**信息工程学院毕业设计课题申报审批表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **教师姓名** | | 仲冰 | **职 称** | 讲师 |
| **课题名称** | | 基于NLP的可视化评价信息收集与分析系统 | | |
| **课题类别** | | 毕业设计 | **课题性质** | 软件开发 |
| **课题来源** | | 自拟 | **课题组人数** | 1人 |
| **课**  **题**  **摘**  **要** | 通过爬虫程序收集商品信息、评价，新闻评论，微博评论等信息，然后对收集的数据进行清洗。然后使用NLP的方法对收集到的信息进行情感分析、分类、观点发掘等。然后对分析得到的结果进行可视化，从而有更为直观的体验和对比性。  对于商品信息的分析，其结果可以给有意向购买该商品的用户提供参考。  对于新闻评论和微博评论等，其结果可以用来观察舆论情况等等。 | | | |
| **课**  **题**  **依**  **据** | scrapy用于从网站中提取所需数据的开源协作框架，是一种快速、简单但可扩展收集信息的方式。  spaCy是一个先进的工业级别自然语言处理Python库。spaCy可以帮助实现分析大量文本，从而从收集的评价中进行观点挖掘等分析工作。  Matplotlib、Seaborn、Pyecharts等Python库都可以很好地实现数据的可视化工作。 | | | |
| **可**  **行**  **性** | 该系统主要通过使用爬虫程序收集数据，利用NLP相关的库对数据进行分析，再利用可视化技术对结果可视化，这些技术都已经较为成熟，因此具有技术可行性。 | | | |
| **专业系**  **意**  **见** | 专业系主任（签字）：  年 月 日 | | | |
| **院**  **部**  **意**  **见** | 院长（签字）：  年 月 日 | | | |