第一次平时练习要求

1. 完成时间：10月29日前。
2. 分组要求：每组最多8人。
3. 练习内容：做一次文本分析。内容包括但不限于：（1）改进课堂上的繁简转换办法。（2）设计课堂内容之外的分词方法（最大匹配、最小匹配、隐形马尔可夫链）。（3）利用课堂介绍的分词方法或其他一些开源分词工具，讨论文本分析的相关问题（如人名、地名、机构名的识别，未登录词的整理，文本相似度，等等）