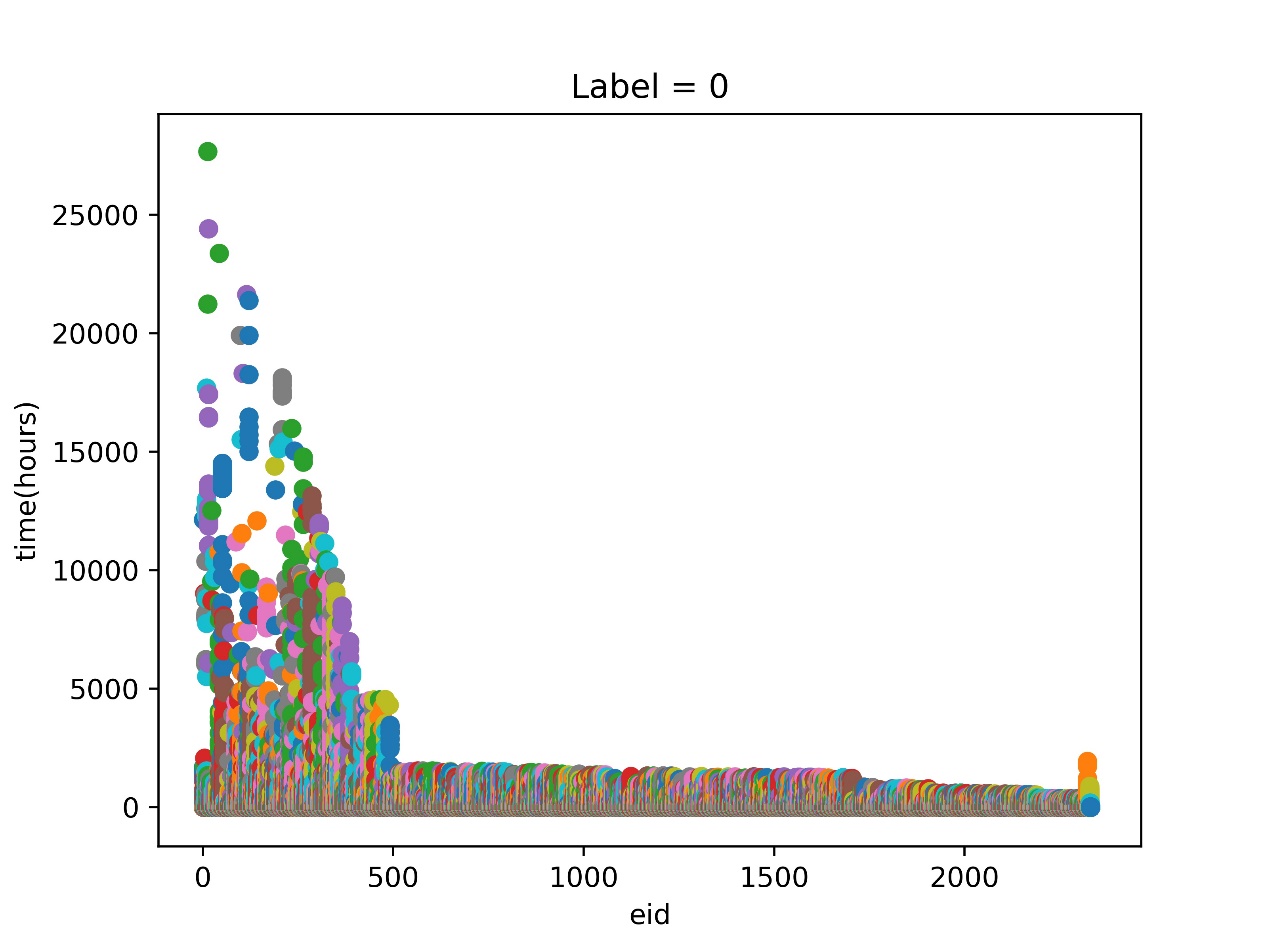
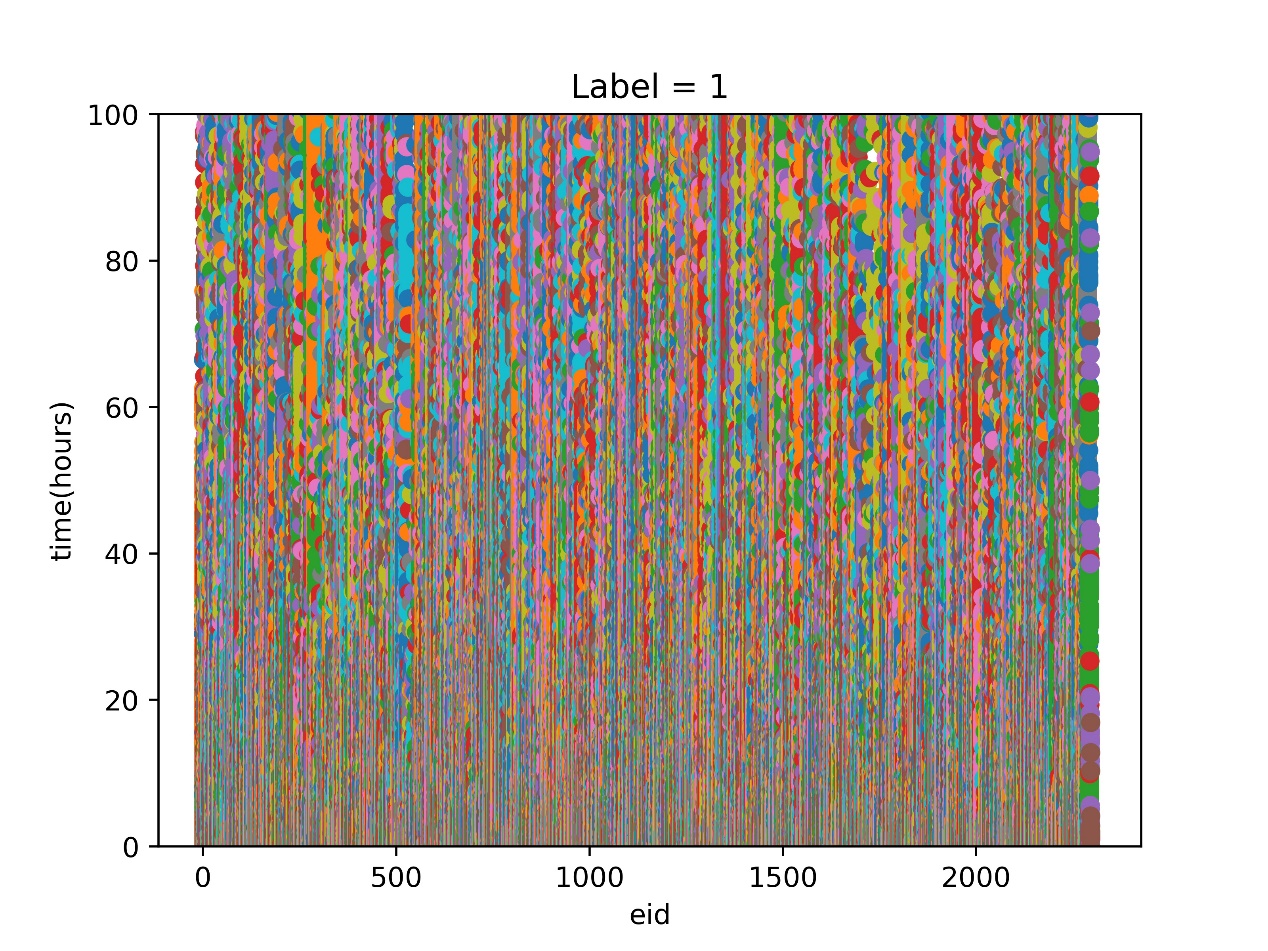
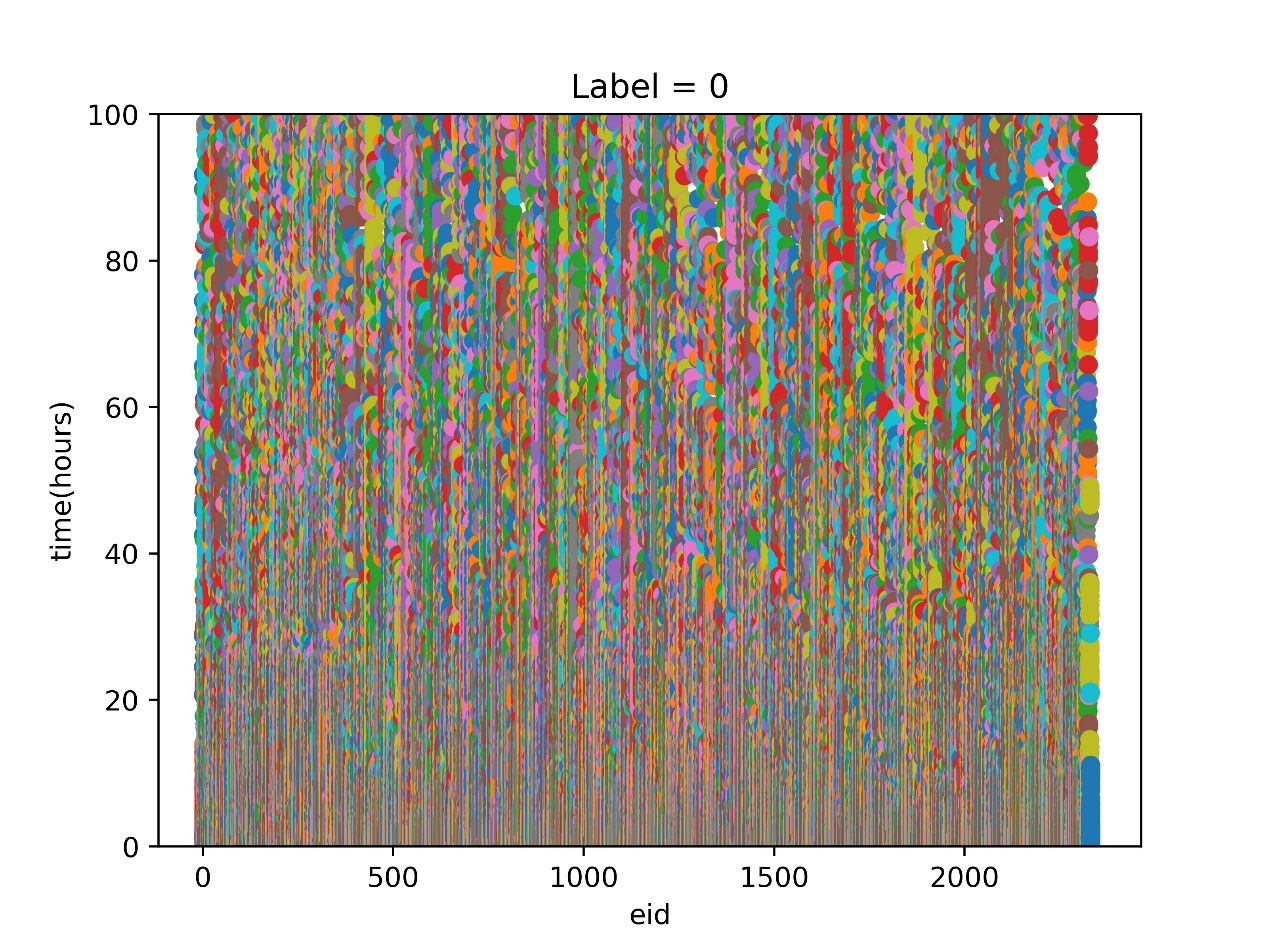
|  |  |  |  |
| --- | --- | --- | --- |
| **NUSTM研究组周报** | | | |
| **姓 名** | 李庆贺 | **日 期** | 2019/03/04– 2019/03/10 |
| **项目及内容** | 毕业设计 | | |
| **上阶段计划** | 在时间流方面对模型的改进方案进行探索；具体的计划为：  1.提取出数据集中每一条post的发布时间，并利用Excel对post的时间分布进行分析，寻找谣言与非谣言事件的发布时间特征；  2.查阅有关时间流方面的资料与论文，尝试结合分析结果获得改进模型的启发； | | |
