**计算机科学与技术专业**

**本科毕业实习日志**

**时间**：2022年3月3日

**地点**：山东科技大学

**本日实习任务**：对获取数据进行进一步处理。

**完成情况：**对爬取到的数据统一整理为了txt文本格式后，利用Python-jieba对其进行了清洗，去除掉了非中文字符，并且对文本进行了切片，同时还对其进行了中文词汇的词性标注。

**总结与明日计划：**今日完成了数据初步处理，明日计划对文本进行词频统计。