

任务：阅读文献，大体了解关键词提取算法

周一：

在网络上通过搜索引擎查找相关博客对该算法有一个粗浅的认识，并将博客收藏以便之后的学习。

周二：

理清各种关键词提取算法之间的差别（有几个不是很理解），以及各算法适用的领域。为之后套用法律文本选择适用的几个算法（因为还没有完全弄清楚所以不是很确定该使用那个算法得到最优化的效果）。

周三：

阅读论文：基于语义的中文文本关键词提取算法

收获：通过词语间的语义距离来定词语间的语义相似度。基本了解了概念以及如何去定义。熟悉了 SKE 算法的逻辑（只是从字面了解了算法的流程，算法的具体实现及原理还在研读）。

周四：

阅读论文：应用于用户兴趣建模的多文本关键词抽取研究

收获：了解了多文本和单文本关键词提取的联系。

周五：

（在家准备周六的浙江省公务员考试）

总结：

通过这周 4 天的学习，大致了解了文本关键词提取的几种算法（不同模型下的不同算法）。同时就研究课题扩展阅读了多文本关键词提取的内容（觉得可能对今后完成课题有所帮助）。但是由于周末需要参加省考而分出了一部分精力，从而学习论文的内容并不是很理想。需要在接下来的时间里好好的加强学习。