| **完成情况** | 1. 编写程序提取出了Weibo数据集中的时间信息，并利用matplotlib绘图工具包将每个事件的发布时间分布绘图，分析结果发现谣言事件与非谣言事件的一些区别，这些差异可能会对以后的改进提供思路； 2. 根据学长提供的资料阅读了4篇概念漂移与检测的论文，希望将这一领域的方法与思路来引入的谣言检测中，其中一篇论文将进一步阅读研究； | | |
| **下周计划** | 1. 阅读《Learning under Concept Drift: an Overview》论文 2. 查阅流数据处理的相关资料与方法； 3. 对数据的时间特征进一步进行分析，并引入到模型中。 | | |
| **收获/困难** | 本周的主要任务是阅读时间流数据处理与概念漂移相关的论文，因为这方面感觉比较抽象，因此论文阅读起来比较难理解，对于如何将其引入到谣言检测中还是没有头绪；在对数据的时间分布特征的分析方面进展比较顺利，通过结构来看两类事件在时间分布上区别还是比较明显的，下一步就是要考虑如何将这差异带到模型中来，如何对差异进行表达，对于这方面的分析还是会继续深入。 | | |
| **参考文献** | 《Learning under Concept Drift: an Overview》  《UnderstandingConceptDrift》  《An overview of concept drift applications》  《Time Weight Collaborative Filtering》 | | |
| **请假** | **请假时间** | **请假原因、地点** | |
|  |  | |
| **其他事项** |  | | |

附：时间分布特征提取结果：

1. 事件在全部时间范围内的分布：



1. 事件在发布后100小时内的分布：



1. 事件在发布后1小时内的分布：

