**6月7日小组会议**

时间：2018年6月7日19:00-21:00

地点：行政楼208室

参会人员：邹乔莎 刘涛 周益 李安贞

1. 上周工作总结：
   1. 语义情感分析（周益、刘涛）

（1）语义情感分析的框架

（2）文本预处理——中文分词

（3）情感分析的两种方法：

a. 基于情感词典和规则

b. 利用神经网络学习

经过讨论，初步方案是先用a方法生成语义正负情感语料库，再用b方法学习到所有文本，从而判断所有微博的语义情感。

* 1. 用户行为分析（邹乔莎、李安贞）

（1）中文分词工具调研结果

（2）用户行为分析的流程图

（3）用提取关键词、统计词频的方式，给出了可能的用户标签

但这种方法受限于关键词的提取算法，并不能准确的给出明确的用户标签，讨论后，决定尝试文本分类的方式，对一个id下所有微博文本进行分类，再进一步给出用户标签。

1. 下周工作安排

周益：生成文本正负情感语料库

刘涛：学习CNN、LSTM

李安贞：对微博评论进行文本分类，形成4类（财经、娱乐、新闻、非上述3种），从而对用户进行画像